Research Square

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

How many is too many? A clinical ceiling on multi-specific antibodies, and why smooth cost compounding does not explain it

William Chi

william.y.chi@warburgtx.com

Warburg Therapeutics Limited https://orcid.org/0009-0006-7047-1037

Research Article

Keywords: multispecific antibodies; bispecific antibodies; trispecific antibodies; valency; clinical development attrition; drug development costs

Posted Date: November 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11836252/v1

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

Additional Declarations: The authors declare potential competing interests as follows: William Y. Chi is the founder and Chief Executive Officer of Warburg Therapeutics Limited, which funded this work and is developing antibody technologies, including multispecific formats. The author declares no other competing interests.

How many is too many? A clinical ceiling on multi-specific antibodies, and why smooth cost compounding does not explain it

William Y. Chi

Warburg Therapeutics Limited, Oxford Science Park, Robert Robinson Avenue, Oxford OX4 4GP, UK; Hong Kong Science and Technology Park, Hong Kong

Correspondence: william.y.chi@warburgtx.com

Abstract

The antibody field has moved from one arm to two, and now to three. The logic behind that move is easy to state and hard to resist: if two arms are good, four must be better. Preclinical pipelines are full of it. We test that logic twice, against the clinical record and against a simple model. A systematic ClinicalTrials.gov census (pull date 17 November 2026) retained 166 multi-specific programmes: 156 bispecific, 8 trispecific, 2 tetraspecific. All 15 FDA-approved multi-specific antibodies are bispecific. Trispecific programmes have entered trials at 0 to 2 per year since 2019 without accumulating, and only 1 of 8 has reached Phase 3; both tetraspecifics have stopped. Scarcity above two specificities is therefore not an artefact of timing. We then build a three-force value model in which benefit saturates while manufacturability and safety compound downward. It reduces to a single step rule: an added arm pays for itself only if its relative benefit gain clears $1/(f \cdot s)$, where f and s are per-arm feasibility and safety retentions. Across 200,000 sampled parameter draws that threshold has a median of 1.69 and a 90% interval of 1.25 to 2.35, the optimum falls at two specificities in 48% of draws and three in 37%, and four or more is favoured in 2%. The model's prediction for four or more matches the observed 1.2% clinical share. Its prediction for three does not: trispecifics are 4.8% of programmes, roughly eightfold below prediction, and an independent valuation of 40 candidate combinations points the same way. Smooth per-arm compounding of manufacturability and safety costs is therefore insufficient to explain the ceiling at two. We report this failure rather than fit around it, and argue that a discrete barrier operates at the third arm. The answer to "how many is too many" is a threshold rule, not a number, and the threshold is set by the platform rather than by the biology.

The maximalist temptation

Monoclonal antibodies target one antigen. Bispecifics target two, and in doing so unlock mechanisms a single arm cannot reach: redirecting a T cell to a tumour cell, mimicking a missing clotting factor, blocking two nodes of one pathway, or requiring two antigens to be present before acting. That step change has been clinically decisive. Fifteen bispecifics are now approved by the FDA, among them blinatumomab, emicizumab, amivantamab, tebentafusp and faricimab, plus a wave of CD3-engaging molecules in lymphoma and myeloma [1 to 3]. Trispecific molecules are the visible frontier [4 to 6].

From here the pipeline logic tends to extrapolate: if adding a specificity helped once, add another. Preclinical disclosures now include tetra- and higher-valent constructs stacking a tumour antigen, a second tumour

antigen for selectivity, a T-cell trigger, a costimulatory arm and a half-life module into one chain. Underneath that is an assumption nobody states, that benefit accumulates faster than cost. We test it two ways: against what has actually reached patients, and against a model of the costs.

The clinical record concentrates at two, and not because the field is young

A systematic census of ClinicalTrials.gov (Methods) retrieved 1,392 unique trials and retained 166 multi-specific programmes. The distribution is extreme (**Figure 1a**, **Table 1**). Bispecifics account for 156 programmes (94.0%), trispecifics for 8 (4.8%), tetraspecifics for 2 (1.2%). Nothing above four specificities has reached a registered trial. All approved agents are bispecific: 15 by the FDA, 16 counting EMA-only approvals, 19 counting China, and none at three or more.

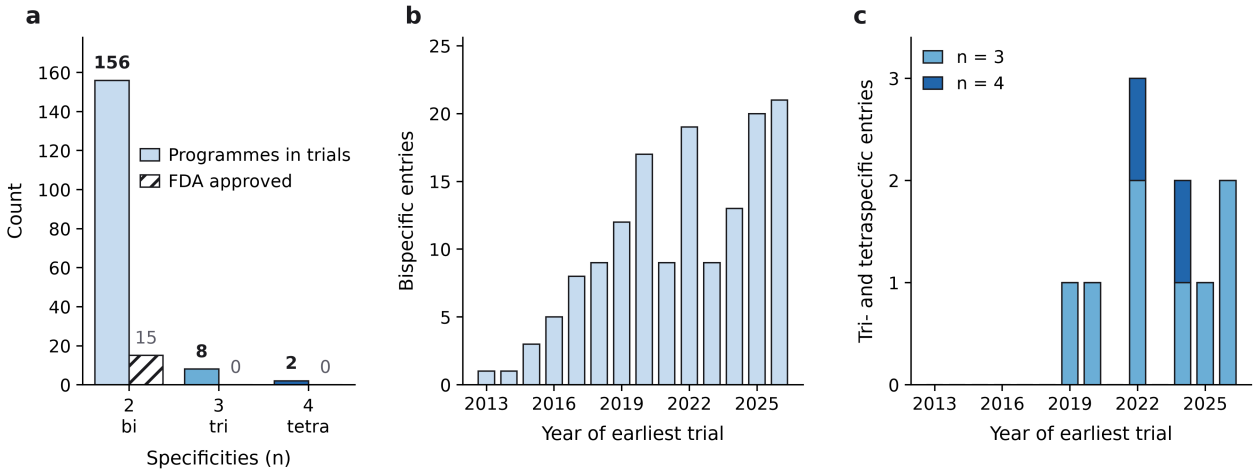


Figure 1 | The clinical record concentrates at two specificities. (**a**) Retained programmes and FDA approvals at each specificity count, census pull 17 November 2026. (**b**) Bispecific programmes by year of earliest retained trial. (**c**) Trispecific and tetraspecific programmes on the same dating and axis, plotted on their own scale as a small multiple rather than a second y-axis. Trispecific entries have run flat at 0 to 2 per year since 2019 while bispecific entries rose.

The obvious objection is timing. Trispecifics might be scarce simply because they are new, in which case the ceiling is a lag and not a limit. The census answers this. Dating each programme by its earliest retained trial, trispecifics first appear in 2019 and have entered at 0 to 2 per year every year since, without accumulating, while bispecific entries rose from single digits to 21 in 2026 (**Figure 1b,c**). Scarcity that persists at a flat rate across eight years, while bispecific entries rose over the same period, is not a cohort effect. This dating is later than the true first-in-human date for programmes whose Phase 1 record was not retrieved, so the flat trispecific rate has run at least as long as shown, not shorter. That is evidence against the timing objection, not for it. **Table 2** dates the same programmes by first-in-human year, the earliest Phase 1 or Phase 1/2 trial, which gives 22 bispecific entries in 2026; the two datings differ for 4 programmes and are stated where each is used.

Attrition is more striking than entry. Of the 8 trispecific programmes, 3 are listed as not yet recruiting past their own planned start dates, 2 are terminated, 2 are recruiting, and 1 is active but no longer recruiting. One of the two recruiting programmes, the GPRC5D by BCMA by CD3 engager JNJ-5322, is also the only trispecific to reach Phase 3. Both tetraspecifics have stopped: MP0533, a DARPin scaffold without an Fc,

completed Phase 1, and IPH6501 was terminated in Phase 1/2. Higher-order molecules are not failing to start. They are starting and not continuing.

The third arm does not rescue the picture. On the 8 trispecific programmes the third arm is a selectivity arm, a second tumour target, in 4, a costimulatory or additional activating arm in 2, and neither in 2. The only Phase 3 trispecific, JNJ-5322, is a selectivity design, but on 8 programmes that cannot test the prediction that selectivity designs should do better. One reclassification moves any share by 12.5 percentage points. We report it as a null result, not as support.

Two details sharpen the picture, and both cut against a maximalist reading. Six of 8 trispecifics are haematological, and 3 of those are the same BCMA, GPRC5D and CD3 concept pursued by three sponsors, so the trispecific frontier is narrower than a count of 8 suggests. The single tetraspecific that completed its trial is the one built on a non-antibody scaffold with no light chains to mispair, which is the format the cost argument below predicts should be the first to tolerate four arms.

Table 1 | Multi-specific programmes by specificity count and furthest phase reached. Census pull 17 November 2026; all values derived quantities except approvals, which are published figures cross-checked against FDA and EMA listings.

n	Phase 1	Phase 1/2	Phase 2	Phase 3	Phase 4	Total	Approved (FDA)
2	58	46	27	17	8	156	15
3	3	2	2	1	0	8	0
4	1	1	0	0	0	2	0
≥5	0	0	0	0	0	0	0

Table 2 | First-in-human starts by specificity count and calendar year. Year is the earliest Phase 1 or Phase 1/2 trial start for the programme, with the earliest retrieved trial used where no Phase 1 or Phase 1/2 trial was retrieved. This differs from Figure 1, which dates programmes by their earliest retained trial; the two differ for 4 programmes. 2026 is a partial year, pull 17 November 2026. The 2027 row is a single programme whose first trial is scheduled to start after the pull date.

Year	n = 2	n = 3	n = 4	Total
2018 and earlier	34	0	0	34
2019	12	1	0	13
2020	17	1	0	18
2021	10	0	0	10
2022	18	2	1	21
2023	10	0	0	10
2024	11	1	1	13
2025	21	1	0	22
2026	22	2	0	24
2027	1	0	0	1
Total, all years	156	8	2	166

The census is a lower bound. Four named multispecifics in development, ISB 2001, SAR442257, EXP011 and AZD9793, never use any search term in their registry records and are absent from the pull. Three of the four are trispecific, so the true clinical trispecific count is at least 11; AZD9793 is bispecific. This matters for the level but not for the shape: adding three programmes to 8 leaves trispecifics under 7% of the retained set.

Three costs that compound

Assembly fidelity. A conventional IgG is built from two identical heavy chains and two identical light chains that assemble with near-perfect fidelity. A bispecific IgG needs two different heavy chains and two different light chains to co-express in one cell and find their one correct arrangement. Under random chain association those four distinct chains generate sixteen combinations, nine mispaired species alongside the one that is wanted, a statistical yield of 12.5% [7,8]. The first approved bispecific, the rat-mouse quadroma catumaxomab, carried exactly this burden, and its hybrid-hybridoma route made the mispaired fraction a defining manufacturing constraint [9]. What followed is often told wrongly, so it is worth getting straight: catumaxomab was withdrawn from the United States in 2013 and from the European Union in 2017 for commercial reasons connected to its manufacturer, not for failure of purification or safety, and it was re-approved by the European Commission in February 2025. The chain-pairing tax is real; it did not kill this molecule. Engineering does not remove the tax either. It converts it into a per-interface fidelity that compounds (**Figure 2**), with knobs-into-holes heterodimerisation the canonical example [10]. If each controlled interface pairs correctly with probability *p*, a molecule needing *m* interfaces yields correct product at p^m. A bispecific IgG controls roughly three interfaces, a trispecific four or five, a tetraspecific about seven, so 95% per-interface fidelity gives about 86% correct product for a bispecific, 77% to 81% for a trispecific, and pushes toward 70% for a tetraspecific. Each lost percentage point is a distinct, structurally similar species that regulators expect quantified and controlled, so the analytical burden grows faster than the molecule.

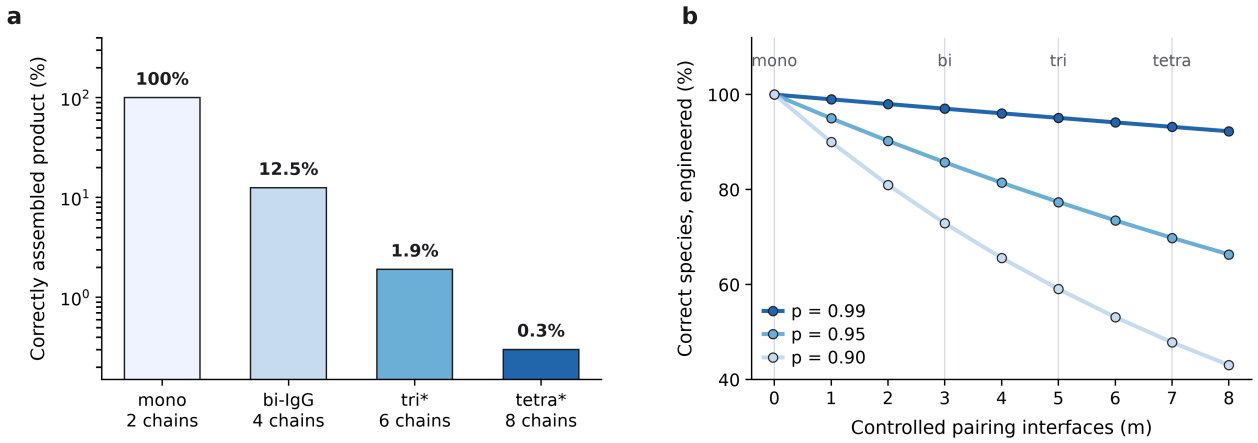


Figure 2 | The assembly-fidelity tax. (**a**) Correctly assembled fraction under random chain association, log scale. The monospecific and bispecific values are arithmetic: one correct arrangement of one, and one of the sixteen combinations available to two distinct heavy and two distinct light chains. The tri and tetra values, marked with an asterisk, are illustrative, not computed: they show the direction of the fall, not a measured yield. (**b**) Compounding of engineered per-interface fidelity, p^m, against the number of controlled interfaces, with the approximate interface count of each format marked. Engineering converts the tax of panel a into the slower decay of panel b; it does not remove it.

Developability and pharmacokinetics. Every appended domain raises aggregation propensity, introduces hydrophobic and charge patches, lowers thermostability and complicates glycosylation. The charge character of what is appended is not a detail. On trastuzumab, metal-chelating polymers carrying zwitterionic pendant groups accumulate markedly less in liver than otherwise comparable polyanionic conjugates [11], so two constructs of similar size and valency can differ in biodistribution for reasons that sit entirely in the appended chemistry rather than in the binding arms. Fragment formats that shed the Fc gain manufacturability and lose neonatal-Fc-receptor recycling, trading a CMC problem for a pharmacokinetic one that must be engineered back.

Safety. Approved CD3 engagers already carry class-defining toxicity. Cytokine release syndrome affects about 72% of patients on teclistamab, ICANS about 3%, both under boxed warnings requiring step-up dosing, and sustained immune activation drives infection in about 76%, with 44.8% at grade 3 or 4, a leading cause of treatment-related death [12]. A dedicated infection analysis of the same trial at longer follow-up puts grade 3 to 4 infections higher still, at 55.2%, with 21 infection deaths [13]. Pharmacovigilance across teclistamab, talquetamab and elranatamab flags the same cluster [14]. Each added arm widens this surface along its own axis: a new tumour antigen brings its own normal-tissue expression, a new activating arm stacked on CD3 risks more frequent and higher-grade CRS, which is why CD28 costimulatory designs are engineered with deliberately attenuated affinity [4], and each non-native junction raises immunogenicity risk. One exception runs the other way. An arm added for selectivity can widen the window rather than narrow it. Trispecific engagers that require two tumour antigens before activating have been reported to kill double-positive cells 70-fold to 750-fold more potently than single-antigen controls, depending on which antigen is compared [15], and placement of the higher-affinity arm, rather than the mere presence of a second tumour binder, determines how much selectivity is gained [16]. That exception is conditional. Dual targeting alone does not confer selectivity: affinity of the individual arms, overall avidity and format valence jointly decide it, and affinity-reduced monovalent formats outperform their bivalent counterparts on selectivity [17]. The window therefore depends on a combination that becomes harder to satisfy as arms multiply.

The model reduces to a threshold, not a number

Let value as a function of specificity count *n* be the product of a benefit term, a make-and-deliver term and a safety term:

$$\mathbf{V(n) = B(n) \cdot F(n) \cdot S(n)}$$

B(n) saturates, because the marginal mechanistic gain shrinks: one to two is categorical, two to three is real but smaller, beyond that marginal. F(n) compounds downward with assembly, developability and PK costs. S(n) compounds downward as each arm widens the liability surface, except where an arm's job is selectivity, in which case its retention exceeds one.

Written per step rather than per molecule, the model collapses to something more useful than a curve. Arm *k* raises value if and only if

$$\mathbf{g_k \cdot f \cdot s_k > 1}$$

where g_k is the relative benefit gain at that step. So the question "how many is too many" is really "does the next arm clear 1/(f·s)". That threshold is an empirical quantity, and it is the paper's answer. Across sampled

parameter ranges (Methods) it has a **median of 1.69 and a 90% interval of 1.25 to 2.35** (**Figure 3b**). Published bi-to-tri benefit gains sit mostly below 1.7, which is why the clinic stops where it does.

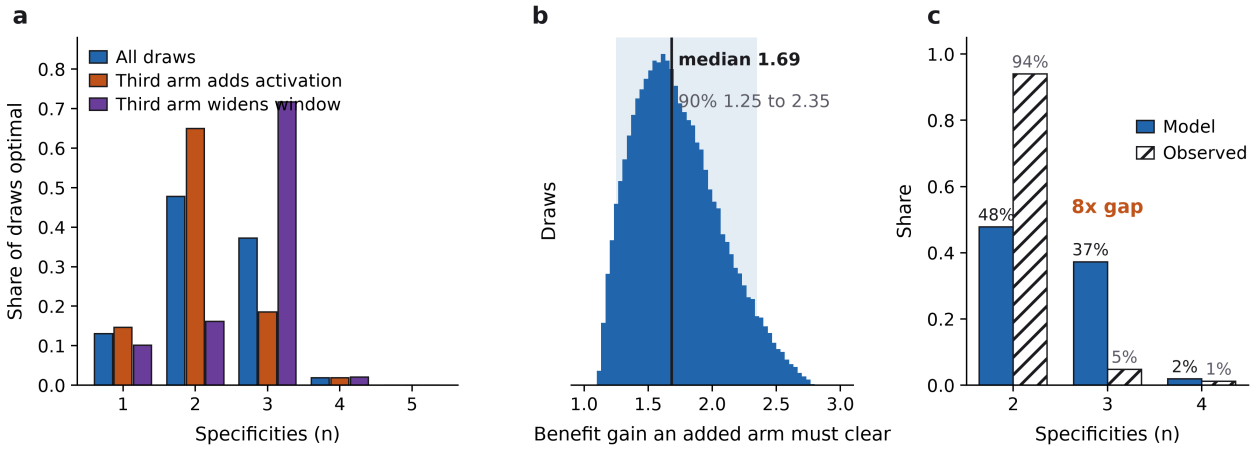


Figure 3 | The value model under sampled uncertainty. (**a**) Share of 200,000 draws in which each specificity count is optimal, split by whether the third arm adds activation or widens the therapeutic window. (**b**) Distribution of the step threshold 1/(f·s) that an added arm's benefit gain must clear, median 1.69, 90% interval 1.25 to 2.35. (**c**) Model against observed shares, showing agreement at four specificities and a 7.7-fold discrepancy at three.

Monte Carlo over 200,000 draws puts the optimum at two specificities in **47.8%** of draws, three in **37.3%**, four in **1.9%** and five in under 0.1%, with one in 13.0% (**Figure 3a**). Conditioning on the third arm's job splits this cleanly: where the third arm merely adds activation, two wins 65.0% of draws; where it widens the window, three wins 71.7%. The optimum is interior, early and conditional, as the three-force argument predicts.

Two further results qualify the framing. First, ablation shows the costs are not co-equal. Removing the safety term moves the optimum little, with two still winning 41.7% of draws. Removing the feasibility term moves the mode to three and gives four or more a 29% share. The assembly and CMC tax carries the result; safety matters mainly through the selectivity exception. Second, sensitivity shows the optimum is set by the platform, not the biology: correlation with an optimum of three or more is 0.58 for the third arm's safety retention and 0.44 for feasibility, against 0.32 for the bi-to-tri benefit gain itself.

The model under-predicts the penalty on the third arm, and we cannot fit it away

Two checks were available that the model was not calibrated against.

The first is the clinical distribution above. The model's prediction for four or more specificities is close: 1.9% of draws against an observed 1.2% of programmes. Its prediction for three is not: 37.3% against an observed 4.8%, a discrepancy of 7.7-fold (**Figure 3c**). The comparison is imperfect and needs care, because a count of programmes reflects cumulative history, sponsor risk appetite and patent timing rather than a per-programme optimality calculation. It is nonetheless the wrong way round by a wide margin, and no reasonable reading of the caveat reverses it.

The second is independent valuation. A companion analysis of 40 candidate combinations built from best-selling antibody franchises, constructed on a different derivation route and for a different purpose, puts the

trispecific median expected net present value at 0.57 of the bispecific median, and the mean ratio at 0.43. The model's median V(3)/V(2) for activating third arms is 0.82, and only 6.4% of draws fall as low as 0.57. The same direction again, with the model too generous to the third arm. That anchor is weak, resting on 8 trispecific entries valued by the same author, and we do not treat it as a calibration target.

Attempts to recover the anchors fail informatively, and we report the failure rather than suppress it. Scaling the per-arm retentions down until median V(3)/V(2) matches 0.57 breaks the other anchor: V(2)/V(1) falls to 0.97 and monospecifics then win 51% of draws, which fifteen bispecific approvals refute. Deriving F(n) instead from counted interfaces, $F(n)=p^{m(n)-1}$, fixes the bispecific result (V(2)/V(1) = 1.61) but erases the third-arm penalty entirely (V(3)/V(2) = 1.01). No specification of a smooth, per-arm multiplicative cost clears both anchors.

The conclusion follows directly. **A smooth compounding of manufacturability and safety costs is not sufficient to explain the clinical ceiling at two specificities.** Something discrete happens at the third arm that per-arm decay does not capture. Candidates are testable and we do not adjudicate between them here: a step change in analytical and regulatory expectation once the impurity matrix passes a threshold of resolvable species; a comparability burden that makes process change prohibitive rather than merely costly; or a clinical-development penalty, since a trispecific must justify its third arm against a bispecific comparator that already works, which is a harder trial to design than a first-in-class bispecific ever faced. Distinguishing these requires programme-level attrition and CMC data that registry records do not carry.

A decision rule that survives the failure

The threshold rule is unaffected by the misspecification above, because it concerns one step rather than the shape of the whole curve. For a team deciding how far to go, three questions in order:

1. Does the next specificity unlock a categorically new mechanism, or a marginal one? Quantify the gain. If it does not plausibly clear 1.7-fold, stop.
2. Can the format pair, express, purify and circulate the added arm without a disproportionate CMC and PK penalty? If not, change format or stop.
3. Does the added arm widen the therapeutic window more than it widens the liability surface? Only a selectivity arm reliably does.

By this rule most oncology programmes should be bispecific, the strongest selectivity-limited programmes may be trispecific on a mature platform, and four or more belongs on scaffolds that abolish light-chain mispairing. The census is consistent with the last point: the one tetraspecific to complete its trial is a DARPin without an Fc. That is a single observation and we draw no inference from it beyond noting the agreement.

The deeper point is that specificity count is the wrong variable to maximise. The right variables are the platform's capacity to absorb complexity and each arm's net effect on the therapeutic window. The clinic reached that conclusion empirically. The model explains part of why, and its failure marks where the remaining explanation must be found. The discipline generalises past antibodies. For conjugates the analogous variable is the payload or isotope, which can likewise be chosen from the tissue architecture rather than maximised, as spatial microdosimetry does in selecting between alpha and beta emitters for a TROP2-targeted conjugate [18].

Limitations

The census is a lower bound. 164 of 1,392 retrieved trials (11.8%) could not be identified well enough to count specificities and no targets were guessed; four named multispecifics are absent because their registry records use none of the search terms, three of them trispecific, putting the true trispecific count at 11 or more; format could not be verified for 20 of 166 programmes and Fc status for 17; and NMPA approvals rest on secondary listings that need primary verification. Registry coverage is weakest for Chinese and preclinical programmes, so the preclinical tail is not characterised here at all.

The model's parameter ranges are assumptions, not fitted values, and the Monte Carlo quantifies their consequences rather than validating them. The benefit gains in particular are the weakest input: if bi-to-tri gains are routinely above 1.7-fold, the threshold conclusion weakens and the discrepancy against the clinical distribution grows rather than shrinks. The clinical distribution is not a direct measurement of per-programme optimality, as noted. The valuation anchor is weak and shares an author with this analysis. The discrete-barrier hypothesis is a conclusion from the failure of alternatives, not a tested mechanism.

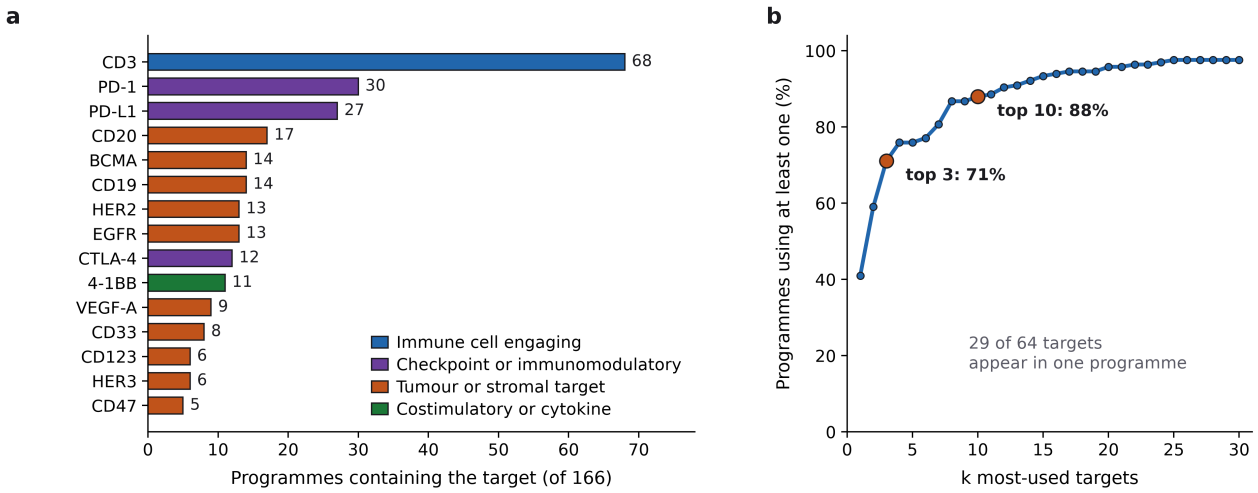
Methods

Census. ClinicalTrials.gov API v2, pull date 17 November 2026. Ten terms were run separately and the matching term retained: bispecific, trispecific, tetraspecific, multispecific, bi-specific, tri-specific, "T-cell engager", "T cell engager", bispecific antibody, "dual targeting antibody". Intervention and title fields were searched with no date filter. 1,392 unique trials were retrieved. Exclusions, each logged with a count: non-interventional (59), cell therapy (98), monospecific antibody-drug conjugate or radioligand (19), comparator only (2), co-administration of separate molecules (233), not multi-specific (194), unidentifiable (164). 623 trials passed screening and collapsed to 166 programmes, keeping the earliest start date per programme. First-in-human year in Table 2 is the earliest Phase 1 or Phase 1/2 trial start per programme, with the earliest retrieved trial used where no Phase 1 or Phase 1/2 trial was retrieved; it differs from the programme start_date in data/census_classified.csv, which is the earliest retrieved trial of any phase, for 4 programmes. Specificities were counted as distinct binding specificities rather than valencies; a biparatopic molecule binding two epitopes on one antigen counts as two; half-life modules and inert Fc domains count as none. Screening flow, classifications and counts, search URLs and ambiguous-record decisions are in the supplementary material.

Model. B(n) is a cumulative product of relative gains sampled per step. Ranges, all assumptions: 1.6 to 3.0 for one to two, 1.10 to 1.70 for two to three, 1.02 to 1.30 for three to four, 1.00 to 1.15 for four to five. Per-arm feasibility retention f is sampled uniformly on 0.55 to 0.95 and per-arm safety retention s on 0.65 to 0.95. In 35% of draws the third arm is treated as window-widening with retention sampled on 1.00 to 1.30, reflecting AND-gated selectivity designs. 200,000 draws, seed 20261117. Specification tests substitute $F(n)=p^{m(n)-1}$ with per-interface fidelity p on 0.90 to 0.98 and interface counts m = 1, 3, 4.5, 6, 7.5 for one to five specificities. Ablations set the relevant retention to unity. Code and outputs are in the supplementary material.

References

1. Klein C, Brinkmann U, Reichert JM, Kontermann RE. The present and future of bispecific antibodies for cancer therapy. *Nat Rev Drug Discov.* 2024;23(4):301-319. doi:10.1038/s41573-024-00896-6
2. Brinkmann U, Kontermann RE. The making of bispecific antibodies. *mAbs.* 2017;9(2):182-212. doi: 10.1080/19420862.2016.1268307
3. Labrijn AF, Janmaat ML, Reichert JM, Parren PWHI. Bispecific antibodies: a mechanistic review of the pipeline. *Nat Rev Drug Discov.* 2019;18(8):585-608.
4. Michaelson JS, Henry CC, Sauer K, Baeuerle PA. Next-generation T cell engagers for cancer and autoimmune diseases. *Front Immunol.* 2026;17:1885909. doi:10.3389/fimmu.2026.1885909
5. Popat R, van de Donk NWCJ, et al. First-in-human study of JNJ-79635322 (JNJ-5322), a novel next-generation trispecific antibody, in patients with relapsed/refractory multiple myeloma: initial phase 1 results. Presented at the European Hematology Association Congress, Milan, 12-15 June 2025, and the ASCO Annual Meeting, Chicago, 30 May to 3 June 2025. NCT05652335.
6. Quach H, Augustson B, Sia H, et al. First results of a phase 1, first-in-human, dose escalation study of ISB 2001, a BCMAxCD38xCD3 targeting trispecific antibody, in patients with relapsed/refractory multiple myeloma. Presented at the American Society of Hematology Annual Meeting, 2024, abstract 1026. NCT05862012.
7. Wang C, Vemulapalli B, Cao M, Gadre D, Wang J, Hunter A, Wang X, Liu D. A systematic approach for analysis and characterization of mispairing in bispecific antibodies with asymmetric architecture. *mAbs.* 2018;10(8):1226-1235. doi:10.1080/19420862.2018.1511198
8. Joshi KK, Phung W, Han G, Yin Y, Kim I, Sandoval W, Carter PJ. Elucidating heavy/light chain pairing preferences to facilitate the assembly of bispecific IgG in single cells. *mAbs.* 2019;11(7):1254-1265. doi: 10.1080/19420862.2019.1640549
9. Linke R, Klein A, Seimetz D. Catumaxomab: clinical development and future directions. *mAbs.* 2010;2(2):129-136. doi:10.4161/mabs.2.2.11221
10. Ridgway JBB, Presta LG, Carter P. "Knobs-into-holes" engineering of antibody CH3 domains for heavy chain heterodimerization. *Protein Eng.* 1996;9(7):617-621.
11. Liu P, Boyle AJ, Lu Y, Adams J, Chi Y, Reilly RM, Winnik MA. Metal-chelating polymers (MCPs) with zwitterionic pendant groups complexed to trastuzumab exhibit decreased liver accumulation compared to polyanionic MCP immunoconjugates. *Biomacromolecules.* 2015;16(11):3613-3623.
12. Moreau P, et al. Teclistamab in relapsed or refractory multiple myeloma. *N Engl J Med.* 2022;387(6):495-505. doi: 10.1056/NEJMoa2203478
13. Nooka AK, Rodriguez C, Mateos MV, Manier S, et al. Incidence, timing, and management of infections in patients receiving teclistamab for the treatment of relapsed/refractory multiple myeloma in the MajesTEC-1 study. 2024. PMID 37960969.
14. Zhou X, Luo E, Chen W, Deng Y, Wu B, Jiang X, Zhang K, Lai F. The silent signals: emerging safety concerns in bispecific antibody therapy for multiple myeloma. *Front Med.* 2025;12:1593405. doi:10.3389/fmed.2025.1593405
15. Shen Y, Jin SJ, Chen YC, Liu WH, Li YM, Zhao WY, Xu YC, Chen SQ, Zhao WB. Improving the tumor selectivity of T cell engagers by logic-gated dual tumor-targeting. *Pharmacol Res.* 2023;192:106781.
16. Dicara DM, Bhakta S, Go MA, Ziai J, Firestein R, Forrest B, Gu C, Leong SR, Yu SF, Polson AG, Agard NJ. Development of T-cell engagers selective for cells co-expressing two antigens. *mAbs.* 2022;14(1):2115213. doi: 10.1080/19420862.2022.2115213
17. Mazor Y, Sachsenmeier KF, Yang C, Hansen A, Filderman J, Mulgrew K, Wu H, Dall'Acqua WF. Enhanced tumor-targeting selectivity by modulating bispecific antibody binding affinity and format valence. *Sci Rep.* 2017;7:40098. doi:10.1038/srep40098
18. Chi WY. Computational pathology and spatial microdosimetry guide radiopharmaceutical selection for TROP2-targeted alpha versus beta radionuclide drug conjugates (RDCs). *bioRxiv* 2026.08.19.745876.



Supplementary Figure 1 | Target concentration. (**a**) The 15 most-used targets among the 166 retained programmes, coloured by the job the arm does: immune cell engaging (blue), costimulatory or cytokine (green), checkpoint or immunomodulatory (purple), tumour or stromal target (orange). The classes are assigned in the figure code, not taken from the census, which does not classify targets. CD47 is classed as a tumour target, by where it sits (on the tumour cell), rather than as a checkpoint, by what blocking it does. (**b**) Cumulative share of programmes touching at least one of the *k* most-used targets: the top 3 targets cover 71% of programmes, the top 10 cover 88%. 29 of the 64 distinct targets appear in a single programme each.