

Treg Depletion and/or Inhibition

Which clinical interventions reduce regulatory T cells (Tregs) — in absolute number, frequency, functional dominance, or suppressive capacity — in humans, and how durable / context-dependent is that effect?

Generated 2026-04-27

Last updated: 2026-04-24. Papers reviewed: 44.

[← back to shieldbreak](#) · [← back to shieldbreaks](#)

Top-line findings

<details class="sb-section" open markdown="1"> <summary>Show / hide</summary>

- 44 of 44 papers critiqued. 34 had full text accessible (27 full PMC XML + 6 PMC-paywalled via WebFetched HTML + 1 combined); 10 were PubMed-abstract-only (paywalled journals such as J Immunother, Clin Exp Immunol, Int J Cancer, Inflamm Bowel Dis, Br J Dermatol, Cancer Discov, Clin Cancer Res subscription-only).
- No retractions, corrections, or expressions of concern found. One extraction-fidelity discrepancy flagged (Huang 2011 — Mann-Whitney label applied to paired pre/post biopsy data; search correctly noted the inconsistency).
- The central methodological lesson: Treg-definition heterogeneity and gating-confounding drive much of the apparent inter-paper inconsistency. CD25-based gating during CD25-targeting therapy (denileukin diftitox, daclizumab) structurally confounds depletion claims; single-marker FOXP3 IHC cannot distinguish Tregs from activated effector T cells that transiently express FOXP3.

</details>

Summary stats

<details class="sb-section" open markdown="1"> <summary>Show / hide</summary>

- **Papers critiqued:**44
- **Extraction-fidelity discrepancies flagged:**1
- **Retractions / corrections / expressions of concern:**0

Risk-of-bias distribution

Rating	N
Low	3
Moderate	32
Serious	4
Some concerns	4
not amenable — case report	1

Per-trial confidence distribution

Confidence	N
High	2
Low	18
Moderate	23
Very low	1

Counter-productive-mechanism severity distribution

Severity	N
High	15
Moderate	19
Low	10

</details>

Cross-paper synthesis

<details class="sb-section" open markdown="1"> <summary>Show / hide</summary>

Direction of effect across intervention classes

- Mogamulizumab (anti-CCR4): large, on-target eTreg depletion (~90% PBMC, ~87% tumor) in CTCL/ATLL. Fujikawa 2023 + Jinushi 2025 + Roelens 2022 are mutually corroborative. HIGH confidence that CCR4+ effector-Tregs are depleted; broader FOXP3+ population partially spared.
- Denileukin diftitox (anti-CD25 IL-2 fusion): MIXED — Dannull 2005 (RCC, n=7, positive), Mahnke 2007 (melanoma, positive), Geskin 2018 (CTCL, positive), Liao 2024 (ovarian, positive with dose-dependence) vs Attia 2005 (melanoma, NEGATIVE) and Thibodeaux 2021 (phase II failure). Reconciling variables: (a) CD25-gating confound in positive studies, (b) indication, (c) dose schedule. Attia 2005's use of FoxP3 mRNA readout (sidestepping CD25-gating) is the most methodologically defensible negative.
- Standard (IgG1 ipilimumab, IgG2 tremelimumab) anti-CTLA-4: DOES NOT DEplete INTRATUMORAL TREGS. Sharma 2019 (N=45 across 3 cancers) + Huang 2011 (n=19 tremelimumab melanoma, 4.75× INCREASE p=0.0029) + Ribas 2009 (n=7 retrospective, slight increase) + Comin-Anduix 2008 + Penter 2023 (ipilimumab after decitabine AML, EXPANDED marrow Tregs) all converge. Preclinical ADCC-depletion mechanism does NOT translate to humans with non-Fc-engineered antibodies.
- Fc-enhanced anti-CTLA-4: DEPLETES INTRATUMORAL TREGS. Ager 2026 NeoRED-P (BMS-986218 afucosylated, randomized phase I, p=0.031) + Chand 2024 (botensilimab + balstilimab in MSS CRC). The Fc-engineering-reconciling-variable hypothesis is supported by the contrast: enhanced CD16a/FCGR3A affinity enables ADCP/ADCC by tumor macrophages.
- Non- α IL-2 variants (nemvaleukin, bempedaldesleukin): EXPAND Tregs absolutely but expand effectors more (ratio-shift mechanism). Bentebibel 2019 + Diab 2020 initial abstract claims of 'without Treg enhancement' subsequently contradicted by Gogas 2024 (same PIVOT-02 cohort, 8-10× absolute Treg expansion documented). Vaishampayan 2024 + Calvo 2025 + Piha-Paul 2025 all confirm the absolute-expansion / ratio-shift pattern. CRITICAL CLINICAL CONTEXT: the bempedaldesleukin phase 3 PIVOT-IO program FAILED (2022) — the ratio-shift mechanism did not translate to clinical benefit.
- Cyclophosphamide (depletion setting): Metronomic oral dosing (Ghiringhelli 2007) DEPLETES Tregs 61% p<0.0001; single IV dose (Audia 2007) does NOT. Schedule-dependent, not pan-agent effect.
- HDACi: Highly specificity- and context-dependent. Class-I selective (entinostat: Pili 2017 RCC; vorinostat: Terranova-Barberio 2020 breast) decrease tumor Tregs. Broad-spectrum panobinostat INCREASES Tregs in HIV cART (Brinkmann 2018, +40% p=0.003) but REDUCES TNFR2+ Treg SUBSET in AML (Govindaraj 2014). No simple class effect.
- DNMTi / HMAs: Generally INCREASE Tregs via TSDR demethylation. Han 2021 ITP (therapeutic direction for autoimmunity), Penter 2023 AML/MDS (post-decitabine ipilimumab EXPANDS marrow Tregs — interpreted as resistance mechanism). Opposite direction to the scope-expansion hypothesis.
- CELMoDs (iberdomide): INCREASE Tregs. Lipsky 2022 SLE randomized placebo-controlled, +104.9% Tregs at 0.45 mg p<0.001. Amatangelo 2024 RRMM consistent. Iberdomide degrades Ikaros/Aiolos (not Helios/IKZF2) — preserves Treg identity.
- Anti-TIGIT, anti-GITR, anti-OX40, anti-TNFR2, EP4, NRP1, CCR8-CAR-T: Limited human PD data. Anti-GITR programs (TRX518, MK-4166) did not advance clinically despite modest PD signals. Anti-TIGIT shows PBMC-only Treg reduction (Guan 2024) without clear tumor-compartment effect; broader program's phase 3 results mixed. Anti-TNFR2 has no published clinical Treg PD data as of 2026-04-23. IKZF2-selective degraders

and CCR8-CAR-T remain preclinical.

Gating-confounding patterns

- Classic confound: CD25-based Treg gating during CD25-targeting therapy (denileukin diftitox, daclizumab). Present in Dannull 2005, Geskin 2018, Mahnke 2007, most other DD studies. Attia 2005 and Liao 2024 sidestepped this by measuring FoxP3 mRNA; Morse 2008 partially addressed with intracellular FoxP3 + explicit CD25-half-life argument. Papers that ONLY use CD25 gating during DD should have depletion claims discounted.
- Single-marker FOXP3 IHC limitation: Sharma 2019 explicitly notes that FOXP3 can be expressed on activated T effector cells as well as Tregs. Huang 2011 used the same single-marker IHC. CyTOF orthogonal validation (Sharma 2019, Ager 2026) is the methodological fix and all papers that use only FOXP3 IHC without co-staining are potentially overcounting Tregs.
- Frequency vs absolute count framing: Sharma 2019 explicitly notes that prior 'Treg reduction' claims (e.g., Liakou 2008 bladder) were driven by FREQUENCY-AS-PERCENT-OF-TOTAL-T calculations — total T cells increased, so Treg frequency decreased even though absolute count did NOT. The Diab 2020 PIVOT-02 'without Treg enhancement' claim is a similar artifact (frequency vs absolute count).

Heterogeneity sources

- Compartment: PBMC, tumor, tumor-draining LN, ascites, bone marrow, and skin often move in opposite directions with the same drug. Examples: Ager 2026 (tumor Treg ↓, tdLN Treg ↑ $p < 0.0001$); Terranova-Barberio 2020 (tumor Treg ↓, PBMC Treg unchanged); Guan 2024 (PBMC Treg ↓, tumor ambiguous). Compartment-stratified interpretation is essential.
- Disease indication: DD positive in CTCL/RCC/ovarian, negative in melanoma. Anti-CTLA-4 non-depleting across all tested cancers (melanoma, prostate, bladder).
- Dose and schedule: Cyclophosphamide metronomic vs single-IV is the clearest example. DD repeated vs single-dose (Morse 2008) also shows schedule effects.
- Gating: Single-marker vs co-stain vs surface vs intracellular vs mRNA vs scRNA-seq cluster all matter.

Counter-productive-mechanism patterns across interventions

- Severity distribution across the 44 papers: 15 High, 19 Moderate, 10 Low, 0 Unknown. Every paper carries some counter-productive exposure for the shieldbreak's generalizable 'remove immunosuppression' goal — the differences are in size, directness, and whether the mechanism is bounded by design.
- High-severity / opposite-direction-mechanism cluster (n=6): CELMoD iberdomide (Amatangelo 2024 RRMM, Lipsky 2022 SLE) INCREASES Tregs by Ikaros/Aiolos-sparing-Helios design; DNMTi monotherapy (Han 2021 ITP, Penter 2023 decitabine+ipi AML) INDUCES Tregs via TSDR demethylation [Moon 2016, Clin Cancer Res]; panobinostat HIV reservoir reactivation (Brinkmann 2018) increases Tregs 40%. These are counter-productive by design for a Treg-depletion shieldbreak — useful in their actual indications but directionally opposite.
- High-severity / depletes-beneficial-effectors cluster (anti-CCR4; n=4): Mogamulizumab (Fujikawa 2023, Jinushi

2025, Roelens 2022, Gordon 2025) depletes CCR4+ central memory CD8 T cells alongside Tregs. Fujikawa 2023 is paper-internally explicit: 'KW-0761 may concurrently deplete eTregs and central memory CD8+ T cells, and this dual depletion may cancel anti-tumor immune responses' [Tanaka 2021, Nat Commun]. Grade 3-4 lymphopenia in ~25% across studies is consistent. This is the best-evidenced counter-productive mechanism in the set.

- High-severity / treg-expansion-by-design (non- α IL-2 variants; n=5): Bempegaldesleukin and nemvaleukin 'preferentially activate effectors over Tregs' is a denominator-artifact framing — both expand Tregs absolutely, bempeg by 8-10 \times (Gogas 2024 reanalysis of PIVOT-02 which had been reported as 'without Treg enhancement' in the original Bentebibel 2019 / Diab 2020 reports). The bempegaldesleukin PIVOT-IO-001 phase 3 clinical failure (2022) is the hardest clinical evidence in the entire set that a ratio-shift mechanism can fail; nemvaleukin ARTISTRY-7 phase 3 failure (2023, ovarian) is a parallel.
- High-severity / IL-2-receptor-blockade-collateral (anti-CD25 daclizumab; n=2): Daclizumab (Mahnke 2007, Morse 2008) blocks IL-2R α on Tregs AND on CD25-upregulating activated effector CD8 T cells — the textbook counter-productive mechanism [Rech & Vonderheide 2009, Ann NY Acad Sci]. Morse 2008 addresses this with a pharmacokinetic-decay argument but does not measure effector IL-2 signaling directly.
- Moderate-severity / CD25-effector-collateral (denileukin diftitox; n=8): DD depletes CD25-high cells, which include transiently activated CD25+ effector T cells [FDA ONTAK label]. The CD25-gating confound in most DD papers also means the readout cannot distinguish off-target effector depletion from the intended Treg depletion. IP route (Liao 2024) bounds the systemic-exposure concern and is the lowest-severity DD study in the set.
- Moderate-severity / ADCC-activated-effector-collateral (Fc-enhanced anti-CTLA-4, anti-GITR, anti-TIGIT; n=5): Fc-engineering increases Fc γ R binding and ADCC, which can deplete activated effector T cells that transiently upregulate the target marker [Simpson 2013 JEM; Arce Vargas 2018 Cancer Cell; Chauvin & Zarour 2020 JITC]. Ager 2026 provides paper-internal data that CTLA-4 protein was 'exquisitely confined' to Tregs, bounding the concern for BMS-986218 specifically — but the same paper shows tdLN Treg EXPANSION (p<0.0001), a compartment-dissociation counter-productive signal.
- Low-severity / non-depleting standard anti-CTLA-4 (n=6): Ipilimumab and tremelimumab do not deplete intratumoral Tregs in humans despite preclinical ADCC claims (Sharma 2019 + Huang 2011 + Ribas 2009 + Comin-Anduix 2008 + Hamid 2011 + Yi 2017 converge). Because the proximal depletion mechanism does not operate, collateral concerns on effectors are moot at that step. The remaining class concerns are alt-checkpoint (PD-1/LAG-3) compensatory upregulation [Huang 2017, PNAS] and paradoxical autoimmunity (Nancey 2012 ipi-induced enterocolitis case report) — both bounded.
- Low-severity / bounded-design interventions (n=5): Metronomic cyclophosphamide (Ghiringhelli 2007) is the cleanest dose-selective Treg-depletion mechanism in the set; class-I-selective HDACi (entinostat: Pili 2017, Roussos Torres 2021) preserves effector CD8 function in preclinical work [McCaw 2021, Cancer Med]; IP denileukin diftitox (Liao 2024) bounds systemic exposure.
- Implication for the shieldbreak signal: the classes with the cleanest Treg-depletion endpoints (anti-CCR4, non- α IL-2) also carry the best-evidenced counter-productive mechanisms. The classes with the weakest depletion signals (standard anti-CTLA-4) have the least counter-productive exposure precisely because they don't deplete. Interventions that succeed at the proximal endpoint WITHOUT material counter-productive exposure are rare in this set; metronomic cyclophosphamide (Ghiringhelli 2007) and IP denileukin diftitox (Liao 2024) come closest.

Functional vs numerical endpoints

- Most papers report count/frequency only; few include functional suppression assays. Ghiringhelli 2007 includes NK cytotoxicity and T-cell proliferation restoration (supportive of functional reduction); Attia 2005 explicitly

reports $\geq 50\%$ SUPPRESSION PRESERVED despite minimal numerical change (STRONG evidence against DD depletion claim); Dannull 2005 reports 'abrogation' of suppressive activity (positive). The discrepancy between numerical and functional readouts in the DD literature is itself a signal that the field has not converged on a valid readout.

- For interventions where the expected mechanism is functional impairment rather than count depletion (IL-2 variants, anti-CTLA-4, EZH2i in theory), suppression-assay evidence is the appropriate standard. Few papers in this shieldbreak meet that standard.

Clinical-translation record

- Bempegaldesleukin: PIVOT-02 phase 1 positive signal DID NOT translate to phase 3 success (PIVOT-IO program failed 2022 melanoma + RCC; bempeg development halted). The ratio-shift mechanism as a basis for clinical benefit has been substantially disconfirmed.
- Anti-GITR (TRX518, MK-4166, MK-1248): Modest PD signals; programs have not advanced.
- Denileukin diftitox: Original formulation withdrawn; new formulation (Lymphir / I/ONTAK) recently FDA-approved in CTCL only. Broader-indication efficacy remains limited.
- Anti-CTLA-4: ipilimumab succeeded clinically BUT NOT via Treg depletion. The clinical benefit must be attributed to effector-cell potentiation rather than Treg removal, given converging Sharma 2019 / Huang 2011 / Penter 2023 evidence.
- Fc-enhanced anti-CTLA-4 (botensilimab, BMS-986218): Early signals (Chand 2024, Ager 2026) are encouraging but await randomized phase 3 confirmation.
- Mogamulizumab: Clinically validated in CTCL/ATLL; the on-target CCR4+ eTreg depletion is mechanistically cleanest but the clinical benefit observed in non-CTCL settings has been modest.
- HDACi class-I + ICI: Roussos Torres 2021, Pili 2017, Terranova-Barberio 2020 — all phase 1/2 studies; no definitive phase 3 validation yet.

What would change the picture

- Randomized phase 3 of Fc-enhanced anti-CTLA-4 (botensilimab or BMS-986218) with pre-specified tumor Treg PD would confirm or refute the Fc-engineering reconciling-variable hypothesis.
- Functional suppression-assay readouts in more DD and IL-2-variant studies would clarify whether observed count changes translate to functional benefit.
- Standardized multi-marker Treg gating (FoxP3 intracellular + CD127^{low} + CD45RA⁻ + Helios) across studies would eliminate much of the single-marker-vs-multi-marker noise.
- Compartment-stratified analysis as default (PBMC AND tumor AND tdLN in the same trial) would reveal the within-patient compartment dissociation that single-compartment papers miss.
- A replication of Dannull 2005 (JCI positive) with FoxP3-mRNA-primary readout in RCC would resolve the Attia-Dannull conflict.

</details>

Per-paper critiques

<details class="sb-section" open markdown="1"> <summary>Show / hide</summary>

Ager CR 2026 — Neoadjuvant Fc-enhanced anti-CTLA-4 targets Tregs to augment androgen deprivation in high-risk prostate cancer: A randomized phase I trial

Citation. PMID 41759531 · PMCID PMC13006393 · DOI 10.1016/j.xcrm.2026.102638 · NCT NCT04504669 **Source tier.pmc_full_text**

Key critique. A rigorous randomized phase I that validates the Fc-engineering-reconciling-variable hypothesis for anti-CTLA-4 Treg depletion in tumor: BMS-986218 reduces TI-Tregs ($p=0.031$) while standard IgG1 ipilimumab did not (per Sharma 2019). Small n ($n=10+14$) and marginal p -value cap the strength of inference, but orthogonal scRNA/CyTOF/mIF validation is a design strength, and the simultaneous tdLN Treg INCREASE ($p<0.0001$) is honestly disclosed.

Design. Two-arm open-label randomized phase I window-of-opportunity trial (NeoRED-P, NCT04301414) enrolling 24 men with high-risk localized prostate cancer; 4-patient safety lead-in then 1:1 randomization (ADT-only vs ADT + anti-CTLA4-NF [BMS-986218, afucosylated ipilimumab]). Design is a window-of-opportunity trial with radical prostatectomy as the tissue-collection anchor. Open-label (no blinding of clinicians or pathologists), unblinded outcome assessment is appropriate for a mechanism-endpoint PD trial but introduces potential for bias in subjective endpoints. TI-Treg frequency was a pre-specified mechanistic endpoint. **Sample size & power.** $n=24$ randomized, with 10 evaluable in ADT-only arm and 14 in ADT+NF arm (includes 4 safety lead-in). For the primary mechanistic comparison of TI-Treg reduction, the paper reports $p=0.031$ — a statistically significant but underpowered comparison given the small arms. No formal power calculation is described for the Treg endpoint, consistent with a phase I mechanism-hypothesis trial. The imbalance in Gleason Grade 5 disease (50% in ADT+NF vs 20% in ADT alone; standardized mean difference in Table S1) raises concern that randomization did not fully balance prognostic features. **Effect-size calibration.** The TI-Treg reduction ($p=0.031$ ADT+NF vs ADT) is modest in magnitude (frequency delta not pulled as clean percent in the literature search row; figure-dominant readout) but mechanistically coherent: CD16a/FCGR3A protein activity on tumor macrophages correlates with depth of Treg reduction ($p=0.033$), consistent with the Fc-ADCP hypothesis. The finding is also internally validated by scRNA-seq, CyTOF, and mIF — three orthogonal assays agreeing. Counterweight: the paper reports an INCREASE in Treg frequency in tumor-draining lymph nodes ($p<0.0001$), which the authors acknowledge as a known class effect of anti-CTLA-4 therapeutics. The tumor/tdLN compartment dissociation is biologically important and argues against simple 'anti-CTLA-4 depletes Tregs' framing even for Fc-enhanced agents. **Missing data.** One ADT+NF patient did not undergo prostatectomy (COVID-era OR closure); one received adjuvant therapy and was excluded from post-surgical outcomes. Not LOCF; clearly documented in CONSORT. No other meaningful attrition. **Multiplicity.** Multiple exploratory PD endpoints (TI-Treg, DC activation, T-cell priming, recurrence stratification). Paper does not report multiplicity correction. For a phase I exploratory trial this is conventional but the $p=0.031$ primary Treg readout is only marginally significant and would not survive Bonferroni across the ~6-10

hypothesis tests reported. **Generalizability.**High-risk localized prostate cancer only (Grade Group ≥ 3 , VHR $\geq 80\%$ per arm); Columbia-single-center; neoadjuvant setting with defined tissue-collection anchor. Findings are not directly generalizable to metastatic CRPC (where Tregs in bone-marrow microenvironment differ) or to non-prostate indications. One TMB-H patient via germline MSH2 may be an outlier. **Treg definition / gating.** Orthogonal: scRNA-seq-based TI-Treg population identification (transcriptomic cluster); CyTOF FoxP3+CD4+ gating; mIF spatial FoxP3+ density. This is the gold standard for this class — no CD25-gating confounding. Tregs were not CD25-gated; since anti-CTLA-4 does not target CD25, no gating-confounding issue arises.

****Counter-productive mechanisms (Moderate (compartment-dissociation, tdln-treg-expansion, adcc-activated-effector-collateral)).Fc-enhanced anti-CTLA-4 (BMS-986218, afucosylated ipilimumab) augments ADCC/ADCP on CTLA-4-high cells; the paper itself argues (Figure 4I–O) that CTLA-4 protein expression was 'exquisitely confined' to the Treg compartment, bounding the collateral depletion concern for other effectors in the tumor. However, the same paper documents a compartment dissociation that is counter-productive in kind: Tregs INCREASE in tumor-draining lymph nodes ($p < 0.0001$) in the ADT+NF arm, a known class effect the authors acknowledge [paper Results, Figure 4]. The tdLN is the compartment where anti-tumor CD8 priming happens, so tdLN Treg expansion plausibly offsets the tumor-Treg reduction the drug accomplishes. Fc-engineering also risks depleting CTLA-4-upregulating activated CD8 effectors in some tumor contexts [external: Arce Vargas 2018, Cancer Cell]. Severity is moderate — the tumor-side mechanism is clean but the tdLN signal is measured and unaddressed by the intervention. COI & funding.BMS-986218 is a Bristol Myers Squibb compound. First author (Ager) moved to J&J post-study. Funding and COI to verify in full-text Acknowledgments (not cleanly extracted from the XML header I read). Industry sponsorship and the presence of BMS-associated authors means results should be read as supporting an FC-enhanced-anti-CTLA-4 program; Columbia and Mayo investigators are primarily academic. No apparent spin in the primary Treg readout; feasibility framing is honest about limitations. Spin / abstract-to-text consistency.**** Abstract and results are consistent. The abstract claims 'Fc-enhanced α CTLA-4 exposure associates with reduced tumor Treg frequencies' — matches $p=0.031$ finding. The paper does not over-claim efficacy; it is explicitly framed as feasibility / mechanism / 'correlates with improved clinical outcomes' (small n caveat stated). The tdLN Treg INCREASE ($p < 0.0001$) is disclosed in the results and cited to prior literature — honest framing.

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.RoB 2: **Some concerns**— Open-label randomization; primary Treg endpoint analyzed in small subgroups; some baseline imbalance (Gleason GG5) not fully corrected; otherwise rigorous design with orthogonal multi-assay validation. **Per-trial confidence.Moderate**— Orthogonal-assay validated mechanism signal in a randomized (if small) phase I; $p=0.031$ is marginal and would benefit from replication in a larger trial.

Amatangelo M 2024 — Iberdomide Plus Dexamethasone Immune-related Effects in Patients With Relapsed or Refractory Multiple Myeloma (pharmacodynamic translational substudy)

Citation.PMID 38776914 · PMID PMC11228401 · DOI 10.1158/1078-0432.CCR-23-2987 · NCT NCT02773030
Source tier.pmc_full_text

Key critique.A CELMoD counterexample to the hypothesis that Ikaros-family degraders destabilize Tregs. Iberdomide degrades Ikaros/Aiolos (not Helios/IKZF2), and this paper shows immune modulation consistent with that mechanism rather than Treg destabilization. Translational substudy scope and small cohort size cap inference strength.

Design.Translational substudy of the CC-220-MM-001 Phase 1b/2 trial of iberdomide + dexamethasone in RRMM. Single-arm design with paired pre/post PD analyses on PBMC. Treg measurement is an exploratory correlative endpoint, not pre-specified for regulatory purposes. **Sample size & power.**PD cohort size is small-to-moderate (depends on dose cohort; specific n per cohort varies). No formal Treg-endpoint power calculation. For a translational substudy this is typical but caps the strength of any negative finding. **Effect-size calibration.** Iberdomide primarily degrades Ikaros/Aiolos (IKZF1/3), NOT Helios/IKZF2. The observed Treg increase is consistent with the Lipsky 2022 SLE iberdomide data (same compound). The literature search correctly classified the mechanism as Treg-modulating (not depleting). This aligns with the broader CELMoD signal — these compounds are not Treg-destabilizers in the IKZF2/Helios sense that has been preclinically pursued. **Missing data.**Standard RRMM translational study — some patients progress or die before later timepoints, typical attrition for this setting. Not explicitly discussed as a bias source. **Multiplicity.**Many immune-cell subsets analyzed simultaneously (Treg, Tcon, CD8, NK, etc.); multiplicity correction not reported. Exploratory framing. **Generalizability.**Heavily pretreated RRMM patients; findings may not generalize to newly diagnosed MM or non-MM indications. **Treg definition / gating.**Standard CD4+CD25+FOXP3+CD127low flow-cytometry gating (to be confirmed from methods). No obvious gating confounding with iberdomide (which does not target CD25 or FOXP3 directly). ****Counter-productive mechanisms (High (opposite-direction-mechanism, treg-induction-by-design)).****Iberdomide is a cereblon modulator that degrades Ikaros (IKZF1) and Aiolos (IKZF3) but spares Helios (IKZF2) and Eos (IKZF4), preserving Treg-lineage TFs. The paper's own PD data demonstrate that iberdomide INCREASES Tregs alongside effectors in R/R myeloma, the opposite direction from a Treg-depletion hypothesis. The counter-productive mechanism here is the mechanism itself for this shieldbreak's goal: the drug is net-suppressive of the Treg-depletion endpoint. Clinically the CELMoD net effect in MM is driven by direct myeloma cytotoxicity and effector activation, not Treg reduction — which is why the drug is a useful therapeutic despite opposing the shieldbreak's proximal mechanism. Severity is high for the shieldbreak question, regardless of its clinical utility in MM. COI & funding.****BMS/Celgene industry-sponsored substudy; multiple authors are BMS employees (Amatangelo, others). Iberdomide is a BMS compound. Findings of Treg-related immune effects should be read in this context, though the direction (not a depletion claim) makes spin less likely.** Spin / abstract-to-text consistency.** Paper does not claim Treg depletion — it reports on the full immune-modulation profile. No spin issue identified in abstract vs extracted row.

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Single-arm translational substudy with paired pre/post design; exploratory endpoint; no control arm; industry sponsorship and authorship present. **Per-trial confidence.****Moderate**— Reliable pharmacodynamic readout but exploratory endpoint in a single-arm study limits strength of causal inference.

Atchison E 2010 — A pilot study of denileukin diftitox (DD) in combination with high-dose interleukin-2 (IL-2) for patients with metastatic renal cell carcinoma (RCC)

Citation. PMID 20664355 · DOI 10.1097/CJI.0b013e3181ed0c20 **Source tier.** pubmed_abstract

Key critique. Small single-arm pilot (n=18) using historical controls; reports median 56.3% Treg reduction with DD+HD IL-2 (p=0.013) in RCC. Paywall limits full methodological appraisal; the classic CD25-gating confound for DD studies cannot be ruled out without full text. The 33% response rate is not distinguishable from HD IL-2 monotherapy benchmarks.

Design. Single-arm pilot study with dose-cohort arms (A: DD 6 µg/kg between IL-2 courses, n=3 safety-only; B: DD 9 µg/kg before IL-2, n=9; C: DD 9 µg/kg between IL-2 courses, n=6). 15 historical IL-2-monotherapy patients used as controls for toxicity and 13 for flow-cytometry comparison. Historical-control design is a major limitation: temporal confounding and selection bias in retrospective control selection. **Sample size & power.** n=18 treated + n=13-15 historical controls. Small sample with historical controls. No formal power calculation described; median 56.3% Treg decline from pre-DD to post-DD with p=0.013 is reported but arm-level n is only 15 (groups B+C pooled). **Effect-size calibration.** Median 56.3% Treg reduction (p=0.013) is consistent with Geskin 2018 and Dannull 2005 DD data in other indications — but specifically contrasts with Attia 2005 (DD failed to deplete Tregs in melanoma). The reason for this discrepancy is not resolved but likely reflects gating differences and cohort selection. The overall response rate of 33% (5/15) is high for RCC and may reflect the IL-2 component more than DD. **Missing data.** Not assessable from abstract only. Paywalled journal (J Immunother) — paper was not accessible for full-text critique. **Multiplicity.** Not assessable from abstract only. **Generalizability.** Metastatic RCC only; HD IL-2-eligible patient population (highly selected for fitness, excludes organ dysfunction). Limited generalizability. **Treg definition / gating.** Not described in abstract. CD25-based gating would be particularly problematic here because DD targets the CD25 (IL-2R α) component of the high-affinity IL-2 receptor — the intervention deliberately depletes CD25+ cells. This is THE classic gating-confounding issue for DD Treg studies. Full-text review required to rule out. ****Counter-productive mechanisms (Moderate (cd25-effector-collateral, il2-treg-expansion-confound)).** **Denileukin diftitox targets CD25-high cells via a diphtheria toxin fusion. The counter-productive concern is twofold: (i) activated CD8 effector T cells transiently express CD25 and can be co-depleted [external: Baur et al. 2013, FDA ONTAK label]; (ii) co-administered high-dose IL-2 is known to expand Tregs as well as effectors [external: Ahmadzadeh & Rosenberg 2006, Blood].** The paper reports high-grade lymphopenia and does not distinguish the cell populations affected beyond a CD25-gated Treg readout, which structurally confounds both the depletion claim and any off-target assessment. Severity is moderate — the CD25-gating confound is a measurement problem rather than a clear counter-productive mechanism, but transient depletion of CD25+ activated effectors is plausible for the dosing schedule used. **COI & funding.** Not accessible from abstract. **J Immunother is subscription-only; paywall-limited critique.** Spin / abstract-to-text consistency. **** Abstract frames the Treg reduction and lymphocyte rebound favorably and ties it to response rate, but the 33% RR matches historical HD IL-2 monotherapy benchmarks, suggesting the DD-specific immune-modulation effect may not translate into clinical benefit beyond IL-2 alone. Caution: abstract does not explicitly make the additive-benefit claim but invites that inference.**

Extraction fidelity. Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Serious**— Historical-control design, small arms, paywalled methods limit verification, likely CD25-gating confounding given DD mechanism. **Per-trial confidence.****Low**— Abstract-only access; historical controls; small n; CD25-gating confound concern unresolved; capped per the framework because full text unavailable.

Attia P 2005 — Inability of a Fusion Protein of IL-2 and Diphtheria Toxin (Denileukin Diftitox, DAB389IL-2, ONTAK) to Eliminate Regulatory T Lymphocytes in Patients With Melanoma

Citation.[PMID 16224276](#) · [PMCID PMC1533764](#) · [DOI 10.1097/01.cji.0000175461.19417.1a](#)**Source tier.**
[pmc_full_text](#)

Key critique.High-integrity negative study: n=13 melanoma patients, DD did NOT eliminate Tregs by FoxP3-mRNA readout (p=0.656 low dose; p=0.031 minimal high dose) and Tregs retained ≥50% suppressive function. The authors specifically chose FoxP3 mRNA over CD25 surface gating to avoid the classic CD25-confound, making this one of the cleanest DD-Treg-depletion studies methodologically. It is the key counterweight to Dannull 2005 (RCC DD positive) and establishes that DD's Treg-depleting activity is indication- and gating-dependent.

Design.Single-arm phase I/II style study with two dose cohorts (9 µg/kg/day × 5 days vs 18 µg/kg/day × 5 days, every 21 days). Paired pre/post Treg analysis. Primary endpoint was FoxP3 expression by qRT-PCR on purified CD4+ lymphocytes, i.e., a DIFFERENT readout than most DD studies (which use flow cytometry for CD4+CD25+ surface-gated Tregs). This methodological choice is critical and likely explains part of the discrepancy with Dannull 2005 and other positive DD studies — Attia avoided the CD25-gating confound by measuring FoxP3 mRNA.

Sample size & power.n=13 total (12 melanoma, 1 RCC); 7 in 9 µg/kg group, 6 in 18 µg/kg group. No formal power calculation — by design a feasibility / PD study. Underpowered to detect small changes; the 18 µg/kg finding (-2.01 copies/10³ β-actin, p=0.031) is statistically significant but 'clinically minimal' per the authors' own framing. **Effect-size calibration.**Negative result: no objective clinical response; minimal FoxP3 change at low dose (p=0.656), small and possibly clinically irrelevant change at high dose (p=0.031). Functional suppression assay: 'at least 50% suppression was present following Denileukin Diftitox' across five evaluated patients — this is critical. The Tregs that remained after DD were FUNCTIONALLY INTACT. This is the strongest evidence against DD as a Treg-depleter in melanoma and is a central counterweight to Dannull 2005 (RCC) and Geskin 2018 (CTCL). Discrepancy with Dannull likely reflects (a) indication (melanoma vs RCC), (b) gating (FoxP3 mRNA vs CD4+CD25hi surface + FoxP3 protein), and (c) small n and different PD-sampling windows. **Missing data.**Not explicitly discussed. Small cohort; attrition pattern not a major concern given sample size. **Multiplicity.**Primary endpoint pre-specified as FoxP3 expression; secondary suppression assay. Multiplicity not a major concern given the n. **Generalizability.**Stage IV melanoma only, KPS ≥60%; intramural NCI Surgery Branch population (highly selected). Negative finding in melanoma does not preclude DD activity in other indications. **Treg definition / gating.**Authors AVOID the CD25-based surface gating that confounds other DD studies by using FoxP3 mRNA as

the primary readout. This is methodologically superior for a CD25-targeting agent. However, FoxP3 mRNA in a bulk CD4+ sort is a blunt readout — it cannot distinguish reduced Treg number from reduced per-cell FoxP3 expression. ****Counter-productive mechanisms (Moderate (cd25-effector-collateral)). Same class concerns as other denileukin diftitox studies: DD depletes CD25-high cells, which include transiently activated CD25+ CD8 and CD4 effector T cells as well as Tregs [external: FDA ONTAK label]. This paper is methodologically unusual in using FoxP3 mRNA as the primary readout, which sidesteps the CD25-gating confound but still cannot distinguish depletion of Tregs from failure to deplete them when the denominator of CD4 T cells itself changes. The Attia 2005 negative result is consistent with the class mechanism of DD being insufficiently Treg-selective to produce a frequency reduction detectable by a FoxP3 readout. Severity is moderate for the shieldbreak question. COI & funding. NCI Surgery Branch intramural; Ligand Pharmaceuticals provided DAB389IL-2. Authors (Rosenberg group) are high-credibility academic oncologists with established independence from industry. COI likely declared absent in the full text. Spin / abstract-to-text consistency.**** Title is unambiguous: 'Inability... to eliminate regulatory T lymphocytes'. The abstract-to-full-text consistency is excellent — authors do not spin a minor high-dose p=0.031 into a depletion claim; they explicitly state 'does not appear to eliminate regulatory T lymphocytes.' High-integrity reporting.

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Low**— Small single-arm study but with rigorous pre-specified Treg endpoint, orthogonal FoxP3-mRNA readout avoiding the CD25-gating confound, functional suppression assay, and honest null-result reporting. **Per-trial confidence.Moderate**— The null result is well-supported by the study's methodological design (avoiding CD25-gating) and by functional-suppression-assay evidence. Confidence is capped at Moderate due to small sample (n=13) and the melanoma-specific setting.

Audia S 2007 — Increase of CD4+ CD25+ regulatory T cells in the peripheral blood of patients with metastatic carcinoma: a Phase I clinical trial using cyclophosphamide and immunotherapy to eliminate CD4+ CD25+ T lymphocytes

Citation.[PMID 17956583](#) · [DOI 10.1111/j.1365-2249.2007.03508.x](#)**Source tier.**pubmed_abstract

Key critique.Honest null-result paper: single IV cyclophosphamide did not reduce Tregs (n=49). Directly contrasts with Ghiringhelli 2007 metronomic oral dosing that showed 61% reduction (p<0.0001) — the lesson being that schedule, not just agent, determines Treg-depleting activity. Paywall limits full-text critique.

Design.Phase I single-arm trial of a single IV cyclophosphamide infusion combined with intratumoral BCG (non-specific immunotherapy) in n=49 metastatic carcinoma patients (plus 24 healthy donor comparisons for baseline Treg characterization). Paired pre/post Treg analysis. **Sample size & power.**n=49 is adequate for

Treg-level comparisons; patient cohort is larger than most in this shieldbreak. No formal power calculation for a 'no-change' finding. **Effect-size calibration.** Null result: single IV cyclophosphamide dose did NOT modulate Treg numbers or function. Directly contrasts with Ghiringhelli 2007 (metronomic oral cyclophosphamide, 61% reduction). The key difference: Ghiringhelli used METRONOMIC oral dosing, Audia used a SINGLE IV infusion. This is an important dose/schedule lesson: cyclophosphamide's Treg-depleting effect appears to require metronomic scheduling rather than single doses. **Missing data.** Not assessable from abstract only. **Multiplicity.** Not assessable from abstract only. **Generalizability.** Metastatic carcinoma mixed cohort; heterogeneous primary tumor types. BCG comparator arm may confound immune-modulation signal. **Treg definition / gating.** CD4+CD25+ with characterization of FoxP3, GITR, and intracellular CD152. Baseline gating quality looks reasonable; no CD25-targeting agent in use, so no gating confound. ****Counter-productive mechanisms (Moderate (lymphopenia-collateral)).** **Single IV cyclophosphamide is cytotoxic-dose chemotherapy, non-selective for Tregs vs effectors at that dose [external: Lutsiak et al. 2005, Blood]. The paper's null Treg result is consistent with this — at cytotoxic dose cyclo depletes lymphoid compartment broadly, including effector T cells needed for the BCG-induced anti-tumor response. Counter-productive concern: the dose regimen chosen depletes the effectors the BCG is trying to activate. Severity is moderate. COI & funding. Not accessible from abstract. Clin Exp Immunol is subscription-only; paywall-limited critique. Spin / abstract-to-text consistency.** Abstract is unusually direct — authors explicitly state that 'cyclophosphamide administration may not represent an optimal therapy to eliminate Treg.' Honest null-result reporting.**

Extraction fidelity. Matches source.

Retraction status. None found as of review date.

Risk of bias. ROBINS-I: **Moderate**— Single-arm with pre-specified Treg endpoint and honest null-reporting; paywall limits full methodological verification. **Per-trial confidence.** **Low**— Abstract-only access caps confidence per framework; n=49 and honest null-reporting suggest higher confidence may be warranted with full-text access.

Bentebibel SE 2019 — A First-in-Human Study and Biomarker Analysis of NKTR-214, a Novel IL2R $\beta\gamma$ -Biased Cytokine, in Patients with Advanced or Metastatic Solid Tumors

Citation. PMID 30988166 · DOI 10.1158/2159-8290.CD-18-1495 · NCT NCT02869295 **Source tier.** pubmed_abstract

Key critique. Industry-sponsored (Nektar) FIH paper claims NKTR-214/bempegaldesleukin promotes 'limited increase' in Tregs — a framing that subsequent data (Gogas 2024, Diab 2020) have substantially refined toward honest 8-10 \times absolute Treg expansion with ratio-shift as the proposed mechanism. Paywall + industry framing limit confidence.

Design. Phase I first-in-human multicenter dose-escalation trial of NKTR-214 (bempegaldesleukin) monotherapy in advanced solid tumors. Paired pre/post PD on PBMC and on-treatment tumor biopsies. **Sample size & power.** Not specified in abstract beyond 'multicenter phase I'; typical FIH dose-escalation cohort sizes. Not powered for a

formal Treg-comparison endpoint. **Effect-size calibration.**Paper's headline claim: NKTR-214 'designed to provide sustained signaling through heterodimeric IL2 receptor $\beta\gamma$ to drive increased proliferation and activation of CD8+ T and natural killer cells without unwanted expansion of T regulatory cells (Treg) in the tumor microenvironment.' Post-hoc literature (Gogas 2024 PIVOT-02 in this shieldbreak; Diab 2020) has established that bempeg DOES expand Tregs absolutely in both blood and tumor, with the ratio-shift being the claimed benefit — so Bentebibel 2019's framing of 'limited increase' in Tregs is an underspecified claim that has been substantially refined by subsequent data. **Missing data.**Not assessable from abstract. **Multiplicity.**Not assessable from abstract. **Generalizability.**Advanced solid tumors mixed cohort; FIH study. **Treg definition / gating.**Not described in abstract. ****Counter-productive mechanisms (High (treg-expansion-by-design, denominator-artifact-framing)).** **Bempegaldesleukin is a PEGylated IL-2 engineered to bias toward IL-2R $\beta\gamma$ (effector/NK) over IL-2R $\alpha\beta$ (Treg) signaling. The counter-productive mechanism for this shieldbreak is now well-evidenced: despite the 'preferential' effector signaling framing, bempeg ABSOLUTELY expands Tregs 8-10 \times in humans [external: Gogas 2024 same PIVOT-02 cohort]; the class claim of 'without Treg enhancement' is a frequency-denominator artifact. The clinical consequence is visible in the phase 3 PIVOT-IO-001 failure (2022) [external: Diab et al. 2023, JCO]. Bentebibel 2019's first-in-human report predates this evidence but the mechanism concern was knowable from the IL-2 biology. Severity is high. COI & funding.Nektar Therapeutics sponsor (drug developer). Heavy industry sponsorship context. The 'designed to not expand Tregs' framing is industry-promotional. Spin / abstract-to-text consistency.** The FIH paper's claim that NKTR-214 promotes 'limited increase' in Tregs is technically true but understates what subsequent Phase I/II data show — absolute Treg expansion occurs, with the ratio-shift being the claim of benefit. This is a classic FIH-biomarker-optimism pattern that later data have moderated.**

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Industry-sponsored FIH biomarker study with optimistic framing in the abstract; paywall limits full verification. **Per-trial confidence.****Low**— Abstract-only access; industry-sponsored framing has been moderated by subsequent data; per the framework abstract-only caps confidence at Low.

Brinkmann CR 2018 — The HDAC inhibitor panobinostat increases frequency and modulates phenotype of regulatory T cells in HIV-infected individuals on suppressive cART

Citation.[PMID 29468194](#) · [PMCID PMC5812898](#) · [DOI 10.1371/journal.ppat.1006827](#) · [NCT NCT01680094](#)**Source tier.**[pmc_full_text](#)

Key critique.A critical HDACi counterexample: panobinostat INCREASED Tregs by 40% at day 4 ($p=0.003$) and day 28 ($p=0.004$) in HIV+cART patients — opposite direction to entinostat (Pili 2017) and vorinostat (Terranova-Barberio 2020) in oncology. This establishes HDACi Treg effects as context- AND drug-specific, not class-general. Valuable honest reporting.

Design.Phase I/II single-arm PD study of panobinostat (a pan-HDAC inhibitor) in HIV-infected patients on suppressive cART. Paired pre/post Treg analysis on PBMC at baseline, day 4, day 28, and follow-up. NOT an oncology study — different context than most rows in this table. **Sample size & power.**Small cohort (~14-15 patients evaluable for PD). Modestly powered; paired-design increases efficiency. **Effect-size calibration.**Treg proportion INCREASED by 40% on day 4 ($p=0.003$), similar elevation at day 28 ($p=0.004$). This is an opposite-sign result from what one might expect for HDACi in oncology contexts (where entinostat, vorinostat DECREASE tumor Tregs per Pili 2017 and Terranova-Barberio 2020). The discrepancy is biologically meaningful: HIV-specific context (chronic viral infection + cART latency reversal) and the specific HDACi (panobinostat, broad HDAC specificity) likely matter. This is a HIGH-VALUE counterexample to 'HDAC inhibitors deplete Tregs' generalizations. **Missing data.**To confirm in full text. Paired-design means attrition affects both timepoints. **Multiplicity.**Multiple timepoints (day 4, 28, follow-up) — this is multiplicity. Authors do not appear to have pre-specified Bonferroni or FDR correction; $p=0.003$ and 0.004 are modestly strong and would survive a 4-test correction. **Generalizability.**HIV+cART context only; not directly generalizable to oncology settings. But highly relevant to the shieldbreak's question of context-dependent HDACi effects. **Treg definition / gating.**Standard CD4+CD25+FoxP3+ with additional phenotype markers (to be confirmed). **Counter-productive mechanisms (High (opposite-direction-mechanism, effector-function-suppression)). **Panobinostat is a pan-HDAC inhibitor used here for HIV reservoir reactivation. Paper reports a 40% INCREASE in Tregs (opposite direction to the shieldbreak hypothesis). Counter-productive concerns: (a) pan-HDAC inhibition can reduce effector CD8 T-cell function and destabilize anti-tumor memory [external: Kroesen et al. 2014, Oncotarget]; (b) the context (HIV cART, not tumor) makes generalization tentative. For the shieldbreak question, the drug's net direction is counter-productive (Tregs UP), placing severity at high; for its actual HIV indication this is a separate question. COI & funding.**Academic PLoS Pathogens publication; funding from HIV-research sources (Danish National AIDS Foundation, etc. — to confirm). **Lower COI risk than oncology industry studies.** Spin / abstract-to-text consistency. ** Title is explicit: panobinostat INCREASES Treg frequency. No spin — this is a contrarian result that the authors report directly.

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Low**— Single-arm paired pre/post design; modest n; honest counterexample reporting; standard gating. **Per-trial confidence.****Moderate**— Methodologically sound but specific to HIV+cART context; small n caps the strength of generalization.

Calvo E 2025 — Phase 1 Study of Nemvaleukin Alfa (ALKS 4230) Monotherapy in Patients with Advanced Solid Tumors

Citation.PMID 40759440· PMCID PMC12323519· DOI 10.1158/1078-0432.CCR-24-2924· NCT NCT02799095
Source tier.pmc_full_text

Key critique.Phase 1 nemvaleukin monotherapy shows Treg ~2× expansion with larger effector expansion (ratio-shift mechanism). Consistent with the 'non- α IL-2 family' pattern: Tregs expand but effectors expand more.

Industry-sponsored; phase 1 n caps precision; honest reporting of direction.

Design.Phase 1 multicenter dose-escalation/expansion trial of nemvaleukin alfa monotherapy. Single-arm paired pre/post PD analyses. **Sample size & power.**Phase 1 standard cohort sizes; not powered for formal Treg-comparison. **Effect-size calibration.**Reported ~2× Treg fold-change-from-baseline at Cmax (FCBmax). Effector cells (CD8+/NK) showed larger fold-change-from-baseline — supporting a ratio-shift rather than depletion mechanism. This is consistent with the nemvaleukin 'non-α IL-2' mechanism (IL-2 variants biased toward βγ signaling, which expands both Treg and effector but preferentially effector). Consistent with Vaishampayan 2024 in this shieldbreak. **Missing data.**Not explicitly discussed; typical phase 1 attrition. **Multiplicity.**Multiple timepoints and cell subsets reported; no multiplicity correction mentioned. Exploratory framing. **Generalizability.**Advanced solid tumors mixed cohort; post-ICI melanoma and RCC subgroups mentioned. **Treg definition / gating.**Standard CD4+CD25+FOXP3+ flow gating (to be confirmed). No CD25-targeting confound because nemvaleukin is designed to bind CD25-CD122-CD132 complex but with βγ bias. ****Counter-productive mechanisms (High (treg-expansion-by-design, denominator-artifact-framing)).****Nemvaleukin alfa is an engineered IL-2 fusion designed for IL-2Rβγ preference. Same counter-productive concern as bempreg: absolute Treg expansion can outpace the 'favorable ratio' [external: Dennis et al. 2022, Lancet Oncol; ARTISTRY-7 phase 3 failed in ovarian cancer, 2023]. Nemvaleukin's clinical failures in large phase 3 programs argue the mechanism shares bempreg's core problem at the clinical level. Severity is moderate-to-high; the absolute-Treg-expansion signal is smaller than bempreg but in the same direction.** COI & funding.**Mural Oncology (formerly Alkermes) sponsorship — drug developer. Industry sponsorship context is relevant for interpretation.** Spin / abstract-to-text consistency.** The 2× Treg expansion is honestly reported; the ratio-shift framing is the claim of benefit. Consistent with the bempreg / NKTR-214 framing pattern.

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Single-arm, industry-sponsored phase 1 PD study; no control arm; standard for this stage of development. **Per-trial confidence.****Moderate**— Consistent with orthogonal data (Vaishampayan 2024, other non-α IL-2 literature); small phase 1 limits effect-size precision.

Chand D 2024 — Botensilimab plus balstilimab in relapsed/refractory microsatellite stable metastatic colorectal cancer: a phase 1 trial

Citation.PMID 39083809· PMCID PMC11609826· DOI 10.1038/s41591-024-03094-4· NCT NCT03860272**Source tier.pmc_full_text**

Key critique.Fc-enhanced anti-CTLA-4 (botensilimab) + anti-PD-1 (balstilimab) in MSS mCRC — reports encouraging responses and intratumoral Treg reduction consistent with the Fc-engineering framework (Ager 2026, Sharma 2019). Industry-sponsored Phase 1; small PD subset; awaits randomized replication.

Design.Phase 1 first-in-human dose-finding trial of botensilimab (AGEN1181, Fc-enhanced multifunctional

anti-CTLA-4) + balstilimab (anti-PD-1) in MSS mCRC. Single-arm open-label with dose cohorts. Treg PD is an exploratory pre-specified endpoint. **Sample size & power.**Phase 1 standard n per dose cohort (typically 3-12); total enrolled ~65 patients with MSS mCRC. Treg-analysis subset is smaller. **Effect-size calibration.**Reports intratumoral Treg reduction — directionally consistent with the Fc-engineering-reconciling-variable hypothesis (cf. Ager 2026, Sharma 2019). Effect-size details to be verified against figures. **Missing data.**Typical phase 1; not a primary concern. **Multiplicity.**Multi-endpoint exploratory; no multiplicity correction mentioned. **Generalizability.** MSS mCRC only; post-chemo heavily pretreated population. **Treg definition / gating.**Flow cytometry with FOXP3 (to be confirmed in methods); no CD25-targeting confound. ****Counter-productive mechanisms (Moderate (adcc-activated-effector-collateral, paradoxical-autoimmunity)).****Botensilimab is an Fc-enhanced (Fc-optimized IgG1) anti-CTLA-4 engineered to increase FcγR engagement and ADCC/ADCP on CTLA-4-high Tregs. The counter-productive concerns: (a) Fc-enhancement can deplete activated effector T cells that transiently upregulate CTLA-4 [external: Simpson et al. 2013, J Exp Med; Arce Vargas et al. 2018, Cancer Cell]; (b) the combination with anti-PD-1 (balstilimab) compounds autoimmune-toxicity risk — the paper documents high-grade irAEs consistent with the class expectation. Unlike Ager 2026, this study does not report compartment-specific Treg data for tumor vs tdLN. Severity is moderate. COI & funding.****Agenus industry-sponsored (botensilimab is Agenus compound); multiple Agenus authors. Major industry context — findings are supportive of an ongoing commercial program and should be read accordingly.** Spin / abstract-to-text consistency. ****** The trial has been prominently cited by Agenus for the botensilimab program; the paper reports clinically meaningful responses in MSS mCRC, a historically ICI-refractory population. Abstract framing is appropriately confident given the effect size but is optimistic about class generalization.

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Single-arm dose-finding; industry sponsorship; exploratory PD endpoint; otherwise competent design. **Per-trial confidence.****Moderate**— Supportive of Fc-engineering hypothesis; industry-sponsored with promotional framing; awaiting randomized confirmation.

Comin-Anduix B 2008 — Detailed analysis of immunologic effects of the cytotoxic T lymphocyte-associated antigen 4-blocking monoclonal antibody tremelimumab in peripheral blood of patients with melanoma

Citation.[PMID 18452610](#) · [PMCID PMC2412852](#) · [DOI 10.1186/1479-5876-6-22](#)**Source tier.**[pmc_full_text](#)

Key critique.Early tremelimumab peripheral-blood PD analysis; aligns with the broader literature (corroborated by Huang 2011, Sharma 2019) that non-Fc-enhanced anti-CTLA-4 does not robustly deplete Tregs. Standard 2008-era translational rigor.

Design.Translational analysis of PBMC from melanoma patients treated with tremelimumab in a phase I/II clinical trial. Paired pre/post samples at multiple timepoints. Detailed flow cytometry focused on T-cell activation, Treg, and suppression. **Sample size & power.**n to confirm (typically 20-30 for this type of translational substudy at UCLA). **Effect-size calibration.**Peripheral blood Treg analysis — historical context: tremelimumab's blood Treg effect has been inconsistently reported across studies; Comin-Anduix is part of the broader 'no clear depletion' literature that aligns with Sharma 2019 (tremelimumab did not deplete intratumoral FOXP3+). **Missing data.**Not a primary concern; paired design. **Multiplicity.**Many flow-cytometry parameters analyzed; multiplicity correction likely not applied (2008 era translational study). **Generalizability.**Melanoma patients only; UCLA cohort. **Treg definition / gating.**Standard CD4+CD25+ FoxP3 flow gating (era-appropriate). Anti-CTLA-4 does not target CD25, so no gating confound. ****Counter-productive mechanisms (Low (alt-checkpoint-upregulation, non-depleting-isotype)).****Same class concerns as Huang 2011 and Ribas 2009: tremelimumab does not deplete intratumoral Tregs in humans, and CTLA-4 blockade induces compensatory alt-checkpoint upregulation [external: Huang 2017, PNAS; Woo et al. 2012]. The paper adds the observation that tremelimumab increases activated effector markers on both Treg and Teff compartments — consistent with the class interpretation that anti-CTLA-4 is a general T-cell co-stimulation release rather than a Treg-selective depletion agent. Severity is low for the Treg-depletion question. COI & funding. Ribas is on Pfizer advisory boards per related papers (Huang 2011 discloses this). Academic-led study but with industry funding for tremelimumab. Spin / abstract-to-text consistency. ** Title says 'detailed analysis' — descriptive framing, not over-selling a positive result.**

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Translational paired-sample analysis; standard rigor for 2008; no major methodological concerns. **Per-trial confidence.****Moderate**— Consistent with subsequent converging evidence (Sharma 2019, Huang 2011) that standard anti-CTLA-4 does not robustly deplete Tregs in humans.

Dannull J 2005 — Enhancement of vaccine-mediated antitumor immunity in cancer patients after depletion of regulatory T cells

Citation.[PMID 16308572](#) · [PMCID PMC1288834](#) · [DOI 10.1172/JCI25947](#)**Source tier.**[pmc_full_text](#)

Key critique.THIS IS A KEY PAPER TO CRITIQUE. Very influential (>2000 citations); n=7 treated with single DAB389IL-2 dose + RNA-DC vaccine in RCC; reports 26-76% Treg reduction with abrogated suppressive function. CRITICAL GATING CONFOUND: CD25-based gating during CD25-targeting therapy is structurally suspect. Authors partially mitigate with FoxP3 mRNA but the 30-80% mRNA-reduction range is also wide. Directly contradicts Attia 2005 (n=13 melanoma, DD failed). The 2005 discrepancy has not been cleanly resolved and the field's pivot to Fc-enhanced anti-CTLA-4 reflects waning confidence in the DD-depletion claim. Confidence: Low despite JCI publication.

Design.Single-arm pilot translational study in J Clin Invest. n=11 enrolled (10 RCC + 1 ovarian); 7 received

DAB389IL-2 (18 µg/kg single IV dose) 4 days before RNA-DC vaccine; 4 received vaccine alone as controls; 1 received DAB389IL-2 alone. Paired pre/post Treg measurement plus functional suppression assay. PMC XML-paywalled; critique based on abstract/WebFetched PMC HTML. **Sample size & power.**Very small: n=7 receiving the combination. No formal power calculation; descriptive comparison to n=4 vaccine-only controls. This is a pilot / proof-of-concept sample size — effect-size estimates will be noisy. **Effect-size calibration.**Reported 26-76% reduction in CD4+CD25high frequency and 30-80% reduction in FoxP3 mRNA (day 4 post-DD). The wide range (26-76%) across 7 patients indicates significant inter-patient heterogeneity. Functional suppression assay: Treg-mediated suppression was 'abrogated' post-DD. This is the central positive DD paper for Treg depletion — directly contrasts with Attia 2005 (melanoma, DD failed). The reconciling variable is gating: Dannull measured CD4+CD25high with FoxP3-mRNA corroboration, while Attia used FoxP3 mRNA alone with different sort purity. Critically, DANNULL'S CD25-BASED GATING IS PARTIALLY CONFOUNDED — DAB389IL-2 targets CD25, so post-treatment the surviving Tregs may appear reduced not because they were depleted but because CD25 was masked or internalized (though the authors' short-half-life argument has merit, it does not fully resolve the issue). The FoxP3 mRNA data are the more gating-independent readout. **Missing data.**Small cohort; not systematically addressed. **Multiplicity.**Multiple immune readouts; no correction. The variability in the 26-76% range weakens the central claim. **Generalizability.**Predominantly RCC (+1 ovarian). The DAB389IL-2 + vaccine pairing is specific to this combined intervention. Standalone DD activity in other indications is addressed better by Attia 2005 (negative in melanoma) and Geskin 2018 (positive in CTCL). **Treg definition / gating.**CD4+/CD25high + intracellular FoxP3 protein flow + FoxP3 mRNA on purified CD4+ sort. The CD25-based gating is the classic confound for DD studies — the intervention targets CD25, so post-treatment CD25+ cell measurements are structurally compromised. Authors mitigate this with FoxP3 mRNA but the 30-80% mRNA-reduction range is also wide. Intracellular FoxP3 protein staining is more gating-independent and worth emphasizing; it is reported to corroborate the finding but at smaller sample. ****Counter-productive mechanisms (Moderate (cd25-effector-collateral, vaccine-priming-window)).** **Denileukin diftitox given before a tumor-RNA DC vaccine in RCC. The counter-productive concern is acute: the vaccine's effect depends on CD25-upregulating newly primed effector CD8 T cells, and DD has a window of anti-CD25 activity that overlaps the priming phase. If DD is administered close to vaccination, it may deplete the very effectors the vaccine is generating [external: Rech et al. 2012 PLoS ONE; Sci Transl Med 2009]. The paper's CD25-gated Treg readout also structurally cannot measure this off-target effect. The Morse 2008 daclizumab paper explicitly addresses this timing question for daclizumab; Dannull 2005 does not. Severity is moderate-to-high for the vaccine-combination context. COI & funding. NIH R21-CA098446 and M01-RR-30. Authors declared no COI. Ligand Pharmaceuticals provided DAB389IL-2. Duke academic-led. PMC paywalled — full acknowledgements read via WebFetch. Spin / abstract-to-text consistency.** Title says 'enhancement of vaccine-mediated antitumor immunity ... after depletion of regulatory T cells' — commits to the depletion claim. The wide reported range (26-76%) is in the body of the paper but the abstract headline is less hedged. This is mild abstract-to-text-optimism, not outright spin.**

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Small pilot cohort; CD25-gating confound partially (but not fully) mitigated by FoxP3 mRNA; wide effect-size range; high-impact journal (JCI) with lasting citation influence. **Per-trial confidence. Low**— High-profile paper but methodologically the CD25-gating confound remains partially unresolved; very small n (7 treated); wide 26-76% range. This paper's positive finding should be weighted lightly until replicated in larger, more rigorously-gated cohorts.

Davar D 2022 — A phase I study of GITR agonist TRX518 alone and in combination with nivolumab in patients with advanced solid tumors

Citation. PMID 35499569 · PMCID PMC9475244 · DOI 10.1158/1078-0432.CCR-22-0339 · NCT NCT02628574
Source tier. pmc_full_text

Key critique. Phase 1 TRX518 anti-GITR with/without nivolumab; modest PD signals in a program that has been largely deprioritized. Confidence in 'anti-GITR depletes Tregs in humans' is Low: the paper's PD signals did not support sufficient efficacy for program advancement.

Design. Phase I dose-escalation/expansion of TRX518 (anti-GITR Ab) monotherapy and combination with nivolumab in advanced solid tumors. Open-label single-arm with dose cohorts; paired PBMC and tumor biopsies.
Sample size & power. Phase 1 standard cohorts; combined n per dose. Two search rows suggest PBMC and tumor compartment analyses. Small cohorts per dose limit precision.
Effect-size calibration. Anti-GITR Treg-depletion claim is weakly supported in humans; TRX518 program has been largely deprioritized clinically. The published PD data on Treg reduction is modest.
Missing data. Standard phase 1; small samples per dose may inflate attrition impact.
Multiplicity. Multiple dose cohorts, multiple timepoints, PBMC + tumor compartments — significant multiplicity; not formally corrected.
Generalizability. Heavily-pretreated advanced solid tumors — poor generalizability to less-pretreated populations.
Treg definition / gating. To confirm in methods; anti-GITR does not target CD25, no gating confound. **Counter-productive mechanisms (Moderate (adcc-activated-effector-collateral, depletes-beneficial-effectors)).
TRX518 is an afucosylated anti-GITR aimed at Treg depletion via ADCC. Counter-productive concerns: (a) GITR is expressed on activated CD8 effector T cells as well as Tregs, so afucosylated anti-GITR can deplete activated effectors [external: Pastor-Fernandez et al. 2023, Front Immunol]; (b) the anti-GITR clinical program did not advance despite modest PD signals, consistent with the collateral concern not being bounded by drug design. Severity is moderate. COI & funding. Leap Therapeutics sponsorship; industry context. Spin / abstract-to-text consistency. Program-level context: TRX518 did not advance past Phase 1; the paper is a terminal-reporting paper. Abstract may frame PD signals more positively than clinical disappointment warrants.**

Extraction fidelity. Matches source.

Retraction status. None found as of review date.

Risk of bias. ROBINS-I: **Moderate**— Phase 1 dose-escalation; industry-sponsored; modest PD effect size in a discontinued program. **Per-trial confidence.** **Low**— Program discontinuation is a real-world signal that the mechanism did not translate; modest effect size and small cohort.

Diab A 2020 — Bempegaldesleukin (NKTR-214) plus Nivolumab in Patients with Advanced Solid Tumors:

Phase I Dose-Escalation Study of Safety, Efficacy, and Immune Activation (PIVOT-02)

Citation.PMID 32439653 · DOI 10.1158/2159-8290.CD-20-0096 · NCT NCT02983045 Source tier.pubmed_abstract

Key critique.PIVOT-02 abstract claims bempreg+nivo was active 'without regulatory T-cell enhancement' — a framing subsequently contradicted by the same cohort's data (Gogas 2024: 8-10× absolute Treg expansion). The abstract-level spin helped advance the phase 3 PIVOT-IO program, which subsequently FAILED clinically. Paywall limits full-text critique but the spin vs downstream-data picture is clear.

Design.PIVOT-02 phase 1 single-arm dose-escalation of bempreg + nivolumab in advanced solid tumors (melanoma, RCC, NSCLC). n=38. Paired pre/post PBMC + tumor biopsy longitudinal analyses. **Sample size & power.**n=38 total; per-cohort analyses smaller. Abstract reports 'without regulatory T-cell enhancement' — which is contradicted by subsequent phase 3 data and by Gogas 2024 CBM-01 follow-on studies showing clear Treg absolute expansion. **Effect-size calibration.**The 'without regulatory T-cell enhancement' claim in the abstract is the key methodological issue — this framing appears to refer to frequency (Treg as % of CD4) rather than absolute count. Gogas 2024 and Bentebibel 2019 both confirm absolute Treg expansion; the 'ratio-shift' framing was the intended interpretation. This is a classic case where abstract simplification (frequency vs absolute count) creates downstream confusion. The phase 3 PIVOT-IO trials subsequently failed, and bempreg development was halted in 2022, which is critical context for interpreting PIVOT-02's optimistic framing. **Missing data.**Not assessable from abstract. **Multiplicity.**Abstract-only, not assessable. **Generalizability.**Phase 1 mixed tumor types; phase 3 failed to replicate the early signal. **Treg definition / gating.**Standard flow; to confirm in full text. No CD25-gating confound. ****Counter-productive mechanisms (High (treg-expansion-by-design, denominator-artifact-framing)).****Same class concerns as Bentebibel 2019: bempreg + nivo in PIVOT-02. The original abstract framed the drug as effector-selective; the subsequent Gogas 2024 reanalysis of the same cohort documented 8-10× absolute Treg expansion, and the PIVOT-IO-001 phase 3 failed its primary endpoints [external: Diab et al. 2023, JCO]. The counter-productive mechanism is the drug's actual MoA: it expands Tregs, just expands effectors slightly more by frequency. Severity is high; this is arguably the cleanest negative-clinical-result evidence in the entire set that the 'favorable ratio' framing is counter-productive to the shieldbreak's goal. COI & funding.Nektar Therapeutics + BMS sponsorship. Major industry context — this was the lead-in paper for the PIVOT-IO phase 3 program that subsequently failed clinically. Spin / abstract-to-text consistency.** SIGNIFICANT SPIN CONCERN: abstract claims 'without regulatory T-cell enhancement' — a claim that conflates frequency-as-%-of-CD4 with absolute count. Gogas 2024 (same PIVOT-02 cohort, subsequent analysis) documents the 8-10× absolute Treg expansion. This is exactly the kind of abstract-to-full-text inconsistency the user flagged — the paper's 'no Treg enhancement' claim is misleading. Importantly, this abstract-level framing helped justify advancement into the phase 3 PIVOT-IO trials, which subsequently FAILED in melanoma and RCC.**

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Serious**— Industry-sponsored phase 1 with abstract-level misleading framing of Treg endpoint (subsequently contradicted by Gogas 2024 from the same cohort); clinically important because it supported failed phase 3 program. **Per-trial confidence.****Low**— Subsequent clinical-trial failure of the phase 3 extension and clearer Treg-expansion data from the same cohort (Gogas 2024) substantially undermine the paper's headline claims.

Fujikawa K 2023 — Mogamulizumab efficacy is underpinned by depletion of CCR4+ regulatory T cells and CCR4+ Sézary cells: a Japanese post-marketing surveillance study

Citation. PMID 37729184 · PMCID PMC10511099 · DOI 10.1002/ijc.34737 **Source tier.** pmc_full_text

Key critique. Mogamulizumab depletes CCR4+ effector-Tregs by ~90% (2.1%→0.20%) in PBMC of n=37 evaluable ATLL/CTCL patients. Rigorous eTreg gating (Miyara Fr. II); on-target mechanism; corroborated by Jinushi 2025 (tumor) and Roelens 2022 (Sézary). One of the strongest Treg-depletion signals in the shieldbreak, capped slightly by industry sponsorship.

Design. Post-marketing surveillance / translational observational study in Japanese ATLL/CTCL patients receiving mogamulizumab. Paired pre/post PBMC analysis of CCR4+ eTreg (effector Treg) subset. **Sample size & power.** n=37 evaluable of 49 enrolled; reasonable sample. Paired pre/post design efficient for within-patient comparison. **Effect-size calibration.** Median ~90% eTreg reduction (from 2.1% to 0.20%); eTreg depletion in all but 2 patients. Effect size is very large — consistent with mogamulizumab's targeted anti-CCR4 mechanism. This is among the cleanest Treg-depletion signals in the shieldbreak (Jinushi 2025 tumor-compartment data from the same class corroborates). **Missing data.** 12 of 49 patients unable to have samples (24% dropout), non-trivial; could represent progressive disease with fewer viable cells. **Multiplicity.** Single primary PD endpoint (eTreg); not a major multiplicity concern. **Generalizability.** Japanese ATLL/CTCL population; mogamulizumab approved in this indication. Generalizability to other CCR4-targeting contexts and populations should be cautious; Roelens 2022 Sézary French data support. **Treg definition / gating.** CCR4+FoxP3+CD45RA-CD25+ effector-Treg gating (Miyara classification). Gating quality is high; effector-Treg definition (Miyara Fr. II: CD45RA-FOXP3hi) is rigorous. ****Counter-productive mechanisms (High (depletes-beneficial-effectors, lymphopenia-collateral)).** **Mogamulizumab anti-CCR4.** The paper itself documents the canonical class counter-productive mechanism: 'unexpected decrease in central memory CD8+ T cells, which also express CCR4 and have anti-tumor immunity' [paper Discussion], citing Kurose et al. 2015 and the Nature Communications 2021 preclinical work [external: Tanaka et al. 2021]. The paper argues 'KW-0761 may concurrently deplete eTregs and central memory CD8+ T cells, and this dual depletion may cancel anti-tumor immune responses.' The paper also documents high-grade lymphopenia (25%). **Severity is HIGH — this is paper-internal, quantitative, and the authors explicitly argue it explains the 'minimal clinical benefit' seen in the trial.** **COI & funding.** Kyowa Kirin (mogamulizumab manufacturer) post-marketing surveillance — industry sponsorship context is strong. **However, the drug mechanism is on-target and the finding is orthogonal to multiple other mogamulizumab studies.** Spin / abstract-to-text consistency. **** Title strongly commits to the depletion claim ('underpinned by depletion');** body of paper quantifies ~90% and is consistent. No inappropriate spin.

Extraction fidelity. Matches source.

Retraction status. None found as of review date.

Risk of bias. ROBINS-I: **Moderate** — Industry-sponsored observational design; large effect size; rigorous eTreg gating; moderate attrition. **Per-trial confidence.** **High** — Large effect size, on-target mechanism, rigorous eTreg gating, and orthogonal corroboration (Jinushi 2025 tumor, Roelens 2022 Sézary). One of the strongest Treg-depletion signals in the table.

Gadi D 2022 — Idelalisib reduces regulatory T cells and activates T helper 17 cell differentiation in relapsed chronic lymphocytic leukaemia patients (correlative study)

Citation. PMID 35170759 · PMCID PMC9263710 · DOI 10.1038/s41375-022-01535-y · NCT NCT01539512 **Source tier.pmc_full_text**

Key critique. Small (n=9) idelalisib CLL correlative study showing Treg decrease — mechanism-aligned (PI3Kδ required for Treg survival) but small-sample.

Design. Correlative translational study in CLL patients receiving idelalisib (PI3Kδ inhibitor). Paired pre/post flow-cytometry of PBMC Tregs; n=9. **Sample size & power.** Very small (n=9); exploratory PD study. **Effect-size calibration.** Direction (Treg decrease) aligns with the PI3Kδi mechanism — PI3Kδ is selectively required for Treg survival and function; preclinical data support. Small n caps effect-size precision. **Missing data.** n=9 is small; attrition not a major concern at this scale. **Multiplicity.** Multiple T-cell subset analyses (Treg, Th17, Th1/2); no multiplicity correction. **Generalizability.** Relapsed CLL only; idelalisib program has been limited by infectious and hepatic toxicity. **Treg definition / gating.** CD4+FOXP3+CD25+CD127low (to confirm); no CD25-targeting, no gating confound. ****Counter-productive mechanisms (Moderate (effector-function-suppression, paradoxical-autoimmunity)).** **Idelalisib is a PI3Kδ inhibitor. PI3Kδ signaling is required for Treg proliferation and function but ALSO for effector CD8 T-cell activation and germinal center formation [external: Okkenhaug et al. 2014, Cancer Discov; Ahmad et al. 2017, Clin Cancer Res]. Counter-productive concerns: (a) idelalisib has a characteristic toxicity profile of hepatitis, colitis, pneumonitis that reflects off-target effector disruption; (b) the drug's label includes black-box warnings for these immune-mediated AEs. Severity is moderate-to-high for the shieldbreak question; idelalisib is approved in CLL despite these concerns because the disease-specific efficacy dominates. COI & funding. Academic; funding to confirm. Spin / abstract-to-text consistency. ** Descriptive correlative paper; no evident spin.**

Extraction fidelity. Matches source.

Retraction status. None found as of review date.

Risk of bias. ROBINS-I: **Moderate**— Small correlative study; mechanism-aligned finding; no multiplicity correction. **Per-trial confidence.** **Low**— n=9 caps effect-size precision; directional finding is consistent with mechanism but requires larger confirmation.

Geskin LJ 2018 — Therapeutic reduction of cell-mediated immunosuppression in mycosis fungoides and Sézary syndrome by denileukin

diftitox (correlative analysis)

Citation. PMID 29204699 · PMCID PMC8274400 · DOI 10.1080/2162402X.2017.1390642 **Source tier.** pmc_full_text

Key critique. n=77 CTCL correlative analysis; median 29% DD-induced Treg reduction (p=0.03) with responder-dependence (responders 20-45% absolute decrease, non-responders 2/3 INCREASE). Response-stratified finding is the most interesting signal. CD25-gating confound (DD targets CD25) caps confidence at Moderate.

Design. Correlative analysis of n=77 CTCL patients (21 stage IA MF, 30 stage IB+ MF, 26 Sézary) plus 9 healthy controls. DD given as 5 consecutive daily IV infusions with 2-week breaks. Paired pre/post. PMC XML-paywalled; critique based on WebFetched PMC HTML. **Sample size & power.** Large cohort (n=77) compared to most DD studies. Patient stratification by disease stage strengthens subgroup comparisons. **Effect-size calibration.** Median 29% Treg reduction (94% CI -83% to -20%, p=0.03) after one DD cycle. Effect-size heterogeneity is substantial (wide CI). Responders showed 20-45% absolute-count decrease; non-responders had 2/3 INCREASE. This response-stratified pattern is the most clinically interesting signal — it aligns with a mechanism hypothesis that DD's immunologic benefit is confined to responders. **Missing data.** Paywall limits verification. **Multiplicity.** Multiple comparisons (disease stage subgroups, responder vs non-responder) — no explicit multiplicity correction but the primary 29% median reduction has p=0.03 which would not survive strict Bonferroni. **Generalizability.** CTCL-specific; disease is already characterized by cutaneous lymphoproliferation with Tregs in flux — baseline immunology is non-standard. **Treg definition / gating.** CD4+CD25+FoxP3+ flow gating. CD25-gating confound applies: DD targets CD25. Post-treatment reduced CD25 staining may reflect epitope masking rather than true Treg depletion. The 'natural Treg' framing in the abstract does not fully resolve this. ****Counter-productive mechanisms (Moderate (cd25-effector-collateral, cd25-tumor-target-confound)).** **In CTCL the tumor cells themselves are often CD25+ or CD4+CD25-intermediate, so DD's anti-tumor activity and its anti-Treg activity are not separable in this disease — a well-evidenced class distinction from the solid-tumor indications [external: FDA ONTAK label]. The counter-productive concern for the shieldbreak question is that the 'Treg depletion' readout in CTCL is contaminated by depletion of CD25+ malignant cells and by depletion of CD25+ non-malignant T cells that participate in anti-tumor immunity. Severity is moderate — the drug clearly works in CTCL but not in a way that cleanly isolates a Treg-depletion mechanism for generalization to other indications. COI & funding. NIH P50CA121973 (SPORE); authors declared no COI. **Academic; lower industry-bias concern.** Spin / abstract-to-text consistency. **** Abstract framing is consistent with results — 29% median reduction and responder-stratified analyses are appropriate framing. No over-claiming.****

Extraction fidelity. Matches source.

Retraction status. None found as of review date.

Risk of bias. ROBINS-I: **Moderate** — Larger cohort than most DD studies but still exploratory correlative; CD25-gating confound applies; paywalled limiting verification. **Per-trial confidence.** **Moderate** — Larger n and response-stratified analyses strengthen inference; CD25-gating confound caps confidence.

Ghiringhelli F 2007 — Metronomic

cyclophosphamide regimen selectively depletes CD4+CD25+ regulatory T cells and restores T and NK effector functions in end stage cancer patients

Citation.PMID 16960692 · PMCID PMC11030569 · DOI 10.1007/s00262-006-0225-8 **Source tier.pmc_full_text**

Key critique.n=9 end-stage cancer study; metronomic oral cyclophosphamide → 61% Treg frequency reduction (7.9→3.1%, p<0.0001), 78% count reduction (28.7→6.4 cells/mm³, p<0.0001) with restored NK/T-cell function. Small n but large effect; direct counter-evidence to Audia 2007 (single IV dose null). The dose-schedule lesson is the central takeaway for cyclophosphamide-Treg strategy.

Design.Small single-arm paired pre/post observational study in n=9 end-stage cancer patients on metronomic oral cyclophosphamide, compared against n=15 healthy-volunteer baseline. Functional suppression assays (NK cytotoxicity, T-cell proliferation) pre/post. PMC XML-paywalled; critique based on WebFetched PMC HTML.

Sample size & power.n=9 treated patients is small. The pairing and within-patient control increase efficiency. No formal power calculation stated; effect size observed is very large (61% frequency reduction, 78% count reduction) relative to noise. **Effect-size calibration.**~61% frequency reduction (7.9% to 3.1%, p<0.0001) and 78% absolute-count reduction (28.7 to 6.4 cells/mm³, p<0.0001). Strong effect, but small n — regression to the mean concern is real. Replicated partly by Scurr 2017 in colorectal (referenced by Liao 2024). Direct contrast with Audia 2007 single-IV-dose null result is the central methodological lesson: metronomic oral scheduling appears essential for cyclophosphamide Treg-depleting activity. **Missing data.**Paywall limits verification; small cohort. **Multiplicity.**Multiple functional endpoints (NK, T-cell proliferation) — no explicit multiplicity correction. Fisher's exact method is mentioned but is not obviously appropriate for paired continuous outcomes; Wilcoxon signed-rank would be more standard. **Generalizability.**End-stage cancer mixed population (6M/3F, range 37-65); very-late-line setting. Effect may not generalize to earlier-line or specific cancers. **Treg definition / gating.**CD3/CD4/CD25 surface + intracellular FoxP3 (4 patients). Cyclophosphamide does not target CD25, so no gating confound.

****Counter-productive mechanisms (Low (lymphopenia-collateral)).****Metronomic (low-dose oral) cyclophosphamide is reported to be preferentially Treg-depleting because Tregs have lower ATP reserves and are more sensitive to low-dose alkylation [external: Lutsiak et al. 2005, Blood; Ghiringhelli et al. 2004, Cancer Immunol Immunother]. The counter-productive concern is bounded by the dosing: at metronomic dose, collateral effector depletion is real but modest, and effector function is reportedly preserved. Severity is low — this is arguably the cleanest dose-selective Treg-depletion mechanism in the set, but it is not effector-free. COI & funding.****Ligue Nationale contre le Cancer (French charity) grants. COI not explicitly disclosed in PMC snippet — full-text verification needed.** Spin / abstract-to-text consistency.** Title firmly commits to 'selectively depletes'. Effect size supports this framing. No evident spin given the observed p<0.0001 numbers.

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Small n (n=9) but large and highly statistically significant effect; paired design; paywall limits full verification; statistical method choice (Fisher's exact for continuous) is non-optimal.

Per-trial confidence.**Moderate**— Large effect size and functional-assay corroboration support inference; n=9 and methodological choices cap at Moderate. Replication in larger colorectal cohorts (Scurr 2017) further supports but is not in this row.

Gogas H 2024 — Reprogramming of the immune microenvironment by bempegaldesleukin + nivolumab in advanced melanoma (PIVOT-02 cohort biomarker analysis)

Citation. PMID 39025948 · PMCID PMC11258232 · DOI 10.1038/s41467-024-49940-4 · NCT NCT02983045 **Source**
tier.pmc_full_text

Key critique. Same-cohort biomarker update to Diab 2020 PIVOT-02. Documents 8-10× ABSOLUTE Treg expansion (contradicting the earlier 'without Treg enhancement' abstract claim) with ratio-shift (CD8:Treg DECREASED — Treg expansion outpaced CD8 2× expansion). Honest reporting. Contextually important: the phase 3 PIVOT-IO program subsequently failed (2022), so this 'ratio-shift' mechanism did not translate to clinical benefit.

Design. Biomarker analysis of the PIVOT-02 phase 1 melanoma cohort (NCT02983045). Two-arm comparison: bempeg+nivo vs nivo monotherapy. Longitudinal PBMC + tumor biopsies at C1D8, C5D1, C5D8. **Sample size & power.** Biomarker subset; n per arm smaller than main trial. Sufficient for directional claims but not formal phase 3 power. **Effect-size calibration.** DIRECTLY CONTRADICTS the Diab 2020 PIVOT-02 abstract claim ('without regulatory T-cell enhancement'). Gogas 2024 documents ~8-10× absolute Treg expansion at C1D8/C5D1/C5D8 vs baseline, with CD8:Treg ratio DECREASED (Treg expansion outpaced CD8 2× expansion). This is the honest-reporting update from the same cohort — a demonstration that early-phase abstract framing can obscure signals visible in later biomarker papers. Notably, the follow-on phase 3 PIVOT-IO trials in melanoma/RCC FAILED clinically, and bempeg development was halted in 2022. **Missing data.** Biomarker subset; not systematically discussed. **Multiplicity.** Many cell-subset and timepoint comparisons; p-values reported as asterisks ($p < 0.00001$ / $p < 0.0001$) without apparent multiplicity correction. The large effect sizes would likely survive Bonferroni even so, but the methodological norm should be explicit correction for this many comparisons. **Generalizability.** Melanoma subset of PIVOT-02; downstream phase 3 failure means findings don't generalize to a clinical-benefit claim. **Treg definition / gating.** Flow cytometry; to confirm. No CD25-gating confound. ****Counter-productive mechanisms (High (treg-expansion-by-design, denominator-artifact-framing)).** **Gogas 2024 is the reanalysis that quantified what Bentebibel 2019 and Diab 2020 framed away: bempeg absolutely expands Tregs 8-10× and the 'favorable ratio' was a denominator artifact. The counter-productive mechanism is paper-internal and well-documented here: the drug expands the population it was designed to spare, and the PIVOT-IO-001 phase 3 clinical failure is the clinical corollary [external: Diab et al. 2023, JCO]. Severity is high, and this paper is methodologically the most valuable for understanding WHY the class counter-productive mechanism matters. COI & funding. Nektar + BMS sponsorship. Industry-sponsored but honest contradiction of the earlier abstract framing is to the authors' credit.** Spin / abstract-to-text consistency. **** Biomarker paper is appropriately framed: absolute expansion documented, ratio-shift framing used. The contrast with Diab 2020 abstract is the spin issue — not in this paper but in the earlier PIVOT-02 publication.**

Extraction fidelity. Matches source.

Retraction status. None found as of review date.

Risk of bias. ROBINS-I: **Moderate** — Biomarker analysis with two-arm comparison; standard rigor;

industry-sponsored; multiplicity handling informal; clarifies earlier abstract-level spin. **Per-trial confidence. Moderate**— Reliable biomarker signal directly from the cohort; context of subsequent phase 3 failure argues against over-interpretation as a clinical-benefit mechanism.

Gordon MJ 2025 — Phase 1 trial of zanubrutinib, obinutuzumab, and venetoclax in chronic lymphocytic leukemia (translational Treg substudy)

Citation.PMID 40546724 · PMCID PMC12182836 · DOI 10.1158/1078-0432.CCR-25-0456 **Source tier.** pmc_full_text

Key critique.Phase 1 three-agent CLL regimen translational substudy with Treg PD as correlative endpoint. Combination makes single-agent Treg-effect attribution impossible; small n; low mechanism-specific confidence.

Design.Phase 1 translational substudy. CLL-specific regimen; Treg PD reported as exploratory. **Sample size & power.**Small phase 1 translational subset. **Effect-size calibration.**Combination of BTKi + anti-CD20 + BCL2i; the Treg effect is likely indirect (lymphocyte depletion from anti-CD20 pulls B cells and some T cells). Mechanism attribution to any single agent is difficult. **Missing data.**Standard phase 1 attrition. **Multiplicity.**Multiple agent combination complicates causal inference. **Generalizability.**CLL only. **Treg definition / gating.**To confirm.

**Counter-productive mechanisms (Moderate (depletes-beneficial-effectors, lymphopenia-collateral)).

Mogamulizumab + rhIL-15 in R/R T-cell malignancies. Class counter-productive concern applies: CCR4+ CD8 central memory collateral [external: Tanaka 2021, Nat Commun; Fujikawa 2023 paper]. The rhIL-15 combination partially addresses this by expanding memory CD8 and NK populations post-depletion, which could rescue the effector compartment — but the paper does not directly measure whether rescue is effective for the depleted subset. Severity is moderate. COI & funding. Industry-supported CLL regimen; obinutuzumab = Genentech/Roche, venetoclax = AbbVie, zanubrutinib = BeiGene. Spin / abstract-to-text consistency. Not a primary Treg-depletion trial; Treg finding is correlative.**

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Small phase 1 combination translational substudy; causal attribution to any one agent difficult. **Per-trial confidence.****Low**— Combination regimen makes single-agent Treg-effect attribution difficult; small n.

Govindaraj C 2014 — Reducing TNF receptor 2+ regulatory T cells via the combined action of

azacitidine and the HDAC inhibitor, panobinostat for clinical benefit in acute myeloid leukemia patients

Citation.PMID 24297862· DOI 10.1158/1078-0432.CCR-13-1756Source tier.pubmed_abstract

Key critique.n=14 AML paper focused on TNFR2+ Treg SUBSET (a minority of total Tregs). Azacitidine + panobinostat reduces this subset in PBMC and BM with associated clinical benefit. Panobinostat is the primary in-vitro driver — but in HIV (Brinkmann 2018) panobinostat INCREASES TOTAL Tregs, so this is a subset-specific / context-dependent effect, not a general 'HDACi depletes Tregs' finding.

Design.Translational study: n=14 AML patients receiving azacitidine + panobinostat with n=30 healthy donor comparisons. Paired pre/post analyses of TNFR2+ Treg subset in PBMC and bone marrow. **Sample size & power.**n=14 is small; healthy-donor comparator is orthogonal rather than matched. **Effect-size calibration.**The paper targets the TNFR2+ Treg SUBSET (a minority of total Tregs enriched for suppressive function). Reduction of this subset is a subtle and subset-specific finding — it does NOT generalize to 'HDACi + HMA reduces total Tregs.' Directionally CONSISTENT with the Treg-reducing HDACi/HMA literature subset but specifically scoped. Panobinostat is identified as the primary driver in vitro — which is consistent with the opposite-direction Brinkmann 2018 HIV-panobinostat data (where panobinostat INCREASED total Tregs) — suggesting panobinostat preferentially targets TNFR2+ subset. Context-dependent, subset-dependent mechanism. **Missing data.**Not assessable from abstract. **Multiplicity.**Subset-specific endpoint reduces multiplicity burden compared to broad immune-profiling. **Generalizability.**AML only; TNFR2+ subset specifically. **Treg definition / gating.** CD4+FOXP3+TNFR2+ gating — high-quality, subset-specific. No CD25-targeting drug; no gating confound. ****Counter-productive mechanisms (Moderate (effector-function-suppression, treg-induction-by-dnmti)).Azacitidine + panobinostat in AML. The counter-productive concern is compound: (a) DNMTi (azacitidine) can demethylate the FOXP3 TSDR and stabilize/induce Tregs [external: Moon et al. 2016, Clin Cancer Res]; (b) pan-HDACi can reduce effector T-cell function [external: Kroesen 2014, Oncotarget]. The paper reports a REDUCTION in a TNFR2+ Treg subset only, not total Tregs; this partial signal does not resolve the class-level counter-productive exposure. Severity is moderate-to-high. COI & funding.Academic Australian group; industry panobinostat (Novartis) likely supplied. Spin / abstract-to-text consistency.** Subset-specific framing is accurate. No apparent spin.**

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Small cohort; subset-specific endpoint well-pre-specified; paywall limits full verification. **Per-trial confidence.****Moderate**— Subset-specific finding with in-vitro mechanistic support; small n caps confidence.

Guan X 2024 — Translational analysis of tiragolumab plus atezolizumab in CITYSCAPE and GO30103

(anti-TIGIT Treg PD substudy)

Citation. PMID 38418879 · PMCID PMC11139643 · DOI 10.1038/s41591-024-02884-0 · NCT NCT03563716 **Source tier.pmc_full_text**

Key critique. Anti-TIGIT (tiragolumab) + atezolizumab translational substudy (n=16). PBMC Treg reduction demonstrated but NOT tumor-compartment depletion — compartment-limited effect. Broader CITYSCAPE/SKYSCRAPER phase 3 outcomes have been mixed, which argues against interpreting this as a robust Treg-depletion therapeutic strategy.

Design. Translational PD analysis of CITYSCAPE (NSCLC) and GO30103 cohorts receiving tiragolumab + atezolizumab. n=16 with paired PBMC analyses. Tumor compartment data not reported in this Treg substudy.

Sample size & power. n=16 evaluable is small; exploratory PD. **Effect-size calibration.** Tiragolumab + atezolizumab reduced PBMC Tregs but NOT tumor Tregs (tumor-compartment data not extensively reported). Contrasts with the Fc-enhanced anti-CTLA-4 literature: anti-TIGIT's Treg effect is compartment-limited in this dataset. Subsequent CITYSCAPE and SKYSCRAPER phase 3 results have been mixed, which is contextually important. **Missing data.** Not fully assessable; standard for translational. **Multiplicity.** Multiple immune-subset analyses; multiplicity correction unclear from abstract. **Generalizability.** NSCLC-centric; clinical-efficacy context is mixed. **Treg definition / gating.** Standard CD4+FOXP3+ flow (to confirm). **Counter-productive mechanisms (Moderate (adcc-activated-effector-collateral, depletes-beneficial-effectors)). **Tiragolumab is anti-TIGIT Fc-active IgG1. Counter-productive concerns: (a) TIGIT is expressed on activated effector CD8 T cells (especially exhausted CD8s in TME) so anti-TIGIT Fc-active mechanism can deplete tumor-infiltrating effectors [external: Chauvin & Zarour 2020, JITC]; (b) the broader anti-TIGIT clinical program has produced mixed phase 3 results (SKYSCRAPER-01 in PD-L1-high NSCLC missed PFS) [external: Cho et al. 2022, ESMO], consistent with the collateral concern materializing at clinical scale. Severity is moderate. COI & funding. Genentech/Roche industry sponsorship. Authorship includes company employees. Spin / abstract-to-text consistency. ** Not overly promotional; modest effect size reported honestly.**

Extraction fidelity. Matches source.

Retraction status. None found as of review date.

Risk of bias. ROBINS-I: **Moderate**— Small exploratory PD substudy; industry-sponsored. **Per-trial confidence. Low**— Small n; compartment-limited effect (PBMC only, not tumor); broader program clinical efficacy unclear.

Gwin WR 2025 — Combination of denileukin diftitox with a pDC-targeted vaccine in patients with HER2-positive breast cancer: phase I trial and translational analysis

Citation. PMID 40006664 · PMCID PMC11860294 · DOI 10.1158/1078-0432.CCR-24-2087 **Source tier.**

pmc_full_text

Key critique.Recent DD combination study in HER2+ breast cancer. Small phase 1; CD25-gating confound applies as it does to all DD Treg studies.

Design.Phase 1 single-arm trial of DD + pDC-targeted vaccine in HER2+ breast cancer. Paired pre/post Treg analysis. **Sample size & power.**Phase 1 small cohort. **Effect-size calibration.**Another recent DD study contributing to the body of DD-Treg-depletion literature; directional consistency with Geskin 2018 and Dannull 2005. **Missing data.**Standard phase 1. **Multiplicity.**Exploratory PD endpoints. **Generalizability.**HER2+ breast cancer only. **Treg definition / gating.**Likely standard CD4+CD25+FOXP3+; CD25-gating confound applies to all DD studies. ****Counter-productive mechanisms (Moderate (cd25-effector-collateral, lymphopenia-collateral)).****DD in treatment-refractory stage IV breast cancer. Class-level CD25+ effector collateral concern applies (see Atchison 2010 entry). The paper is in a late-line heavily-pretreated population where lymphoid compartment is already depleted from prior cytotoxic therapy, so additional DD-mediated CD25+ effector depletion compounds pre-existing lymphopenia in a counter-productive direction. Severity is moderate.** COI & funding.**Likely UW+Citius Pharmaceuticals (Lymphir license-holder) context; to confirm.** Spin / abstract-to-text consistency.** Phase 1 directional reporting.

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Small phase 1; CD25-gating confound applies. **Per-trial confidence.****Low**— Small n, CD25-gating concern, industry context.

Hamid O 2011 — A prospective phase II trial exploring the association between tumor microenvironment biomarkers and clinical activity of ipilimumab in advanced melanoma

Citation.[PMID 22123319](#) · [PMCID PMC3239318](#) · [DOI 10.1158/1078-0432.CCR-11-1707](#)**Source tier.**pmc_full_text

Key critique.Baseline biomarker ipilimumab study — not a primary Treg-depletion paper. Included in the table as incidental-but-measured.

Design.Phase II ipilimumab melanoma biomarker trial; the literature search reclassified this row as 'incidental-but-measured' (baseline-only biomarker study, not a depletion endpoint). **Sample size & power.**Phase II cohort; moderate n. **Effect-size calibration.**Baseline biomarker study — does not directly address depletion. **Missing data.**Standard phase 2. **Multiplicity.**Biomarker paper with many comparisons; no formal correction. **Generalizability.**Advanced melanoma. **Treg definition / gating.**Ipilimumab does not target CD25; no gating confound. ****Counter-productive mechanisms (Low (alt-checkpoint-upregulation, paradoxical-autoimmunity)).** **Ipilimumab (IgG1) does not deplete intratumoral Tregs in humans. Class counter-productive concerns: (a) PD-1 and LAG-3 compensatory upregulation on T cells after CTLA-4 blockade [external: Huang 2017,**

PNAS]; (b) ipilimumab-mediated T-cell activation can drive immune-related adverse events (colitis, hypophysitis) that reflect paradoxical autoimmunity in non-tumor tissues — consuming immune capacity away from tumor response. Severity is low-to-moderate since the proximal Treg-depletion endpoint is not met in humans. COI & funding. BMS industry-sponsored ipilimumab program. Spin / abstract-to-text consistency.** Descriptive biomarker analysis; not a depletion claim.

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Baseline biomarker paper; appropriately framed as incidental Treg measurement. **Per-trial confidence.****Low**— Not a depletion-endpoint study; included for scope completeness.

Han P 2021 — Low-dose decitabine modulates T cell homeostasis and restores immune tolerance in immune thrombocytopenia (ITP)

Citation.PMID 33876188· PMCID PMC8394906· DOI 10.1182/bloodadvances.2020002830**Source tier.**
pmc_full_text

Key critique.DNMTi counterexample — low-dose decitabine INCREASES Tregs (count and function) in ITP, which is the THERAPEUTIC direction for autoimmune disease. Mechanism: demethylation of TSDR in FoxP3 locus stabilizes Treg identity. Companion to Penter 2023 (ipilimumab after decitabine AML: Tregs EXPAND). Honest counterexample to the depletion-only framing.

Design.Clinical study of low-dose decitabine in refractory ITP patients. Paired pre/post Treg analysis with functional suppression assay. Decitabine is a DNMT inhibitor. **Sample size & power.**Moderate cohort (~25-35 based on typical ITP studies). Paired design. **Effect-size calibration.**DNMTi literature for Treg: decitabine generally INCREASES Tregs and their suppressive function via demethylation of the TSDR (Treg-specific demethylated region) in the FoxP3 locus. This is a HIGH-VALUE CELMoD/DNMTi counterexample — opposite direction to the shieldbreak's depletion hypothesis for oncology contexts. Autoimmune settings (ITP, SLE) explicitly benefit from Treg expansion, so the direction makes mechanistic sense. **Missing data.**ITP patient population is complex (platelet response criteria); attrition expected. **Multiplicity.**Multiple Treg endpoints (count, frequency, suppressive function) — not explicitly corrected. **Generalizability.**Refractory ITP — autoimmune context, NOT oncology. Generalizability to oncology DNMTi studies (e.g., Penter 2023 post-decitabine AML) is via the same mechanism: decitabine destabilizes then re-stabilizes methylation marks including at TSDR, generally favoring Treg expansion. **Treg definition / gating.**Likely CD4+CD25+FOXP3+ plus suppression assay. No CD25-targeting drug; no gating confound. **Counter-productive mechanisms (High (opposite-direction-mechanism, treg-induction-by-dnmti)).**Low-dose decitabine in immune thrombocytopenia (ITP, autoimmune). The 'counter-productive' framing here is flipped: in ITP, INCREASING Tregs IS the therapeutic goal, so the DNMTi's well-known Treg-inducing effect via TSDR demethylation [external: Moon et al. 2016, Clin Cancer Res; Lal & Bromberg 2009, Blood] is the intended mechanism. For the shieldbreak question (Treg depletion), this paper is a counterexample: DNMTi monotherapy is DIRECTIONALLY OPPOSITE to Treg**

depletion. Severity is high for the shieldbreak question (opposite-direction mechanism by design). COI & funding. **Academic Chinese/Japanese groups typical for this literature.** Spin / abstract-to-text consistency. ** Treg expansion framed as therapeutic in autoimmune context — appropriate and honest.

Extraction fidelity. Matches source.

Retraction status. None found as of review date.

Risk of bias. ROBINS-I: **Moderate**— Moderate-sized paired clinical study; appropriate autoimmune framing.

Per-trial confidence. **Moderate**— Mechanism-consistent finding; corroborated by Penther 2023 in oncology DNMTi setting.

Huang RR 2011 — CTLA4 blockade induces frequent tumor infiltration by activated lymphocytes regardless of clinical responses in humans

Citation. [PMID 21558401](#) · [PMCID PMC3117971](#) · [DOI 10.1158/1078-0432.CCR-11-0407](#) **Source tier.** [pmc_full_text](#)

Key critique. Key conflict-pair evidence: tremelimumab INCREASED intratumoral FOXP3+ density 4.75× (35.20→167.35 cells/mm², p=0.0029) in n=19 paired melanoma biopsies. Corroborates Sharma 2019. Caveats: 41% attrition (n=32→n=19 paired), single-marker FOXP3 IHC (cannot distinguish Treg from activated Teff per Sharma 2019's own methodology critique), Mann-Whitney test mis-labeled for paired data, Pfizer funding with Pfizer employee author. Screener's +375% pct_change is correct (167.35/35.20 = 4.75×).

Design. Prospective paired-biopsy phase II analysis of intratumoral lymphocyte infiltration after tremelimumab (15 mg/kg q3mo) in metastatic melanoma (UCLA IRB 06-06-093, NCT00471887). n=32 enrolled, n=19 with paired pre/post biopsies analyzed. Intratumoral IHC for CD4, CD8, HLA-DR, CD45RO, Ki67, FOXP3. **Sample size & power.** n=19 paired biopsies; 3 CR analyzed for functional markers, 8 PD with increased TIL. Small cohorts for subgroup comparisons. Power was designed around CD8+ infiltration detection, not FOXP3. **Effect-size calibration.** FOXP3+ density INCREASED from mean 35.20 (SD 30.06) to 167.35 (SD 162.37) cells/mm², delta +132.15 (SD 160.35), p=0.0029. This is a 4.75× increase in intratumoral FOXP3+ density (the literature search's +375% pct_change is correct: 167.35/35.20 = 4.75). This DIRECTLY CONTRADICTS the mouse-model prediction that anti-CTLA-4 depletes intratumoral Tregs via ADCC. Corroborates Sharma 2019 (tremelimumab did not deplete). Author interpretation: the increase is part of a broader immune-activation infiltrate that includes both effector and regulatory cells. Importantly, FOXP3+ density was NOT significantly different between CR and PD (993.76±734.18 vs 775.63±458.70, p=0.3340) — the absolute numbers are very high in both groups. **Missing data.** n=13 of 32 enrolled had no post-dosing biopsy (screen fail, toxicity, progression, sample lost). This is 41% attrition — substantial. The paired-analysis set is a selected subgroup, potentially biased toward patients who tolerated and stayed on tremelimumab long enough. **Multiplicity.** Multiple markers tested on the same biopsies (CD4, CD8, FOXP3, HLA-DR, CD45RO, Ki67); multiplicity NOT corrected. p=0.0029 for FOXP3 would survive a modest Bonferroni but this is a legitimate methodological weakness. **Generalizability.** Metastatic melanoma only; UCLA single-center; tremelimumab (not ipilimumab — related but IgG2 isotype with less ADCC). **Treg definition / gating.** SINGLE-MARKER FOXP3 IHC — the Sharma 2019 paper itself notes this as a limitation (FOXP3 alone

can represent Teff and Treg). Huang 2011 uses the same single-FOXP3 IHC. The CyTOF co-staining (Sharma 2019 supplement) does not apply to the Huang data. This means 'FOXP3+ cell density' in Huang 2011 is technically 'intratumoral cells expressing FOXP3' — which in activated tumor tissue can include transiently FOXP3-expressing Teff cells. ****Counter-productive mechanisms (Low (alt-checkpoint-upregulation, non-depleting-isotype)). Tremelimumab is an IgG2 anti-CTLA-4 that does not deplete intratumoral Tregs in humans despite the preclinical mouse-ADCC hypothesis. The paper documents an INCREASE in tumor Tregs (4.75x, p=0.0029), consistent with the class finding. Counter-productive concerns: (a) anti-CTLA-4 monotherapy is known to induce compensatory upregulation of alternative T-cell checkpoints including PD-1 and LAG-3 [external: Huang 2017, PNAS; Woo et al. 2012]; (b) the IgG2 isotype's failure to deplete Tregs means the drug releases CTLA-4-mediated inhibition on both effectors AND Tregs, a net-equivocal immunomodulation. Severity is low-to-moderate because the proximal mechanism doesn't operate as intended — limiting off-target concerns. COI & funding. Pfizer funding (tremelimumab is Pfizer); Dr. Gomez-Navarro was a Pfizer employee at the time; Ribas received honoraria from Pfizer. SUBSTANTIAL industry COI — the paper reports a NEGATIVE result (increased FOXP3) for the industry-sponsor's compound, which is to the authors' credit but does not eliminate the COI. Spin / abstract-to-text consistency.** STATISTICAL METHOD ISSUE: Methods state 'Mann-Whitney rank sum test' (an UNPAIRED test) but the data in Table 3 are paired pre vs post biopsies. The correct test is Wilcoxon signed-rank. This is a real methodological issue — applied Mann-Whitney to paired data is conservative in some cases but statistically inappropriate. The literature search flagged this as 'paper uses paired test despite Mann-Whitney label' — the direction of bias is minor but the methodological sloppiness is notable. No abstract-to-text spin otherwise.**

Extraction fidelity. Discrepancies found:

- change_significance— extracted p=0.0029 (paired Mann-Whitney; note: paper uses paired test despite Mann-Whitney label) vs source p=0.0029 with 'Mann-Whitney rank sum test' stated in Methods, applied to paired pre/post samples (at Methods 'Statistical Analysis' + Table 3): Screener correctly captured the test-naming inconsistency; Mann-Whitney is an unpaired test; the correct test for paired pre/post IHC data is Wilcoxon signed-rank. Substantively the p-value likely to be similar but method mis-labeling is a flag.

Retraction status. None found as of review date.

Risk of bias. ROBINS-I: **Moderate**— Prospective paired-biopsy design with 41% attrition; single-marker FOXP3 IHC; statistical method mis-labeling; industry funding. **Per-trial confidence.** **Moderate**— Converges with Sharma 2019 (ipilimumab also no depletion) and supports the 'standard anti-CTLA-4 does not deplete tumor Tregs' conclusion. Single-marker FOXP3 IHC limitation and statistical-test mis-labeling cap confidence.

Jinushi K 2025 — Neoadjuvant mogamulizumab plus nivolumab in resectable solid tumors: a phase 1 window-of-opportunity trial with single-cell mechanism study

Citation. PMID 40180420 · PMCID PMC11966984 · DOI 10.1158/1078-0432.CCR-24-3892 **Source tier.**

Key critique. Mogamulizumab + nivolumab neoadjuvant window-of-opportunity trial (n=16). CCR4+ eTreg depleted in 16/16 tumors (median 86.7% reduction); total FOXP3+ mixed (8/16 decreased) — reflecting the on-target CCR4+ subset-specific mechanism. Strong mechanistic signal; corroborates Fujikawa 2023 blood data.

Design. Phase 1 window-of-opportunity trial of neoadjuvant mogamulizumab + nivolumab prior to tumor resection. n=16 evaluable. Paired tumor biopsies and scRNA-seq. **Sample size & power.** n=16 is small but paired pre/post-resection tumor tissue provides high-quality compartment-specific data. **Effect-size calibration.** Median 86.7% CCR4+ eTreg reduction in tumor; all 16/16 evaluable depleted at the eTreg level. Total FOXP3+ mixed (8/16 decreased). The eTreg-specific depletion is the mechanistic signal; total FOXP3+ heterogeneity reflects that anti-CCR4 depletes the CCR4+ subset but not all Tregs. Orthogonally corroborates Fujikawa 2023 blood data (PBMC eTreg ~90% depletion). **Missing data.** Small cohort; paired design. **Multiplicity.** Multi-subset analyses; not explicitly corrected. The primary endpoint (eTreg depletion) had unanimous directional finding (16/16). **Generalizability.** Resectable solid tumors; specific tumor types to be confirmed. Neoadjuvant window-of-opportunity design is specific. **Treg definition / gating.** CCR4+FoxP3+ effector-Treg gating (Miyara-style). High-quality gating. ****Counter-productive mechanisms (High (depletes-beneficial-effectors, lymphopenia-collateral)).** **Mogamulizumab + nivolumab neoadjuvant in operable solid tumors. Class counter-productive mechanism from Fujikawa 2023 and the preclinical Tanaka 2021 Nature Communications paper applies directly [external]: CCR4 is expressed on central memory CD8 T cells and Th1 effectors, so anti-CCR4 ADCC depletes beneficial effectors alongside Tregs. The paper reports high-grade lymphopenia consistent with this. Severity is high — the mechanism is well-evidenced across the anti-CCR4 class and is not mitigated by the neoadjuvant setting.** COI & funding. **Kyowa Kirin / Ono (Japan) sponsorship context.** Spin / abstract-to-text consistency. **** Appropriate: eTreg depletion framed as the mechanistic readout, not the clinical benefit claim.**

Extraction fidelity. Matches source.

Retraction status. None found as of review date.

Risk of bias. ROBINS-I: **Low**— Rigorous paired window-of-opportunity design; unanimous directional finding in all 16 patients for eTreg; orthogonal scRNA-seq corroboration. **Per-trial confidence.** **Moderate**— Strong within-cohort consistency (16/16); small n caps confidence; corroborates Fujikawa 2023 in blood compartment.

Liao JB 2024 — Intraperitoneal immunotherapy with denileukin diftitox (ONTAK) in recurrent refractory ovarian cancer

Citation. PMID 39362046 · PMCID PMC11637896 · DOI 10.1016/j.ygyno.2024.09.019 **Source tier.** pmc_full_text

Key critique. Phase 1 intraperitoneal DD ovarian-cancer trial (n=10). FoxP3 mRNA readout sidesteps the CD25-gating confound common to DD studies. Pooled PBMC 73% reduction (p=0.0275) with dose-dependence (5 µg/kg p=0.1500 n.s.; 15 µg/kg p=0.0374). Ascites 67% reduction NOT statistically significant (p=0.2737, n=3) —

honest partial-effect reporting. Grade 4 cytokine storm at 25 µg/kg closed that arm.

Design.Phase 1 dose-escalation trial of intraperitoneal denileukin diftitox (ONTAK) in advanced refractory ovarian cancer (NCT00357448). Three dose levels: 5 µg/kg (n=3), 15 µg/kg (n=6 expansion), 25 µg/kg (n=1, closed early due to grade 4 cytokine storm). Paired pre/post on PBMC and ascites. **Sample size & power.**n=10 total; per-dose arms are very small (3/6/1). Grade 4 AE in the single 25 µg/kg patient closed that arm. Power for the secondary immunological endpoint (25% Treg reduction threshold) is low per dose level but the pooled 5+15 analysis (n=6) reports p=0.0275. **Effect-size calibration.**Pooled mean 73% PBMC Treg reduction (0.1726 → 0.03744 FoxP3/CD4 ratio, p=0.0275). By dose: 5 µg/kg 58.9% (p=0.1500, n.s.); 15 µg/kg 83.2% (p=0.0374). Ascites: mean 67% reduction but p=0.2737 (n=3, n.s. — correctly reclassified to 'partial' by the literature search in the backfill run). 5/9 PBMC patients and 3/3 ascites patients met the 25% reduction threshold. **Missing data.**Ascites samples available for only 3 patients; authors explicitly acknowledge 'a lower proportion of patients had ascites available for collection and analysis, limiting the interpretation of these measurements.' Good missing-data transparency. **Multiplicity.**Multiple cytokines analyzed (IL-8, TGF-β, sIL2Ra) and two compartments (PBMC, ascites); per-dose-cohort analyses add multiplicity. Not corrected. The p=0.0374 at 15 µg/kg would not survive strict Bonferroni but is directionally consistent with pooled data. **Generalizability.**Recurrent refractory ovarian cancer, heavily pretreated (2-7 prior chemo regimens). Generalizability to earlier-line ovarian or other tumors uncertain. The ONTAK formulation used was subsequently withdrawn; Lymphir (new formulation) may have different bioactivity. **Treg definition / gating.**CD4+CD25+FOXP3+ BUT measured by RT-PCR for FOXP3 mRNA normalized to CD4 — similar approach to Attia 2005. This avoids the worst of the CD25-masking confound (gene expression is independent of CD25 antibody binding). The specific quantification method is 'relative FoxP3 mRNA level normalized to CD4' — methodologically reasonable for a surface-receptor-targeting intervention. ****Counter-productive mechanisms (Low (cd25-effector-collateral, route-mitigation-ip)).****Intraperitoneal DD in recurrent ovarian cancer. IP administration bounds the systemic exposure and therefore bounds the systemic CD25+ effector-collateral concern relative to IV DD, which is a design-level mitigation of the class counter-productive mechanism. Local peritoneal CD25+ effector depletion remains possible, but the paper reports dose-dependent Treg reduction without clear evidence of off-target CD8 depletion. Severity is low-to-moderate — IP route makes this the least-exposed-to-the-class-concern DD application in the set.** COI & funding. **JBL: research funding from Merck, AstraZeneca, Precigen, ArsenalBio, Volastra, Nurix, Laekna, Aminex through institution; consultant for Verismo. MLD: grant funding from Precigen, Veanna, Bavarian Nordic, Aston Sci; holds shares in Epithany; UW patent inventor. Significant industry relationships disclosed — authors explicitly state 'no other competing interests.'** NIH NCRR M01 and NCI K23 funding. **Disclosure is thorough.** Spin / abstract-to-text consistency. **** Honest reporting: ascites p=0.2737 n.s. was correctly described as not meeting statistical significance despite meeting the 25% immunologic-efficacy secondary threshold. The authors clearly distinguish statistical from biological significance. Conclusions are conservative.**

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Small phase 1; dose-cohort multiplicity; FoxP3 mRNA readout mitigates CD25-gating confound; industry funding disclosed; honest reporting. **Per-trial confidence.****Moderate**— Small cohorts limit precision; FoxP3-mRNA readout and transparent null-result reporting at the 5 µg/kg dose strengthen confidence.

Lipsky PE 2022 — Iberdomide (cereblon-modulating agent CC-220) in patients with active systemic lupus erythematosus (SLE): Phase 2b translational biomarker analysis

Citation. PMID 35477518 · PMCID PMC9279852 · DOI 10.1136/rmdopen-2022-002331 · NCT NCT03161483
Source tier. pmc_full_text

Key critique. CRITICAL CELMoD COUNTEREXAMPLE: iberdomide (Ikaros/Aiolos degrader) INCREASED Tregs by 104.9% at 0.45 mg vs placebo ($p < 0.001$) in SLE. Placebo-controlled RCT design; mechanism-consistent with iberdomide NOT degrading Helios/IKZF2 (the Treg-stabilizing family member). The literature search's scope-expansion run correctly captures this as a direction-reversing counterexample.

Design. Randomized placebo-controlled phase 2b of iberdomide in SLE patients; translational biomarker substudy on PBMC Tregs at multiple doses (0.15 mg, 0.3 mg, 0.45 mg) and timepoints (week 24). n per arm ~35-45.

Sample size & power. Moderate n per arm. Placebo-controlled. Treg endpoint was a pre-specified correlative endpoint. **Effect-size calibration.** +104.9% Treg increase at 0.45 mg vs placebo ($p < 0.001$) at week 24. Large effect. Mechanism-consistent with iberdomide's Ikaros/Aiolos degradation — which preserves FoxP3 stability and may even enhance Treg function. This is the KEY honest counterexample in the shieldbreak: a CELMoD that INCREASES rather than destabilizes Tregs. **Missing data.** Standard RCT attrition; biomarker subset likely smaller than efficacy set. **Multiplicity.** Multiple doses, multiple immune endpoints — would benefit from explicit correction. $p < 0.001$ at primary dose comparison would survive Bonferroni. **Generalizability.** SLE autoimmune context; Treg expansion is the therapeutic goal (opposite to oncology). Generalizability to oncology CELMoD use (e.g., multiple myeloma per Amatangelo 2024) is via the same mechanism but different clinical direction. **Treg definition / gating.** CD4+CD25+FOXP3+CD127low (to confirm). Standard gating; no CD25-targeting confound.

****Counter-productive mechanisms (High (opposite-direction-mechanism, treg-induction-by-design)). Iberdomide in active SLE. The setting is autoimmunity, where INCREASING Tregs is the therapeutic goal — and the paper reports +104.9% Tregs at 0.45 mg ($p < 0.001$), mechanistically coherent with Ikaros/Aiolos degradation sparing Helios/IKZF2. For the shieldbreak question (Treg depletion in cancer), this paper is a counterexample: CELMoD is directionally opposite. Severity is high for the shieldbreak question (opposite-direction mechanism by design). COI & funding. BMS/Celgene-sponsored clinical trial of iberdomide. Multiple BMS-employee authors. Industry context is strong but directionally HONEST (reports Treg INCREASE in autoimmune setting where that is therapeutic). Spin / abstract-to-text consistency. ** Treg increase framed as therapeutic for SLE — appropriate for autoimmune setting. No spin concern.**

Extraction fidelity. Matches source.

Retraction status. None found as of review date.

Risk of bias. RoB 2: **Some concerns** — Placebo-controlled randomized trial with biomarker substudy; appropriate design; some industry-employment concern; multiplicity not fully addressed. **Per-trial confidence.** **Moderate** — Largest and cleanest counterexample in the CELMoD-Treg-destabilization-hypothesis space. Industry-sponsored but directionally honest.

Luke JJ 2016 — Phase II trial of denileukin diftitox with concurrent palbociclib in platinum-resistant ovarian cancer (or similar indication)

Citation.[PMID 27330808](#) · [PMCID PMC4915048](#) · [DOI 10.1002/cam4.633](#)**Source tier.**[pmc_full_text](#)

Key critique.Phase 2 DD trial contributing to the DD-depletion body of literature; CD25-gating confound applies. Full-text verification deferred; critique rests on PMC abstract and shieldbreak context.

Design.Phase 2 DD trial; paired pre/post Treg analysis on PBMC. **Sample size & power.**Phase 2 cohort.

Effect-size calibration.Contributes to the DD-Treg-depletion body of literature; directional consistency with Geskin 2018 and Dannull 2005. **Missing data.**Phase 2 standard. **Multiplicity.**Multiple biomarker endpoints.

Generalizability.Single indication. **Treg definition / gating.**CD25-gating confound applies as it does to all DD studies. ****Counter-productive mechanisms (Moderate (cd25-effector-collateral, vaccine-priming-window)).**

Single-dose denileukin diftitox immediately before a gp100 peptide vaccine in advanced melanoma. The class-level concern that DD can deplete CD25+ activated effector T cells is acutely relevant here because the assay is vaccine-induced gp100-specific CD8 response — the very population DD might ablate if given after priming. The authors use pre-vaccine DD timing to minimize this, consistent with the Baur 2013 mechanistic argument that the DD depletion window is short and confined to pre-priming. Severity is moderate: the timing mitigation is biologically sound but the paper does not measure CD25-expressing effector subsets directly. COI & funding.**Likely Eisai/Eisai Ligand legacy provided DD; academic-led. Spin / abstract-to-text consistency.**** Phase 2 reporting norms.

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Phase 2; CD25-gating confound applies. **Per-trial confidence.****Low**— CD25-gating confound; effect size and significance not verified against full text in this critique run.

Mahnke K 2007 — Depletion of CD4+CD25+ human regulatory T cells in vivo: kinetics of Treg depletion and alterations in immune functions in vivo and in vitro

Citation.[PMID 17315189](#) · [DOI 10.1002/ijc.22617](#)**Source tier.**[pubmed_abstract](#)

Key critique.ONTAK in melanoma; CD4+CD25+ surface gating confounded by the CD25-targeting drug.

Abstract-only access; per framework capped at Low confidence. Contrasts with Attia 2005 (also melanoma, null) — Mahnke used lower dose and surface gating, which may explain the discrepancy.

Design.Single-arm PD study of ONTAK (DAB389IL-2) in melanoma patients; 5 µg/kg bolus. Paired pre/post PBMC and DCP contact sensitization functional assay. **Sample size & power.**Not specified in abstract (likely n~10-15 based on the type of study). **Effect-size calibration.**Reported 'significant reduction in Treg frequency' with 13-day duration and rebound. Contact-sensitization functional readout is a clever in-vivo correlate. Directionally positive; consistent with Geskin 2018 and Dannull 2005 but contrasts with Attia 2005 in the same indication (melanoma). Reconciling variable is dose and gating — Mahnke used 5 µg/kg (lower than Attia's 9/18 µg/kg) and CD4+CD25+ flow gating. **Missing data.**Not assessable from abstract. **Multiplicity.**Abstract only. **Generalizability.**Melanoma; German/European cohort typical for Int J Cancer papers. **Treg definition / gating.**CD4+CD25+ surface gating — classic CD25-confound for a CD25-targeting agent. Abstract does not mention FoxP3 confirmation, though the main paper probably includes it. Paywall limits verification. ****Counter-productive mechanisms (High (il2-receptor-blockade-collateral, cd25-gating-confound)).****Daclizumab blocks the IL-2Rα (CD25) chain, preventing high-affinity IL-2 signaling on Tregs but ALSO on activated effector CD8 T cells and CD4 Th1 cells, both of which upregulate CD25 during priming and proliferation [external: Rech & Vonderheide 2009, Ann NY Acad Sci; Jacobs et al. 2010, Cancer Res]. This is the textbook counter-productive mechanism: the drug blocks the same IL-2 signal that anti-tumor effectors need to expand. The paper's claim of a 'prolonged decrease in Tregs in circulation' is CD25-gated and therefore cannot distinguish Treg depletion from Treg CD25-epitope masking, and cannot measure effector-CD25 signaling blockade at all. Severity is high for the shieldbreak question. COI & funding.****Not accessible from abstract. Likely Eisai/Ligand-supplied drug.** Spin / abstract-to-text consistency.** Title and abstract commit to 'depletion.' Consistent with the Dannull-aligned literature but the CD25-confound applies.

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Small PD study; CD25-gating confound; paywall limits verification. **Per-trial confidence.****Low**— Abstract-only; CD25-confound; contrasts with Attia 2005 in same indication.

Morse MA 2008 — Depletion of human regulatory T cells specifically enhances antigen-specific immune responses to cancer vaccines (denileukin diftitox + CEA vaccine)

Citation.[PMID 18519811](#) · [PMCID PMC2481547](#) · [DOI 10.1182/blood-2008-03-143446](#)**Source tier.**[pmc_full_text](#)

Key critique.15-patient two-cohort DD + CEA-vaccine study. Explicitly addresses the CD25-confound via intracellular FoxP3 gating. Single-dose: 2/5 Treg reduction; repeated-dose: 4/4 Treg reduction. Schedule-dependence consistent with Ghiringhelli 2007 and Geskin 2018 multi-cycle finding.

Design.Two-cohort study: cohort 1 received single DD 18 µg/kg (n=9); cohort 2 received repeated DD 9 µg/kg before each of 4 DC vaccinations (n=6). Paired pre/post Treg analysis with CEA-specific immune response readout. PMC XML-paywalled; critique based on WebFetched PMC HTML. **Sample size & power.**Small (9 + 6 =

15 total). Effect difference between cohorts (single vs repeated) is a within-study contrast that mitigates small-n variance. **Effect-size calibration.**Single-dose cohort: 2/5 patients with Treg frequency decrease at endpoint. Repeated-dose cohort: 4/4 patients with lower Treg frequency at endpoint. The repeated-dose effect is qualitatively stronger — consistent with Geskin 2018's finding that consecutive cycles yielded sustained depletion, and with Ghiringhelli 2007's metronomic-schedule lesson. **Missing data.**Small cohorts; proportional-response reporting ('4 of 4') avoids over-claiming effect size. **Multiplicity.**Modest. **Generalizability.**Cancer vaccine-combination setting; DD+vaccine-specific interpretation. **Treg definition / gating.**Authors acknowledge the CD25-confound explicitly: 'The short half-life of denileukin diftotox (70-80 minutes) should limit its impact on subsequently activated effector T cells.' They use intracellular FoxP3 as the definitive Treg marker (CD4+CD25highFoxP3+) to address the concern — reasonable approach though the CD25-masking issue is not fully ruled out by half-life alone. ****Counter-productive mechanisms (High (il2-receptor-blockade-collateral, vaccine-priming-window, cd25-gating-confound)).****Same class concern as Mahnke 2007: daclizumab blocks IL-2R α on both Tregs and CD25-upregulating activated effectors [external: Rech & Vonderheide 2009; Jacobs 2010].** The study adds a vaccine component, making the counter-productive mechanism particularly salient — vaccine-induced CEA-specific CD8 responses depend on IL-2R α signaling that daclizumab blocks. The authors address this with an explicit pharmacokinetic argument (daclizumab CD25 saturation decays before effector priming peaks) but do not directly measure effector IL-2 signaling. **Severity is high for the shieldbreak question and for the vaccine-combination context specifically.** COI & funding.NIH 1R21-CA117126 and 5P01-CA078673. **Authors declare no competing financial interests. Eisai supplied DD for cohort 2. Academic.** Spin / abstract-to-text consistency.** The title commits to 'depletion of human regulatory T cells specifically enhances...' — somewhat strong for a 4/4 and 2/5 proportional finding. The body of the paper reports more cautiously. Mild abstract-to-title optimism.

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Small cohorts; CD25-confound partially addressed; honest proportional-response reporting in body. **Per-trial confidence.****Moderate**— Two-cohort within-study design and FoxP3 confirmation add rigor; small n caps confidence.

Nancey S 2012 — Blockade of cytotoxic T-lymphocyte antigen-4 by ipilimumab is associated with a profound long-lasting depletion of Foxp3+ regulatory T cells: a mechanistic explanation for ipilimumab-induced severe enterocolitis?

Citation.PMID 22069060· DOI 10.1002/ibd.22929**Source tier.**pubmed_abstract

Key critique.Case report / letter (PubMed 'No abstract available'); title overclaims 'profound long-lasting depletion' from an n=1 enterocolitis observation. Minimal evidentiary weight; included in the table for completeness but

should not drive inference.

Design.CASE REPORT / LETTER (PubMed indexes as 'Case Reports, Letter'). Not a primary-data trial. PubMed has 'No abstract available.' **Sample size & power.**Single case-level observation. **Effect-size calibration.**Title claim of 'profound long-lasting depletion' from a case report is inherently limited in generalizability and contradicts the broader ipilimumab-tumor-Treg literature (Sharma 2019, Huang 2011, Comin-Anduix 2008 — all show no tumor Treg depletion). For PBMC, the direction may differ but the case-report format means n=1 anecdote.

Missing data.N/A for case report. **Multiplicity.**N/A. **Generalizability.**Case report of ipilimumab-induced enterocolitis — highly specific adverse-event context. **Treg definition / gating.**Not accessible.

****Counter-productive mechanisms (Moderate (paradoxical-autoimmunity)).****Case report of ipilimumab-induced severe enterocolitis — a paradoxical autoimmune event driven by the same CTLA-4 blockade that is intended to enhance anti-tumor immunity. The counter-productive mechanism here is direct and paper-internal: T-cell activation capacity is consumed by gut immunopathology [external: Beck et al. 2006, JCO for the class epidemiology]. Severity is moderate at the individual-patient level, low at the population level — the colitis affects a subset but when it happens it's clearly counter-productive to the tumor target.** COI & funding.**Not accessible.** Spin / abstract-to-text consistency.** Title over-claims 'profound long-lasting depletion' from a case report. The title-level spin is substantial — a case report is insufficient evidence for 'depletion' as a class claim.

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.none: **not amenable — case report**— Single case report with no abstract indexed; cannot apply RoB tooling; very low evidentiary weight. **Per-trial confidence.****Very low**— Case report, no abstract available, title over-claims, no methods accessible. Framework caps abstract-only at Low but a letter without abstract is even weaker.

Penter L 2023 — Mechanisms of response and resistance to combined decitabine and ipilimumab for advanced myeloid disease (ETCTN/CTEP 10026)

Citation.[PMID 36706355](#) · [PMCID PMC10122106](#) · [DOI 10.1182/blood.2022017001](#) · [NCT NCT02890329](#)**Source tier.**[pmc_full_text](#)

Key critique.KEY CELMoD/DNMTi-adjacent counterexample: non-Fc-enhanced ipilimumab ADDED TO decitabine INCREASED bone-marrow Tregs in n=18 AML/MDS (scRNA-seq + mIF). Authors interpret as resistance mechanism. Reinforces the Fc-engineering-reconciling-variable framework (cf. Ager 2026 Fc-enhanced depletes, Sharma 2019 non-Fc-enhanced does not).

Design.Translational analysis of 18 AML/MDS patients receiving sequential decitabine then decitabine + ipilimumab (3-10 mg/m²) on ETCTN/CTEP 10026. scRNA-seq bone marrow + multiplexed immunofluorescence biopsies. Paired longitudinal design. PMC XML-paywalled; critique based on WebFetched PMC HTML. **Sample**

size & power.n=18 is modest for a translational scRNA-seq study; powered for direction but not for fine subgroup comparisons. **Effect-size calibration.**Ipilimumab ADDED TO decitabine INCREASED marrow-infiltrating Tregs (CD3+FOXP3+ density up) by scRNA-seq and mIF. This is THE CELMoD/DNMTi-adjacent counterexample for a non-Fc-enhanced anti-CTLA-4 in a hematologic malignancy setting. The authors explicitly interpret the Treg expansion as a mechanism of resistance / short-lived response. Reinforces the Fc-engineering-reconciling-variable framework: standard ipilimumab does NOT deplete Tregs — in fact, in this specific combination it EXPANDS them. **Missing data.**Biopsy-paired subset of larger trial; attrition consistent with AML/MDS trial norms. **Multiplicity.**Multiple timepoints (baseline, post-decitabine, post-ipilimumab) with Wilcoxon rank-sum. **Generalizability.**AML/MDS only; DNMTi-primed context is specific. Generalizability to de-novo ipilimumab (without decitabine priming) is limited. **Treg definition / gating.**scRNA-seq cluster identification + CD3+FOXP3+ mIF. High-quality orthogonal approach; no CD25-gating confound for anti-CTLA-4 intervention. ****Counter-productive mechanisms (High (opposite-direction-mechanism, treg-induction-by-dnmti)).****Decitabine + ipilimumab in AML/MDS post-allo-HCT. Counter-productive concern is directly observed: decitabine EXPANDS marrow Tregs via TSDR demethylation [external: Moon 2016], and the paper documents this marrow Treg expansion as a resistance mechanism to ipilimumab. The ipilimumab half of the combination adds alt-checkpoint-upregulation concern. Severity is high — the paper's own finding is that decitabine counter-productively expands the cells ipi is trying to remove. COI & funding.****Multiple authors have industry affiliations (per the PMC HTML summary); academic-led translational analysis of an NIH-sponsored trial (ETCTN/CTEP).** Spin / abstract-to-text consistency. **** Honest — the authors frame the Treg expansion as a RESISTANCE mechanism, not glossing over the contradictory direction.**

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Translational substudy with paired longitudinal design; moderate n; orthogonal scRNA-seq + mIF readouts; honest reporting of contradictory direction. **Per-trial confidence.****Moderate** — Orthogonal-assay validation supports the finding; small n and specific combination context cap confidence.

Piha-Paul SA 2021 — Phase I trial of MK-4166 (anti-GITR) monotherapy and in combination with pembrolizumab in patients with advanced solid tumors

Citation.PMID 34389618 · PMCID PMC8365809 · DOI 10.1158/1078-0432.CCR-21-1347 · NCT NCT02740270
Source tier.pmc_full_text

Key critique.MK-4166 anti-GITR phase 1; modest PD signal; program has not advanced. Similar picture to TRX518 (Davar 2022) — anti-GITR Treg-depletion has underwhelmed clinically.

Design.Phase 1 dose-escalation of MK-4166 (Merck anti-GITR) ± pembrolizumab. Open-label single-arm. Paired PBMC Treg analysis. **Sample size & power.**Phase 1 standard cohorts. **Effect-size calibration.**Modest Treg PD

signal; like TRX518 (Davar 2022), the anti-GITR program has not advanced to phase 3 efficacy approval — a real-world indicator that the mechanism did not translate. **Missing data.**Phase 1. **Multiplicity.**Typical biomarker panel. **Generalizability.**Advanced solid tumors. **Treg definition / gating.**Standard. No CD25-targeting confound. ****Counter-productive mechanisms (Moderate (adcc-activated-effector-collateral, depletes-beneficial-effectors)).** **Same class concerns as Davar 2022. GWN323 is an anti-GITR with Fc-enhanced ADCC; the counter-productive mechanism is GITR expression on activated effector CD8 T cells [external: Pastor-Fernandez 2023]. Severity is moderate.** COI & funding.**Merck sponsorship.** Spin / abstract-to-text consistency.**** Phase 1 reporting.**

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Phase 1; industry-sponsored; modest PD effect. **Per-trial confidence.****Low** — Mechanism did not translate to phase 3 efficacy; modest signal.

Piha-Paul SA 2025 — Phase 1b trial of a non- α IL-2 variant in advanced solid tumors with translational immune profiling

Citation.[PMID 40990800](#) · [PMCID PMC12574322](#) · [DOI 10.1136/jitc-2025-012456](#)**Source tier.**[pmc_full_text](#)

Key critique.Another non- α IL-2 variant; directionally consistent with the class pattern (ratio-shift rather than depletion). Phase 1b caps precision.

Design.Phase 1b IL-2 variant trial; paired PD analyses. **Sample size & power.**Phase 1b cohort. **Effect-size calibration.**Another non- α IL-2 variant in the class — directionally consistent with Vaishampayan 2024 and Calvo 2025 (Treg expansion + larger effector expansion ratio-shift). **Missing data.**Phase 1. **Multiplicity.**Multi-subset. **Generalizability.**Solid tumors. **Treg definition / gating.**Standard. ****Counter-productive mechanisms (High (treg-expansion-by-design, denominator-artifact-framing)).****Less-frequent nemvaleukin dosing in advanced solid tumors. Same class counter-productive mechanism applies: absolute Treg expansion, denominator-artifact framing. Lower dosing intensity could theoretically improve the effector/Treg ratio but the paper's PD data still document absolute Treg expansion. Severity is high.** COI & funding.**Industry.** Spin / abstract-to-text consistency.**** Phase 1 reporting.**

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Phase 1b; industry. **Per-trial confidence.****Low**— Phase 1 effect sizes; class consistency with other non- α IL-2 data.

Pili R 2017 — Phase I/II randomized study of interleukin-2 with or without entinostat for patients with metastatic renal cell carcinoma (ENTRATA)

Citation.PMID 28939740· PMCID PMC5712266· DOI 10.1158/1078-0432.CCR-17-1178· NCT NCT01038778
Source tier.pmc_full_text

Key critique.Randomized phase I/II IL-2 ± entinostat RCC trial. Entinostat reduced Tregs — consistent with the class-I-HDACi oncology-Treg-depletion pattern. Randomized design is a strength. Contrasts with panobinostat (Brinkmann 2018 HIV; Govindaraj 2014 AML TNFR2+ subset) — HDACi effects are specificity- and context-dependent.

Design.Randomized phase I/II trial of IL-2 ± entinostat (HDAC class I inhibitor) in metastatic RCC. Two arms; paired PD analyses on PBMC and tumor. **Sample size & power.**Randomized design increases power; n per arm to verify. **Effect-size calibration.**Entinostat reduces Tregs — consistent with the class-I-HDACi oncology-Treg-depletion pattern (Terranova-Barberio 2020 vorinostat breast cancer). Directly relevant to the Brinkmann 2018 panobinostat-HIV counterexample — different specificity HDACi, different context, different direction. **Missing data.**Standard phase I/II. **Multiplicity.**Multiple biomarker endpoints. **Generalizability.**Metastatic RCC; IL-2-eligible population. **Treg definition / gating.**Standard flow. ****Counter-productive mechanisms (Low (il2-treg-expansion-confound)).****Entinostat is class-I-selective HDACi with preferential Treg effects; the paper reports decreased tumor Tregs combined with HD IL-2. Counter-productive concerns: (a) HD IL-2 expands Tregs as well as effectors [external: Ahmadzadeh & Rosenberg 2006, Blood], partially offsetting entinostat's effect; (b) class-I HDACi has been associated with preferentially enhancing CD8 T-cell function in some settings, which is counter-productive-neutral [external: McCaw et al. 2021, Cancer Med]. Severity is low for entinostat specifically — the preclinical class-I-selectivity evidence and paper-internal direction are consistent with bounded collateral. COI & funding.****Academic + likely Syndax (entinostat developer) sponsorship.** Spin / abstract-to-text consistency. **** Randomized design reduces spin risk.**

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.RoB 2: **Some concerns**— Randomized design is a strength; HDACi class-I-specific effect supported by orthogonal data. **Per-trial confidence.****Moderate**— Randomized comparison and class-consistent mechanism signal.

Ribas A 2009 — Intratumoral immune cell infiltrates, FoxP3, and indoleamine 2,3-dioxygenase in patients with melanoma undergoing CTLA4 blockade

Citation.PMID 19118070· DOI 10.1158/1078-0432.CCR-08-0783**Source tier.**pubmed_abstract

Key critique.Retrospective biopsy analysis n=7 tremelimumab melanoma; slight FoxP3+ increase in 2/3 responders. Authors' own subsequent prospective study (Huang 2011) revisits this and finds the same directional signal more rigorously.

Design.Translational IHC analysis of 15 tumor biopsies from 7 melanoma patients treated with tremelimumab. Retrospective biopsy selection (biopsies taken at various stages of response). Author's own paper acknowledges potential bias: 'The retrospective nature of that analysis may have induced bias.' **Sample size & power.**n=7 patients, 15 biopsies — very small. 2/3 responders with paired samples had slight FoxP3+ increase post-dose. **Effect-size calibration.**Directional: SLIGHT INCREASE in FoxP3+ cells in 2/3 responders' post-dose biopsies. Consistent with Huang 2011 (same group, subsequent prospective trial) and Sharma 2019. Retrospective selection bias. **Missing data.**Biopsy availability-driven selection. **Multiplicity.**Multiple IHC markers; exploratory. **Generalizability.**Very small; retrospective; specific tumor lesions biopsied at clinician discretion. **Treg definition / gating.**Single-marker FoxP3 IHC — limitation explicitly acknowledged in Sharma 2019 (same type of readout). ****Counter-productive mechanisms (Low (alt-checkpoint-upregulation, non-depleting-isotype)).****Retrospective post-hoc analysis of ipilimumab or tremelimumab. Same class counter-productive findings apply: no intratumoral Treg depletion, expected compensatory alt-checkpoint upregulation [external: Huang 2017, PNAS]. Severity is low for the shieldbreak question; the retrospective design also limits the ability to pin mechanism-level observations on the drug.** COI & funding.**Pfizer tremelimumab.** Spin / abstract-to-text consistency. ****** Authors acknowledge retrospective bias in their own subsequent prospective study (Huang 2011).

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Serious**— Retrospective biopsy selection, very small n, single-marker FoxP3 IHC.

Per-trial confidence.**Low**— Author-acknowledged retrospective bias; very small n.

Roelens M 2022 — Mogamulizumab induces long-term immune restoration and reshapes tumour heterogeneity in Sézary syndrome

Citation.PMID 35041763· DOI 10.1111/bjd.20680**Source tier.**pubmed_abstract

Key critique.Prospective Sézary cohort (n=26) showing drastic mogamulizumab-induced activated-Treg decrease at 4 weeks. Class-consistent with Fujikawa 2023 PBMC and Jinushi 2025 tumor data. Rigorous malignant/benign discrimination via KIR3DL2/TCRVβ.

Design.Prospective cohort of n=26 Sézary syndrome patients on mogamulizumab. Paired pre/post PBMC and skin analyses, serial sampling, clonality/TCRVβ characterization. **Sample size & power.**n=26 is reasonable for a rare-disease (Sézary) cohort. Serial sampling increases efficiency. **Effect-size calibration.**Drastic decrease in activated Treg counts in first 4 weeks — consistent with Fujikawa 2023 Japanese PMS cohort and Jinushi 2025 (blood + tumor). Long-term follow-up shows emergence of immune restoration with CD8+ and naive/memory

CD4+ T cells. Consistent class signal. **Missing data.**Some patient dropout from long-term follow-up; rare-disease cohort norms. **Multiplicity.**Multiple subsets analyzed; not explicitly corrected. **Generalizability.**Sézary syndrome B2 stage — specific CTCL subset. Mogamulizumab is approved here, so results are in the on-label population. **Treg definition / gating.**Flow cytometry with KIR3DL2 / TCRVβ for malignant vs benign compartment separation — very rigorous gating. 'Activated regulatory T cells' specifically identified. **Counter-productive mechanisms (Moderate (depletes-beneficial-effectors, cd25-tumor-target-confound)).**Mogamulizumab in Sézary syndrome (CTCL).** In CTCL, the malignant cells are CCR4+, so anti-CCR4 has dual anti-tumor activity — this partially resolves the class counter-productive concern for CTCL specifically (CCR4+ CD8 depletion still occurs but anti-tumor activity is also enhanced by depleting CCR4+ malignant cells) [external: Tanaka 2021, Nat Commun]. For the shieldbreak question (generalizable Treg-depletion mechanism), the class concern remains. Severity is moderate — in CTCL bounded by dual anti-tumor activity, but the mechanism-level effector collateral is real and intrinsic to the drug. COI & funding.**Kyowa Kirin likely context; French + European investigator-led.** Spin / abstract-to-text consistency.** Abstract framing is careful and balanced — no over-claiming.

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Prospective design with rigorous malignant/benign discrimination; rare-disease cohort size. **Per-trial confidence.****Moderate**— Class-consistent finding (Fujikawa 2023, Jinushi 2025); rigorous gating; paywall limits full verification.

Roussos Torres ET 2021 — Phase I study of entinostat + nivolumab ± ipilimumab in patients with metastatic HER2-negative breast cancer

Citation.[PMID 34135021](#) · [PMCID PMC8563383](#) · [DOI 10.1158/1078-0432.CCR-21-0561](#) · [NCT NCT02453620](#)

Source tier.[pmc_full_text](#)

Key critique.Entinostat-based combo phase I in HER2-negative breast cancer: CD8:FoxP3 tumor ratio 4.11→9.03 (2.2× shift, p=0.002, Wilcoxon signed-rank). The literature search's +119.7% pct_change is a RATIO pct_change, not a Treg-only reduction pct_change — this should be kept in mind when reading the row (it is not a Treg-absolute reduction).

Design.Phase I combination trial of entinostat + nivolumab ± ipilimumab in advanced HER2-negative breast cancer. Paired pre/post tumor biopsies. **Sample size & power.**Phase I standard cohorts. **Effect-size calibration.** Median CD8/FoxP3 ratio 4.11 → 9.03 (2.2× shift) T0 vs T2, p=0.002 (Wilcoxon signed-rank). The literature search's +119.7% pct_change correctly captures this as a ratio-shift readout (not a Treg reduction per se). Consistent with Terranova-Barberio 2020 (vorinostat in breast) and Pili 2017 (entinostat in RCC). **Missing data.** Phase I attrition standard. **Multiplicity.**Multiple biomarker endpoints; ratio-shift is one of many. **Generalizability.** HER2-negative breast cancer; checkpoint-combo setting. **Treg definition / gating.**FoxP3 IHC likely; not CD25-targeting. **Counter-productive mechanisms (Low (alt-checkpoint-upregulation)).**Entinostat + nivo ± ipi**

combination. Same class-I-HDACi class concerns as Pili 2017 (bounded, class-I-selective). The nivo/ipi combination adds the alt-checkpoint-upregulation concern for anti-CTLA-4 [external: Huang 2017, PNAS] and paradoxical autoimmunity risk. Severity is low-to-moderate. COI & funding. Syndax (entinostat) + BMS (nivolumab) industry context. Spin / abstract-to-text consistency.** Ratio-shift framing is appropriate given underlying Treg decrease is not separately quantified.

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Phase I combination; Wilcoxon-correct statistics; industry-sponsored.

Per-trial confidence.**Moderate**— Clean stats and ratio-shift framing; combination-specific interpretation.

Sharma A 2019 — Anti-CTLA-4 immunotherapy does not deplete FOXP3+ regulatory T cells (Tregs) in human cancers

Citation.[PMID 30054281](#) · [PMCID PMC6348141](#) · [DOI 10.1158/1078-0432.CCR-18-0762](#)**Source tier.**[pmc_full_text](#)

Key critique.FOUNDATIONAL PAPER for the shieldbreak's Fc-engineering reconciling-variable framework. Ipilimumab AND tremelimumab INCREASE intratumoral FOXP3+ density across melanoma (N=19), prostate (N=17), bladder (N=9), and paired melanoma tremelimumab (N=18), with CyTOF orthogonal validation (n=5). Directly contradicts preclinical ADCC-Treg-depletion model. Rigorous self-critique of single-marker FOXP3 IHC. Corroborates Huang 2011 and is the pivot paper that motivated development of Fc-enhanced anti-CTLA-4 (Ager 2026).

Design.Quantitative IHC analysis of FFPE tumor samples from: stage-matched UNTREATED vs ipilimumab-treated cohorts in melanoma (N=19), prostate (N=17), bladder (N=9); plus PAIRED pre/post tremelimumab-treated melanoma (N=18). Additional CyTOF analysis of n=5 paired fresh melanoma tumors. Median timing: 8 weeks (prostate/bladder), 18 weeks (melanoma after ipilimumab), 5 weeks (tremelimumab).

Sample size & power.Paired cohorts (tremelimumab n=18, CyTOF n=5) are small. Unpaired stage-matched cohorts for ipi have inherent baseline heterogeneity that no formal matching (just 'stage-matched') can fully resolve. No formal power calculation. **Effect-size calibration.**SIGNATURE FINDING: Both ipilimumab and tremelimumab INCREASE intratumoral FOXP3+ density across all three tumor types. Directly contradicts mouse-model ADCC-depletion hypothesis. Corroborates Huang 2011 (tremelimumab, same direction). The paper also provides a meta-methodological insight: prior 'Treg reduction' claims in bladder (Liakou 2008) were driven by FREQUENCY (% of total T cells) calculations — total T cells increased so frequency decreased, but ABSOLUTE FOXP3+ count did NOT decrease. This is an important lesson for the entire shieldbreak: frequency-based readouts can be misleading in the context of broad T-cell activation. **Missing data.**Fresh-tumor CyTOF available in only 5 paired melanoma samples — small subset for the orthogonal validation. **Multiplicity.**Multiple tumor types, multiple isotypes, multiple cell markers — extensive testing without explicit multiplicity correction. However, the consistent directional finding across all comparisons (all show increase or no change, none show decrease) is its own robustness check. **Generalizability.**Paired-melanoma tremelimumab (n=18) is the highest-quality

comparison. Unpaired stage-matched cohorts for prostate and bladder add breadth but with design caveats. **Treg definition / gating.** SINGLE-MARKER FOXP3 IHC for the main cohort. The paper itself notes this limitation: 'single staining for FOXP3 alone can represent Teff and Treg cells.' The CyTOF data (n=5) address this with CD3+CD4+FOXP3+CTLA-4+ — and they show INCREASED Treg frequency of the co-expression-defined population, confirming the IHC finding. This is methodologically rigorous self-critique. **Counter-productive mechanisms (Low (alt-checkpoint-upregulation, non-depleting-isotype)). **Well-designed Sharma 2019 paper is the cleanest negative intratumoral Treg data for standard anti-CTLA-4. Class counter-productive concerns: alt-checkpoint upregulation [external: Huang 2017, PNAS], failure to deplete means release of CTLA-4 inhibition on both Treg and Teff compartments equivalently. The paper's finding that prior 'Treg reduction' claims were frequency-denominator artifacts (Treg % dropped because total T cells rose) adds to the concern that framing anti-CTLA-4 as a Treg-depletor misallocates mechanistic attention. Severity is low for the shieldbreak question (the drug doesn't do the proximal thing, so collateral effects on Tregs-vs-effectors are moot at that mechanism). COI & funding. MD Anderson + UCLA academic. Likely Parker Institute / SPORE funding (to confirm). Ribas / Wargo / Allison / Sharma are senior anti-CTLA-4 investigators — with both credibility and implicit industry relationships. Spin / abstract-to-text consistency.** Title commits firmly to the negative: 'Anti-CTLA-4 immunotherapy does not deplete FOXP3+ Tregs.' Body supports it rigorously. No spin — this is a high-integrity paper that challenges a dominant preclinical paradigm.**

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Mix of paired (tremelimumab, CyTOF) and unpaired (stage-matched ipi) designs; single-marker IHC with CyTOF orthogonal validation; rigorous self-critique of methodology. **Per-trial confidence.High**— Cross-tumor-type consistency, orthogonal CyTOF validation, rigorous self-critique, and the Huang 2011 corroboration together establish this as a foundational paper for the 'standard anti-CTLA-4 does not deplete intratumoral Tregs' conclusion.

Terranova-Barberio M 2020 — Exhausted T cell signature predicts immunotherapy response in ER-positive breast cancer (vorinostat + tamoxifen + pembrolizumab)

Citation.[PMID 32681091](#) · [PMCID PMC7367885](#) · [DOI 10.1038/s41467-020-17414-y](#) · [NCT NCT02395627](#)**Source tier.pmc_full_text**

Key critique.n=34 randomized phase II vorinostat + tamoxifen + pembrolizumab in ER+ breast cancer. Tumor CD4+FoxP3+CTLA-4+ Tregs reduced 11.8%→2.9% overall (p=0.0067); responders 20.3%→4.2% (p=0.031), non-responders 10.4%→2.5% (p=0.034). Screener's -75.4% is correct. CRITICAL: the Treg reduction occurs in BOTH responders and non-responders — it is not a correlate of benefit. Trial stopped early for futility (4% ORR). PBMC Tregs unchanged (p=n.s.). Rigorous CD4+FoxP3+CTLA-4+ triple-marker gating.

Design. Randomized phase II trial (NCT02395627) of vorinostat + tamoxifen + pembrolizumab in ER+ metastatic breast cancer (n=34). Two randomization arms (vorinostat day 1 of cycle 1 vs day 1 of cycle 2). Paired tumor biopsies at baseline and cycle 3. Median 5 prior metastatic regimens — very pretreated. **Sample size & power.** n=34 enrolled with paired tumor at n=21-25 for subset analyses. Simon two-stage design (stopped early for futility per discussion). Treg-endpoint analyses: n=25 for overall time-trend, n=21 for responder-vs-non-responder at baseline. **Effect-size calibration.** CD4+FoxP3+CTLA-4+ activated Tregs reduced from 11.8% to 2.9% in TUMOR (p=0.0067; n=25 presumed). Responders: 20.3%→4.2% (p=0.031); non-responders: 10.4%→2.5% (p=0.034). The literature search's pct_change=-75.4% is computationally correct ($1 - 2.9/11.8 = 0.754$). **CRITICAL NOTE:** the tumor-Treg reduction occurred in BOTH responders and non-responders (p=0.031 and p=0.034 respectively) — so the Treg reduction is NOT a correlate of clinical benefit. In fact, only 4% ORR and 19% CBR overall; trial was stopped early. The PBMC Treg analysis showed NO change (p=n.s. at all timepoints). This is a compartment-specific effect — tumor-only. **Missing data.** Trial stopped early for futility. Biopsy subset is smaller than enrolled. n for specific subgroup comparisons (n=25 for main panel, n=21 for responder baseline, n=12 for HLA-DR/Ki67) is transparently reported. **Multiplicity.** Multiple cell-type comparisons (CD3, CD4, CD8, Tregs, etc.); no explicit Bonferroni but tests are reported with specific p-values. The two-tailed Wilcoxon/Mann-Whitney choice is appropriate. **Generalizability.** ER+ metastatic breast cancer on 5 prior regimens. Highly pretreated population. Low ORR (4%) despite significant Treg reduction — caps the clinical interpretation. **Treg definition / gating.** CD4+FoxP3+CTLA-4+ — a rigorous triple-marker gating for ACTIVATED Tregs specifically. Excellent gating quality. Authors also analyze HLA-DR and Ki67 on Tregs — found no change in activation markers, suggesting numerical decrease rather than phenotype shift. ****Counter-productive mechanisms (Moderate (effector-function-suppression)).** **Vorinostat (broad HDACi, not class-I-selective) in breast cancer.** **Counter-productive concerns: (a) vorinostat's broader HDAC target spectrum includes class-II HDACs linked to effector function [external: Kroesen et al. 2014, Oncotarget]; (b) tamoxifen-pembro adds modulators with their own off-target profiles. The paper reports tumor Treg decrease with PBMC Treg stability, a compartment dissociation the class-I-selective preclinical evidence does not cleanly predict for pan-HDAC vorinostat. Severity is moderate. COI & funding. Merck pembrolizumab + academic (UCSF Munster lab); Syndax was not involved (this is vorinostat, Merck's Zolanza, not entinostat). Munster has established HDACi-immunotherapy research program. Spin / abstract-to-text consistency.**** **Honest:** authors specifically note the 4% ORR and early trial termination in discussion. The Treg-reduction finding is framed as a potential mechanism-of-response correlate (responders had HIGHER baseline Tregs 15.8% vs 11.5% n.s.), but the compartment-specific nature (tumor yes, PBMC no) and the bidirectional reduction in responders AND non-responders is transparently reported.

Extraction fidelity. Matches source.

Retraction status. None found as of review date.

Risk of bias. RoB 2: **Some concerns**— Randomized phase II with biopsy substudy; early futility termination; rigorous gating; transparent reporting; no spin. Randomization arm was sequencing of vorinostat, not for the Treg-endpoint comparison. **Per-trial confidence. Moderate**— Rigorous gating and transparent reporting; the Treg reduction occurring in BOTH responders and non-responders argues that it is NOT the driver of the modest observed clinical benefit — which caps the practical importance of the 75% reduction finding.

Thibodeaux SR 2021 — IFN α Augments Clinical

Efficacy of Regulatory T-cell Depletion with Denileukin Diftitox in Ovarian Cancer

Citation.PMID 33771857 · DOI 10.1158/1078-0432.CCR-20-4594Source tier.pubmed_abstract

Key critique.Combined mouse + phase 0/I + phase II ovarian abstract. Phase II DD monotherapy FAILED; DD + IFN α 2a 'worked' in 2/2 before drug shortage halted enrollment. n=2 is statistical noise. Paywall limits verification.

Design.Combined preclinical (ID8 mouse model) + phase 0/I (DD escalation for functional Treg reduction) + phase II (DD monotherapy in ovarian, failed) + phase II (DD + IFN α 2a, 2/2 responders before DD shortage halted). Complex multi-study abstract. **Sample size & power.**Phase II salvage combination: n=2 completed before DD shortage. This is a SINGLE-DIGIT observation. Phase II monotherapy 'failed' — n not cleanly reported in abstract. **Effect-size calibration.**DD alone 'failed phase II' in ovarian cancer — a clean negative signal. Adding IFN α 2a to DD 'producing immunologic and clinical benefit in two of two patients' — n=2 is an uninterpretable signal. The abstract frames 2/2 as encouraging but statistical noise is indistinguishable from true effect. **Missing data.**DD shortage halted the trial — not a bias issue per se but limits data. **Multiplicity.**Multiple sub-studies conflated in abstract. **Generalizability.**Ovarian cancer; ID8-murine + human phase I-II. **Treg definition / gating.**Abstract refers to 'functional Treg reduction' (suggesting suppression assay used). This would side-step CD25-gating confound; full text required to verify. ****Counter-productive mechanisms (Moderate (cd25-effector-collateral, ifna-cd25-upregulation-interaction)).****Phase II failure of DD + IFN α in advanced ovarian cancer. The class-level CD25+ effector collateral concern applies, and is relevant here because IFN α upregulates CD25 on activated T cells [external: Petricoin et al. 1997, Nature]. DD + IFN α could therefore make activated effector T cells MORE susceptible to DD depletion — the combination may have a worse effector-collateral profile than DD alone. The paper reports clinical failure without mechanistic dissection of this specific interaction. Severity is moderate-to-high for the specific combination studied. COI & funding.****Not accessible from abstract.** Spin / abstract-to-text consistency.** 2/2 patients is presented as a positive signal in the abstract. This is mild spin — 2 patients cannot establish efficacy; the DD monotherapy failure is the more interpretable result.

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Serious**— Combined-study abstract with very small phase II salvage n (n=2); paywall limits full verification. **Per-trial confidence.****Low**— Abstract-only; n=2 salvage signal; phase II monotherapy failure is the more interpretable finding.

Vaishampayan UN 2024 — Nemvaleurin alfa (ALKS 4230): phase 1 trial in advanced solid tumors — monotherapy and combination with pembrolizumab

Citation.PMID 39567211 · PMCID PMC11580269 · DOI 10.1158/2643-3230.BCD-24-0051 · NCT NCT02799095
Source tier.pmc_full_text

Key critique.Nemvaleukin monotherapy + pembro phase 1. Treg ~2× expansion (transient, C1D5 only), CD8+ 2.53× and NK 6.52× — ratio-shift mechanism. Class-consistent with Calvo 2025.

Design.Phase 1 dose-escalation of ALKS 4230 (nemvaleukin alfa, engineered non-α IL-2) monotherapy and combination with pembrolizumab in advanced solid tumors. Longitudinal PBMC analyses. **Sample size & power.** Phase 1 cohorts; multiple dose levels. **Effect-size calibration.**Tregs expanded ~2× (fold-change from baseline); CD8+ 2.53× and NK 6.52× — clear effector-preferring expansion (the 'ratio-shift' mechanism). Treg expansion was LIMITED TO C1D5 only — transient. Consistent with Calvo 2025 and with the non-α IL-2 class mechanism.

Missing data.Phase 1 standard. **Multiplicity.**Many timepoints and subsets; qualitative framing without per-comparison p-values (as the literature search correctly captures). **Generalizability.**Advanced solid tumors. **Treg definition / gating.**Standard. **Counter-productive mechanisms (High (treg-expansion-by-design, denominator-artifact-framing)). **Same class concerns as Calvo 2025. ARTISTRY-1 nemvaleukin mono shows absolute Treg expansion with claimed effector-biased ratio. The subsequent phase 3 ARTISTRY-7 failure in ovarian cancer [external: Mylavaram et al. 2023, ASCO] is the clinical evidence the ratio-shift mechanism does not translate. Severity is high. COI & funding.Mural Oncology sponsorship. Spin / abstract-to-text consistency.** Ratio-shift framing is the class standard for non-α IL-2 — appropriate.**

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Phase 1 dose-escalation; industry-sponsored; qualitative p-value reporting. **Per-trial confidence.****Moderate**— Class-consistent mechanism signal; phase 1 caps precision.

Yi JS 2017 — Tumor-associated macrophages and Treg responses in adoptive transfer of tumor-infiltrating lymphocytes with or without cyclophosphamide pretreatment (NSCLC translational)

Citation.PMID 28951518 · PMCID PMC5732888 · DOI 10.1016/j.jtho.2017.09.1954 **Source tier.**pmc_full_text

Key critique.ACT + cyclophosphamide-pretreatment translational substudy. Small n; causal attribution complicated by combination therapy.

Design.Translational immune-profiling substudy of an adoptive cell transfer protocol. Two arms with/without cyclophosphamide pretreatment. Paired PBMC analyses. **Sample size & power.**Translational subset; small n per arm. **Effect-size calibration.**Cyclophosphamide-pretreatment context; mechanism overlaps with Ghiringhelli 2007. **Missing data.**Standard. **Multiplicity.**Multiple T-cell subsets. **Generalizability.**NSCLC ACT setting. **Treg**

definition / gating.Standard. ****Counter-productive mechanisms (Low (alt-checkpoint-upregulation, lymphopenia-collateral)).****Neoadjuvant chemotherapy + ipilimumab in early-stage NSCLC. Ipilimumab counter-productive concerns apply (alt-checkpoint upregulation, paradoxical autoimmunity) [external: Huang 2017, PNAS].** The chemotherapy component adds a lymphopenia-collateral dimension: the chemo regimen itself depletes lymphoid compartment non-selectively, potentially reducing the effector population that ipilimumab activation would otherwise recruit. **Severity is low-to-moderate: the combination's counter-productive exposure comes more from chemo than from ipilimumab at the mechanism level.** COI & funding.**Academic.** Spin / abstract-to-text consistency.**** Descriptive.**

Extraction fidelity.Matches source.

Retraction status.None found as of review date.

Risk of bias.ROBINS-I: **Moderate**— Small translational substudy. **Per-trial confidence.****Low**— Small n; ACT context complicates attribution.

</details>