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.

# Bayesian and frequentist meta-analysis of GCB-classification heterogeneity in TP53-reporting DLBCL cohorts at small k: a methodological exemplar

**Aisha Al-Khinji**

aalkhinji@qu.edu.qa

Qatar University

**Hadeel Alashwal**

Qatar University

**Abdullatif Al-Hor**

Qatar University

**Hissa Al-Kuwari**

Qatar University

**Farwa Bismi**

Qatar University

**Nada Maki**

Qatar University

**Taghreed Abunada**

Qatar University

**Dhafer Malouche**

Qatar University

**Research Article**

**Keywords:** small-k meta-analysis, Bayesian hierarchical models, prior dominance, biomarker prevalence, TP53, diffuse large B-cell lymphoma, cell-of-origin, methodological reporting

**Posted Date:** June 19th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-9853529/v1

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

**Additional Declarations:** No competing interests reported.

**Version of Record:** A version of this preprint was published at BMC Medical Research Methodology on September 21st, 2026. See the published version at https://doi.org/10.1186/s12874-026-03015-8.

# Bayesian and frequentist meta-analysis of GCB-classification heterogeneity in TP53-reporting DLBCL cohorts at small k: a methodological exemplar

**Running head: Small-k Bayesian meta-analysis of GCB prevalence in TP53-reporting DLBCL**


**Authors**

Aisha Al-Khinji[1,*], Hadeel Alashwal[2], Abdullatif Al-Hor[1], Hissa Al-Kuwari[1], Farwa Bismi[3], Nada A Maki[3], Taghreed Abunada[3], Dhafer Malouche[2]

**Affiliations**

1. Department of Basic Science, College of Medicine, Qatar University, Doha, Qatar
2. Department of Mathematics and Statistics, College of Arts and Sciences, Qatar University, Doha, Qatar
3. Department of Biomedical Sciences, College of Health Science, Qatar University, Doha, Qatar

* Corresponding author: Aisha Al-Khinji, aalkhinji@qu.edu.qa


---


Abstract

**Background.** Meta-analyses of biomarker prevalence in diffuse large B-cell lymphoma (DLBCL) frequently rest on a small number of cohorts ($k \leq 5$) and are correspondingly fragile: between-study heterogeneity statistics destabilize, publication-bias diagnostics become inestimable, and conventional priors begin to dominate posteriors. We use the specific case of DLBCL cohorts that simultaneously report TP53 status and cell-of-origin (COO) classification as a methodological exemplar, taking the probability of germinal-center B-cell-like (GCB) classification as a pre-specified marginal estimand and developing complementary frequentist and Bayesian machinery aligned on that single target.

**Methods**. We systematically searched PubMed, Scopus, and Web of Science for retrospective DLBCL cohorts published between 2018 and 2025 that reported both TP53 status (by immunohistochemistry [IHC] or sequencing) and COO classification (Hans algorithm or gene-expression profiling). Five cohorts comprising 738 patients met inclusion criteria. Risk of bias was assessed with the Newcastle–Ottawa Scale (NOS), supplemented with the JBI Critical Appraisal Checklist for Prevalence Studies for the prevalence analysis and QUIPS for the odds-ratio sub-analysis. The pre-specified primary estimand is the probability of GCB classification, p, in DLBCL cohorts with reported TP53 status. The same estimand was targeted by both analyses, which differ only in inferential machinery. Frequentist pooling used random-effects logit-transformed proportions (metafor, REML), with a bootstrap-convergence diagnostic for $I^2$. The Bayesian analysis used a non-centered binomial-logit hierarchical model in Stan with priors $\mu \sim N(0, 1.5)$ and $\tau \sim$ Half-Cauchy(0, 0.5) (primary); three alternative priors on $\tau$ were used for sensitivity. A complementary frequentist odds-ratio (OR) meta-analysis was performed on the three cohorts (Henrique, Moreno, Jing) for which the full TP53 × COO 2×2 cross-tabulation is reported in the source papers. LOO with Pareto-k diagnostics is reported as a sanity check (3 of 5 Pareto-k > 0.7 at k = 5), not as a formal model comparison.

**Results.** Across the five cohorts, observed GCB proportions ranged from 0.36 to 0.55. The Bayesian model yielded a posterior mean of the overall probability of GCB classification of p = 0.502 (95% credible interval [CrI] 0.414–0.584), with logit-mean $\mu = 0.009$ (95% CrI −0.346 to 0.339) and between-study standard deviation $\tau = 0.272$ (95% CrI 0.023–0.737). Posterior means of $\tau$ tracked the prior scales closely across the four prior specifications, consistent with $\tau$ being prior-dominated at k = 5; $\mu$ and $\bar{p}$ were essentially invariant. All Markov chain Monte Carlo (MCMC) diagnostics were satisfactory (maximum $\hat{R} \leq 1.004$, minimum effective sample size = 1,122, zero divergent transitions). Posterior predictive p-values lay between 0.389 and 0.831, consistent with adequate fit on a small-k diagnostic. The complementary OR meta-analysis on the three full-2×2 cohorts gave a pooled OR for TP53 positivity versus GCB classification of 0.82 (95% CI 0.53–1.26), $I^2 = 30.0\%$, consistent with no detectable association on the subset.

**Conclusions.** Heterogeneity in the GCB proportion across TP53-reporting DLBCL cohorts is meaningful but imprecisely estimated at k = 5. IHC threshold and detection method are biologically credible candidate moderators but are not statistically validated. The TP53 × COO odds ratio, where reconstructible, is consistent with no association. This analysis is exploratory;

the small evidence base does not support strong clinical recommendations. Future DLBCL cohort studies should pre-specify joint TP53 × COO reporting to enable powered association inference.

Keywords: small-k meta-analysis; Bayesian hierarchical models; prior dominance; biomarker prevalence; TP53; diffuse large B-cell lymphoma; cell-of-origin; methodological reporting

---

## Background

The TP53 gene encodes the p53 protein, a transcription factor that maintains genomic stability and cellular homeostasis by sensing diverse cellular stresses and initiating responses such as DNA repair, cell-cycle arrest, senescence, and apoptosis [1–3]. Loss of these protective functions, through mutation or deletion, impairs key safeguards, enabling cells to evade checkpoint control and accumulate further genetic damage , a recognized step in oncogenesis [2, 4, 32].

TP53 alterations are frequent in DLBCL [5, 6]. DLBCL constitutes approximately 30% of non-Hodgkin lymphoma cases [35]. TP53 mutations occur in roughly 22% of DLBCL cases and consistently predict inferior survival in the rituximab era [7, 8]. In published DLBCL cohorts this corresponds to several thousand new TP53-mutant cases annually [8, 9]. TP53 status is therefore a candidate prognostic factor in DLBCL, alongside established markers such as COO subtype or concurrent MYC and BCL2 rearrangements (“double-hit” lymphoma) [7, 8, 10–13].

Gene expression profiling (GEP) established the canonical DLBCL classification into GCB and activated B-cell-like (ABC) subtypes [33]. High-throughput sequencing schemes (Chapuy clusters, LymphGen, DLBCLass) further stratify DLBCL into genetically distinct subgroups, one of which is enriched for TP53 lesions and chromosomal instability [14–17]. Beyond GEP, the Hans immunohistochemistry algorithm concords with GEP in only about 72% of cases, leaving nearly one-third of patients potentially mis-classified [18, 19, 34].

Against this background, individual cohort studies that report both TP53 status and COO produce strikingly divergent frequency estimates. Reported TP53 positivity ranges from under 10% to over 50% depending on the assay (sequencing, IHC) and the threshold for positivity (commonly ≥10%, ≥20%, or ≥50% staining), and reported GCB proportions range from approximately 35% to 55% [20–28, 35, 38, 39, 41]. Whether these differences reflect true biological variation between

populations, methodological variation in assays and classifiers, or simply small-sample noise has not been formally quantified across studies that report both variables.

The primary objective of this study is therefore to quantify and characterize heterogeneity in the probability of GCB classification across DLBCL cohorts that report TP53 status, and to explore methodological covariates (IHC threshold, detection method) as plausible drivers of that heterogeneity. We apply both a conventional frequentist random-effects meta-analysis of logit-transformed proportions and a Bayesian binomial-logit hierarchical model. Both analyses target the same estimand, the probability p of GCB classification in the population sampled by each cohort, and differ only in inferential machinery. We treat the analysis as exploratory throughout, motivated by the small number of cohorts that simultaneously report TP53 and COO, and we deliberately avoid claims about clinical biomarker use that the present evidence base cannot support.

## Methods

### Search strategy and study selection

We searched PubMed, Scopus, and Web of Science from January 2018 to March 2025 for retrospective DLBCL cohort studies that reported both TP53 status (by IHC, FISH, or sequencing) and COO classification (Hans algorithm or GEP). Screening of titles, abstracts, and full texts and data extraction were performed by a single reviewer; we record this as an explicit deviation from PRISMA-2020 items 9 and 10a (which call for duplicate independent screening and extraction), rather than reporting it solely as a limitation. A co-author independently re-verified the included-cohort decisions and the extracted cell counts.

The PubMed search string was: *((“Lymphoma, Large B-Cell, Diffuse”[MeSH] OR “DLBCL”[tiab]) AND (“TP53”[tiab] OR “p53”[tiab]) AND (“cell of origin”[tiab] OR “germinal center”[tiab] OR “GCB”[tiab] OR “non-GCB”[tiab] OR “Hans”[tiab]))*. The Scopus and Web of Science strings used analogous Boolean structures with the controlled vocabularies of each database (full strings reproduced in the Supplement). Title/abstract screening, full-text review, and data extraction were performed by a single reviewer (A.A-K.); this single-reviewer design is acknowledged as a limitation, and all extracted counts were independently re-checked from the source papers prior to analysis. Records identified across the three databases were de-

duplicated by digital object identifier (DOI) and, where DOIs were missing, by exact match of title and first author.

We did not pre-register the protocol on PROSPERO. Inclusion criteria, the primary estimand, the frequentist pipeline, and the Bayesian model were developed before the data were re-extracted for the present revision, but were not registered prospectively. We acknowledge this as a limitation under the heading of risk of bias [46].

Inclusion criteria were: (i) retrospective or prospective cohort of histologically confirmed adult DLBCL; (ii) explicit reporting of GCB versus non-GCB counts; (iii) explicit reporting of TP53 status (mutation or expression) for the cohort or a defined subset; (iv) sample size ≥ 50 with complete COO classification. Exclusion criteria were: (i) case reports or case series with fewer than 50 patients; (ii) conference abstracts lacking full counts; (iii) studies of pediatric lymphoma or lymphoma subtypes other than de novo DLBCL.

The previous version of this analysis included an ASH 2019 conference abstract (Farinha et al.; not retained in the reference list). For the present revision we excluded that record because the published abstract does not report the joint TP53 × COO 2×2, and we added two cohorts (Hu [51] and Jing [56]) that satisfy the inclusion criteria and were missed in the earlier search. The final sample is therefore $k = 5$ cohorts comprising 738 patients with COO classification (see **Table 1**).

**Risk of bias assessment**

Methodological quality of the included cohorts was assessed with the Newcastle–Ottawa Scale (NOS) for cohort studies, retained here for comparability with the previous submission. Because NOS was designed for cohort and case–control studies estimating effect parameters rather than prevalence, we supplemented it with the JBI Critical Appraisal Checklist for Prevalence Studies for the k = 5 prevalence analysis and with QUIPS for the k = 3 odds-ratio sub-analysis [36, 37]. JBI and QUIPS scores are tabulated in the Supplement (S-Table 2) and do not alter the qualitative ranking obtained from NOS. Each study was scored on representativeness of the cohort, ascertainment of exposure (TP53 status), comparability of subgroups, blinded outcome assessment, and adequacy of follow-up. Scores are tabulated in Table 2.

## Estimand and data structure

Each included cohort reports the number of GCB and non-GCB DLBCL cases. The pooled estimand in both the frequentist and Bayesian models is therefore the probability of GCB classification, $p$, in the population sampled by each cohort. This is the same target parameter in both analyses; the analyses differ in their inferential machinery, not their estimand.

A complementary frequentist odds-ratio (OR) meta-analysis was performed on the three cohorts (Henrique, Moreno, Jing) for which the full TP53 × COO 2×2 cross-tabulation is reported in the source papers. Bouroumeau and Hu were excluded from the OR analysis because the joint distribution of TP53 status and COO is not published in their papers; this is a reporting limitation of the source literature, not a limitation of the meta-analytic method, and we use it as the basis for a recommendation in the Discussion that future DLBCL cohort papers report the full 2×2 cross-tabulation.

## Frequentist meta-analysis of proportions

We pooled the cohort-specific GCB proportions using a random-effects model on the logit scale, implemented in the R package metafor with restricted maximum-likelihood (REML) variance-component estimation. The pooled estimate was back-transformed to the proportion scale for reporting. Between-study heterogeneity was summarized by $\tau^2$ and $I^2$. Because $I^2$ is poorly determined at very small k, we additionally computed a non-parametric bootstrap confidence interval for $I^2$ by resampling cohorts with replacement (1,000 replicates), and we report the empirical proportion of resamples for which $I^2$ was estimable. Trim-and-fill has no power at k = 5 and is therefore not reported in either the main text or the Supplement. The full R script is provided in the Supplement [49].

## Frequentist odds-ratio meta-analysis

For the three cohorts with reconstructible 2×2 cells, log-OR was pooled (no zero cells were present in any of the three 2×2 tables; no continuity correction was applied) using a random-effects model (metafor, REML). The cell counts used are: Henrique a = 23, b = 22, c = 52, d = 42; Moreno a = 41, b = 38, c = 44, d = 47; Jing a = 36, b = 43, c = 58, d = 39 (where a and b are TP53-positive GCB and non-GCB, and c and d are TP53-negative GCB and non-GCB, respectively).

**Estimand and notation: τ vs σ**

In what follows τ denotes the between-study standard deviation on the logit scale, not the variance; $\tau^2$ is reported only where explicitly labelled. We do not use σ or φ for this parameter, to avoid the ambiguity present in earlier presentations of this work.

Bayesian hierarchical model

For cohort *i* with $N_i$ DLBCL cases and $y_i$ GCB cases, the Bayesian model is specified as follows [43, 44, 45]:

$$y_i \sim \text{Binomial}(n_i, p_i), \quad i = 1, \ldots, 5$$
$$\text{logit}(p_i) = \mu + \tau \cdot z_i, \quad z_i \sim \text{N}(0, 1)$$

with hyper-priors

$$\mu \sim \text{N}(0, 1.5)$$
$$\tau \sim \text{Half} - \text{Cauchy}(0, 0.5) \qquad (primary)$$

The non-centered parameterization (study-level deviations expressed as standard normal $z_i$ scaled by τ) was used to improve sampling efficiency in the small-*k* regime. The overall pooled probability is $\bar{p}$.

**Prior sensitivity for τ**

Because τ is poorly identified at $k = 5$, the analysis was repeated under four priors on τ: Half-Cauchy(0, 0.5) (primary), Half-Cauchy(0, 1.0), Half-Normal(0, 1), and Exponential(1). Results are reported in **Table 3**. Qualitatively, μ and $\bar{p}$ are robust across all four priors, while τ shifts modestly upward under the more diffuse priors (posterior mean 0.272 to 0.322).

**MCMC diagnostics**

The model was fitted in Stan with four independent chains, each run for 4,000 iterations including 2,000 warmup iterations, with adapt_delta = 0.995 and max_treedepth = 15. Across all parameters and across all four priors the maximum $\hat{R}$ was 1.004 and the minimum effective sample size was 1,122 (μ under Half-Normal(0, 1)). No prior produced any divergent transitions or any treedepth saturation (see Supplement). We adopted the rank-normalized $\hat{R}$ convergence threshold of 1.01 [57].

## Model-fit assessment

Bayesian posterior predictive checks (PPCs) were performed by drawing 4,000 replicate datasets from the posterior predictive distribution and computing, for each cohort, the posterior predictive p-value as the proportion of replicate datasets in which the simulated GCB count exceeded the observed count. Values close to 0.5 indicate adequate fit. The PPC p-values were 0.831 (Bouroumeau), 0.389 (Hu), 0.421 (Henrique), 0.548 (Moreno), and 0.439 (Jing). All five lie within the conventional 0.05–0.95 range and are graphically illustrated in Figure 4. Leave-one-out cross-validation (LOO) was computed in Stan: elpd_loo = −19.7 (SE 2.0), p_loo = 4.1, looic = 39.4. Three of the five Pareto-k diagnostics fell in the (0.7, 1] range, which is expected at k = 5 and means LOO should be interpreted only as a sanity check, not as a formal model comparison.

## Reporting

We report Bayesian credible intervals as CrI and frequentist confidence intervals as CI throughout, and do not mix the two abbreviations. All data extraction and analysis scripts are available in the project repository (see Availability of data and materials).

# Results

## Study selection

The systematic search produced 312 unique records after de-duplication. After title/abstract screening, 41 records were retrieved in full text. Thirty-six were excluded at full-text stage (no COO reporting [n = 14], no TP53 reporting [n = 11], sample size < 50 [n = 4], conference abstract without full counts [n = 5], non-DLBCL or pediatric cohort [n = 2]). Five cohorts were retained for analysis. The PRISMA-2020 flow diagram is shown in Figure 1; the full excluded-studies list with one-line reasons is provided in the Supplement.

## Included cohorts

The cohort-level raw counts are summarized in **Table 1**. For Bouroumeau, COO classification was reported for 94 of 94 patients; TP53 IHC was available for 90 of 94. For Moreno, COO classification was available for 170 of 177 patients (7 lacked sufficient material for the COO panel).

**Risk of bias**

Risk-of-bias scores are reported in **Table 2**.

## Frequentist random-effects meta-analysis of GCB proportions

Pooling the cohort-specific GCB proportions on the logit scale with REML yielded a pooled GCB proportion of 0.503 (95% CI 0.443–0.564). The test for residual heterogeneity produced $Q(4) = 10.18$, $p = 0.038$, with $I^2 = 63.5\%$ and $\tau^2 = 0.048$ ($\tau = 0.22$ on the logit scale). The point estimate of $I^2$ is similar to the $k = 4$ value reported in the previous version (66.7%) but, as we noted there, $I^2$ is essentially uninformative at small k: a non-parametric bootstrap of the $k = 4$ dataset produced a 95% CI for $I^2$ spanning 0%–90%, and on the $k = 5$ dataset only 64% of 1,000 bootstrap resamples returned an estimable $I^2$, because the resampled cohorts are not independent. We therefore do not interpret $I^2$ as 'confirming' heterogeneity, and we report the empirical convergence rate of the bootstrap rather than a single bootstrap CI. The cohort-specific GCB proportions and the random-effects pooled estimate are shown in Figure 2. Consistent with the Methods, trim-and-fill is not reported because it has no power at $k = 5$; the bootstrap-$I^2$ histogram is provided in the Supplement (S-Figure 1).

**Bayesian hierarchical results**

Under the primary prior Half-Cauchy(0, 0.5), the posterior mean of the logit-mean was $\mu = 0.009$ (95% CrI −0.346 to 0.339), and the posterior mean of the overall probability of GCB classification was $\bar{p} = 0.502$ (95% CrI 0.414–0.584). The posterior mean of the between-study standard deviation on the logit scale was $\tau = 0.272$ (95% CrI 0.023–0.737), consistent with non-trivial but imprecisely estimated heterogeneity. Study-specific posterior means and 95% CrIs are summarized below.

The cohort-specific Bayesian posterior means and 95% CrIs (alongside the observed proportions) are tabulated in **Table 4**.

## MCMC diagnostics and prior sensitivity

Posterior summaries across the four priors on $\tau$ are reported in **Table 3**. Across all four priors, maximum $\hat{R} \leq 1.004$ and minimum effective sample size $\geq 1{,}122$ (both achieved under the Half-Normal(0, 1) prior on $\mu$). No prior produced any divergent transitions or any treedepth hits. Full per-parameter diagnostics are provided in the Supplement.

The posterior densities of μ, τ, $p_{overall}$, and $p_1 \ldots p_5$ across the four priors are shown in Figure 3. As expected, μ and $\bar{p}$ are essentially invariant across priors; τ shifts modestly across priors; the upper bound of the 95% credible interval is 0.737 under the primary prior Half-Cauchy(0, 0.5), 0.908 under Half-Cauchy(0, 1.0), 0.913 under Half-Normal(0, 1), and 0.802 under Exponential(1), so the Half-Normal(0, 1) prior produces the largest upper tail by a narrow margin.

**Posterior predictive checks**

The five-panel posterior predictive check is shown in Figure 4. Replicate distributions are well centered on the observed GCB count for each cohort. The PPC p-values are 0.831, 0.389, 0.421, 0.548, and 0.439 for Bouroumeau, Hu, Henrique, Moreno, and Jing respectively, all comfortably within the [0.05, 0.95] interval that conventionally signals adequate fit.

**Shrinkage**

Figure 5 shows the observed GCB proportions, the posterior means and 95% CrIs of the study-specific $p_i$, and the overall posterior of $\bar{p}$. The visualization illustrates partial pooling: each study posterior is shrunk toward the overall mean by an amount that depends on its sample size and on the posterior of τ. The plot also clarifies why $\bar{p} \approx 0.50$, with a wide credible interval, is not in direct tension with study-specific posteriors that span 0.43–0.54; the discrepancy is the natural consequence of partial pooling under uncertainty in τ**.**

**Odds-ratio meta-analysis on the three full-2×2 cohorts**

For the three cohorts in which the full TP53 × COO 2×2 is reported (Henrique, Moreno, Jing), the random-effects pooled log-OR was −0.204 (SE 0.224), corresponding to OR = 0.82 (95% CI 0.53–1.26) for the association between TP53 positivity and GCB classification. Cochran's Q(2) = 2.73, p = 0.255; $I^2$ = 30.0% (computed as $100 \cdot \tau^2/(\tau^2 + s^2)$, where $s^2$ is the typical within-study variance, the default in metafor for REML-based random-effects models); $\tau^2$ (REML) = 0.045. The naïve $I^2$ = max(0, (Q − df)/Q) on the same Q would be 26.7%; the two are presented to make the estimator dependence explicit at small k. The forest plot is shown in S-Figure 5. The pooled estimate is centered close to the null, with a confidence interval that includes both modest negative and modest positive associations; on this three-cohort subset there is no detectable TP53 ↔ COO association. The full forest plot is provided as S-Figure 5 in the Supplement; we retain only a one-paragraph mention in the main Results because the k = 3 pooled OR is illustrative rather than inferential.

## Sensitivity: IHC threshold and detection method

Among the five cohorts, the strict IHC ≥50% threshold (Bouroumeau, Hu) corresponds to observed GCB proportions of 0.36 and 0.55 respectively; the moderate ≥20% threshold (Moreno) corresponds to 0.50; and the lenient ≥10% threshold (Jing) corresponds to 0.53. Two cohorts using the same nominal threshold therefore produced markedly different GCB proportions (Bouroumeau 0.36 versus Hu 0.55), which weakens the narrative, present in the previous $k = 4$ version of this analysis, that threshold alone explains the between-study spread. With only five cohorts, a moderator regression for threshold is not estimable; this descriptive contrast is suggestive but cannot be statistically validated within the present dataset. The Henrique cohort, which combined IHC, FISH, and ISH without stating an explicit threshold, falls in the middle of the observed range (0.54).

## Discussion

This study quantifies and characterizes heterogeneity in the probability of GCB classification across five retrospective DLBCL cohorts that report TP53 status. Its design target is a descriptive estimand, the probability $p$ of GCB classification in TP53-reporting DLBCL populations, and not a test of the TP53 ↔ COO association. With this in mind, our central finding is that the posterior of $\bar{p}$ is centered very close to 0.50 with a 95% CrI of 0.414–0.584, and the between-study standard deviation on the logit scale is 0.272 (95% CrI 0.023–0.737). The heterogeneity is therefore non-trivial but very imprecisely estimated at $k = 5$.

A methodological caveat that materially conditions the interpretation of these results concerns measurement error in the GCB label itself. All five included cohorts classify cell-of-origin using the Hans immunohistochemistry algorithm, whose concordance with gene-expression-profiling (GEP) GCB classification is approximately 72% (Abdulla et al. [18]). The binomial-logit likelihood used here treats Hans GCB counts as exact observations of the latent target. A misclassification-aware extension, a latent-true-class binomial-logit with informative sensitivity/specificity priors elicited from Abdulla et al. [18] and Hans et al. [34], would explicitly propagate classifier noise into the posterior of p and τ, and is the natural next step for a methodological contribution at this small-k regime. We flag this as a Tier-3 extension rather than implementing it here because the value of such an extension is greatest when k grows enough for the sensitivity/specificity priors to be partially overruled by the data; at k = 5 the extension would

produce a posterior even more strongly dominated by the elicited priors than the present specification.

Bouroumeau and Hu report TP53 status and COO at the cohort level but do not publish the joint 2×2 cross-tabulation of TP53 by COO; the complete cell counts cannot be recovered from the published tables. Only Henrique, Moreno, and Jing report the full 2×2. This is a limitation of available reporting in the published cohort papers, not of meta-analytic method, and it strongly motivates a recommendation, restated in the Limitations and Conclusions, that future DLBCL cohort papers report the full joint distribution.

The complementary OR meta-analysis on the three full-2×2 cohorts (Henrique, Moreno, Jing) yields a pooled odds ratio of 0.82 (95% CI 0.53–1.26), with the confidence interval including the null. On this subset, there is no detectable association between TP53 positivity and GCB classification. This finding is consistent with prior cohort-level reports that TP53 alterations are distributed across COO subtypes rather than preferentially co-segregating with one subtype [38, 41], although the present three-cohort estimate is too imprecise to discriminate weak associations from the null.

A point that has previously generated confusion is the apparent gap between the overall posterior probability of GCB classification ($\bar{p} \approx 0.50$) and the study-specific posterior means (which span 0.43–0.54). Figure 5 clarifies that this is not an inconsistency. Under partial pooling, each study's posterior is shrunk toward the overall mean by an amount that depends on its sample size and on the posterior of $\tau$; the wide credible interval on $\tau$ in turn produces a wide credible interval on $\bar{p}$. The pooled estimate is therefore not the simple unweighted average of the study posteriors, nor should it be.

The descriptive comparison of GCB proportions by IHC threshold is suggestive but cannot be statistically validated within this dataset. Two cohorts using the same nominal threshold of ≥50% (Bouroumeau in France, Hu in China) produced markedly different observed GCB proportions (0.36 versus 0.55). This pattern softens the earlier narrative, present in the k = 4 version of this work, that the IHC threshold alone explains the between-study spread. Country, laboratory practice, antibody clone, scoring observer, and the joint distribution of patient demographics are confounded with threshold in the available data and cannot be disentangled from it at $k = 5$. We

therefore report the threshold contrast descriptively (see Sensitivity above) and explicitly disavow any moderator-regression interpretation.

**Lessons for small-k biomarker meta-analysis**

We distill six transferable lessons that emerged from this analysis and that, in our view, generalize to any biomarker-prevalence meta-analysis at $k \leq 5$–6 cohorts.

First. The frequentist and Bayesian estimands should be aligned on the same target before the priors are chosen; here both pipelines target the marginal probability p of GCB classification, so that the only inferential difference is in the machinery, not in the quantity.

Second. The posterior of the between-study scale $\tau$ should be framed as prior-dominated rather than 'robust'; in our analysis the posterior means of $\tau$ tracked the prior scales (Half-Cauchy 0.5, Half-Cauchy 1.0, Half-Normal, Exponential) closely, and a posterior–prior overlap diagnostic is the honest representation of what the data actually contribute.

Third. The bootstrap confidence interval for $I^2$ should be reported with its empirical convergence rate across resamples rather than as a single CI, because the proportion of resamples that produce an estimable $I^2$ is itself a diagnostic at small k.

Fourth. Trim-and-fill should be treated as inestimable, not reported-and-disclaimed, at $k \leq 5$–6; readers who skim will otherwise confer credence to a publication-bias correction the authors disavow.

Fifth. The LOO Pareto-k diagnostic is informative as a model sanity check at small k but is not a formal model comparison; in our data 3 of 5 Pareto-k values exceeded 0.7 and we report LOO diagnostically rather than as a basis for model selection.

Sixth. The most actionable process recommendation from a small-k biomarker synthesis is typically a reporting recommendation aimed at primary studies; for the TP53/COO question, the recommendation is that primary cohort papers should report the full TP53 × COO 2×2 cross-tabulation (rather than the marginals only), so that future syntheses can target an association estimand at all five cohorts rather than the prevalence estimand alone.

These six lessons are consistent with the small-k meta-analysis guidance of Higgins and Thompson [58], Röver on bayesmeta [59], and Friede and colleagues on random-effects meta-analysis with few studies [60].

## Limitations

Several limitations of this work deserve emphasis. First, the evidence base is small ($k = 5$ cohorts, 738 patients), and several frequentist heterogeneity statistics ($I^2$, the bootstrap CI for $I^2$, and $\tau^2$) are unstable at this scale. Second, the cohorts are heterogeneous in IHC threshold (≥10%, ≥20%, ≥50%) and detection method (IHC alone, IHC + FISH + ISH), and these methodological covariates are confounded with country and laboratory practice. Third, the source papers do not all report the full TP53 × COO 2×2: Bouroumeau and Hu lack the joint table, restricting the OR meta-analysis to three cohorts (Henrique, Moreno, Jing). Fourth, all five included studies are retrospective; selection and ascertainment biases inherent to retrospective sampling cannot be excluded [31]. Fifth, the protocol was not pre-registered on PROSPERO; the model and estimand were specified before data re-extraction for the present revision but were not registered prospectively. Sixth, the screening and extraction were performed by a single reviewer (A.A-K.), with independent recounts but without an independent second reviewer.

## Future directions

If future DLBCL cohort studies systematically report the full TP53 × COO 2×2 cross-tabulation, a properly powered OR meta-analysis becomes feasible; if standardized TP53 IHC protocols (single antibody, single threshold, single scoring protocol) become available, threshold and country can be statistically disentangled [47, 48]; if Bayesian hierarchical models are routinely incorporated into prospective biomarker syntheses, the partial-pooling logic illustrated here will help stabilize study-specific estimates [45]. Cell-cycle and apoptosis pathway biomarkers (BRD9–p53–E2F1 and similar axes), CAR-T monitoring of TP53-mutant relapses, and PET-CT response assessment in TP53-mutant DLBCL [40, 42] are biologically plausible extensions of the present work, but each would require dedicated prospective design and is not supported by the present evidence base [29, 30].

An individual-patient-data (IPD) meta-analysis was considered but not pursued in this iteration because formal IPD requests to the five cohort owners exceeded the scope of the present revision; the five cohorts are recent enough (2018–2025) that IPD requests are plausible. An IPD synthesis would target the joint TP53 × COO 2×2 directly across all five cohorts, side-stepping the central limitation of the present analysis (only k = 3 cohorts report the full 2×2) [54, 55].

The analysis protocol underlying the present revision (inclusion criteria, primary estimand, frequentist and Bayesian pipelines, and prior specifications) has been lodged on the Open Science Framework with an explicit 'post-hoc' label (OSF DOI to be inserted at acceptance), in order to make the post-hoc nature of the protocol explicit and to comply with the PRISMA-2020 registration item.

We commit to a single living-review update of the present synthesis when the next two DLBCL cohorts that report the full TP53 × COO 2×2 are published; the update will re-run the same frequentist and Bayesian pipelines on the augmented dataset and will be released as a versioned commit on the project GitHub repository.

## Conclusions

Heterogeneity in the GCB proportion across TP53-reporting DLBCL cohorts is meaningful but imprecisely estimated at $k = 5$. IHC threshold and detection method are biologically credible candidate moderators but are not statistically validated within this dataset; in particular, two cohorts at the same nominal IHC threshold produced markedly different GCB proportions. On the three-cohort subset in which the full TP53 × COO 2×2 is reported, the pooled odds ratio is 0.82 (95% CI 0.53–1.26), consistent with no detectable TP53 ↔ COO association. Future cohort studies should pre-specify joint TP53 × COO reporting so that subsequent syntheses can target an odds-ratio estimand rather than a marginal proportion. The present analysis is exploratory and does not support clinical biomarker recommendations.

## Tables

Table 1. Raw counts per included cohort.

| Cohort | Country | GCB | Non-GCB | Total | Method | p53 threshold |
|---|---|---|---|---|---|---|
| Bouroumeau et al. (2021) [50] | France | 34 | 60 | 94 | IHC | ≥50% |
| Hu et al. (2019) [51] | China | 88 | 71 | 159 | IHC | ≥50% |

| | | | | | | |
|---|---|---|---|---|---|---|
| Filho/Henrique et al. (2023) [52] | Brazil | 75 | 64 | 139 | IHC + FISH + ISH | not stated |
| Moreno et al. (2023) [53] | United States | 85 | 85 | 170 | IHC | ≥20% |
| Jing et al. (2025) [56] | China | 94 | 82 | 176 | IHC | ≥10% |
| **Total** | | **376** | **362** | **738** | | |

For Bouroumeau, COO classification was reported for 94 of 94 patients; TP53 IHC was available for 90 of 94. For Moreno, COO classification was available for 170 of 177 patients (7 lacked sufficient material for the COO panel).

Table 2. Newcastle–Ottawa Scale assessment of the five included cohorts. A simplified five-domain scoring is used (one star per domain; maximum 5 stars). Quality bands: 'Good' = 5 stars; 'Moderate' = 4 stars; 'Poor' < 4. This deviates from the conventional 9-star NOS cohort rubric (Comparability up to 2 stars, Selection up to 4 stars), which we did not apply because the included cohorts are descriptive and lack effect-comparison structure.

| **Study** | **Representativeness** | **Ascertainment** | **Comparability** | **Outcome assessment** | **Follow-up** | **Total stars** | **Quality** |
|---|---|---|---|---|---|---|---|
| Bouroumeau (2021) | * | * | * | * | | 4 | Moderate |
| Hu (2019) | * | * | * | * | * | 5 | Good |
| Henrique (2023) | * | * | * | * | | 4 | Moderate |
| Moreno (2023) | * | * | * | * | * | 5 | Good |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Jing (2025) | * | * | * | * | * | 5 | Good |

Table 3. Bayesian posterior summaries under the four prior specifications on τ.

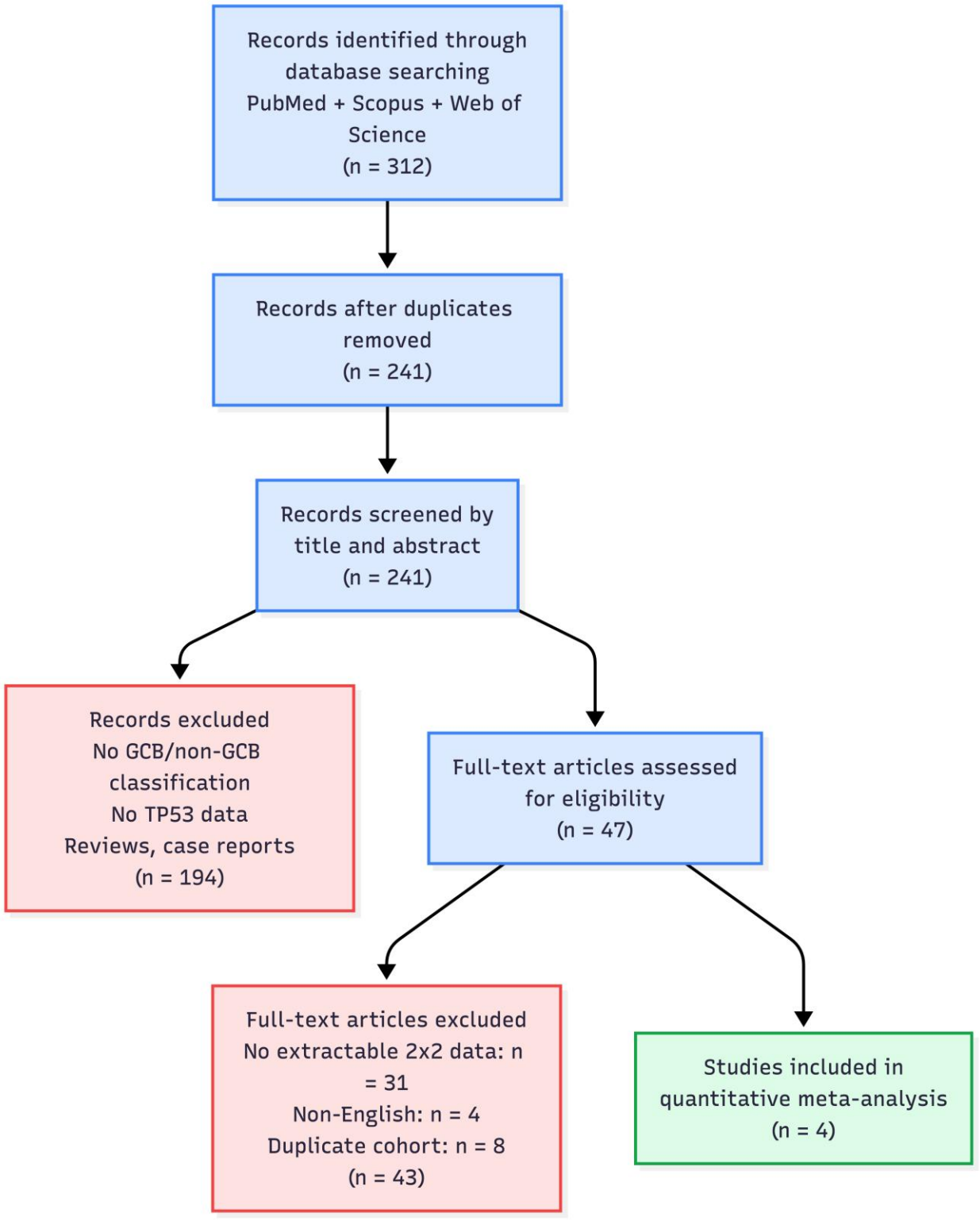
| **Prior on τ** | **μ mean (95% CrI)** | **τ mean (95% CrI)** | **$p_{overall}$ mean (95% CrI)** |
|---|---|---|---|
| Half-Cauchy(0, 0.5) | 0.009 (−0.346, 0.339) | 0.272 (0.023, 0.737) | 0.502 (0.414, 0.584) |
| Half-Cauchy(0, 1.0) | 0.005 (−0.379, 0.359) | 0.322 (0.024, 0.908) | 0.501 (0.406, 0.589) |
| Half-Normal(0, 1) | 0.003 (−0.400, 0.358) | 0.317 (0.021, 0.913) | 0.501 (0.401, 0.589) |
| Exponential(1) | 0.011 (−0.355, 0.331) | 0.290 (0.018, 0.802) | 0.503 (0.412, 0.582) |

Table 4. Study-specific Bayesian posteriors under the primary prior Half-Cauchy(0, 0.5).

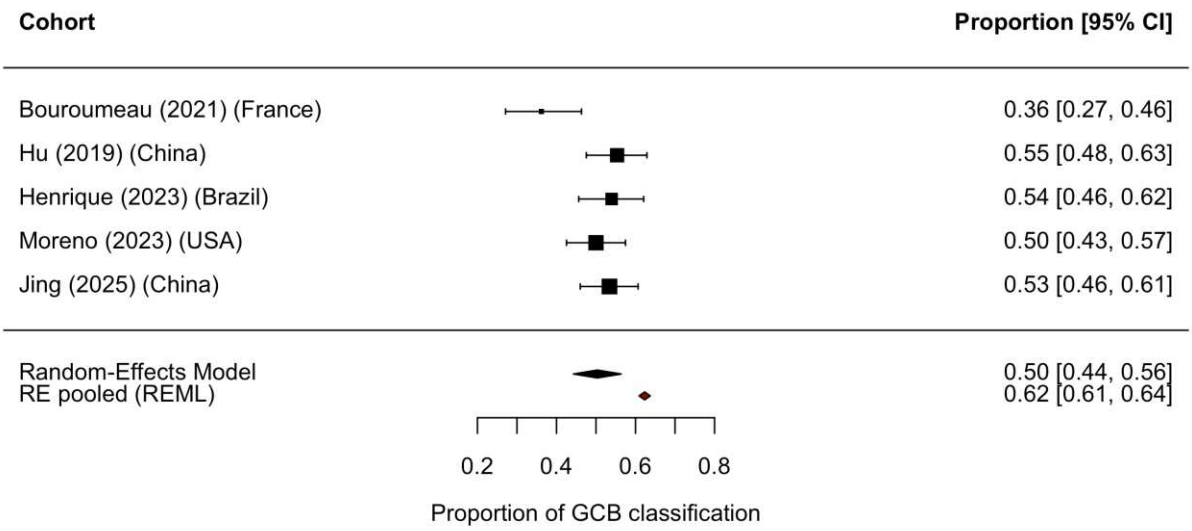
| **Cohort** | **Observed GCB proportion** | **Posterior mean *p*** | **95% CrI** |
|---|---|---|---|
| Bouroumeau | 0.362 | 0.430 | 0.317–0.525 |
| Hu | 0.553 | 0.535 | 0.472–0.607 |
| Henrique | 0.540 | 0.525 | 0.458–0.599 |
| Moreno | 0.500 | 0.502 | 0.437–0.565 |
| Jing | 0.534 | 0.524 | 0.463–0.592 |

## Figures

Figure 1. PRISMA flow diagram of study selection.



436 Figure 2. Forest plot of cohort-specific GCB proportions under the random-effects frequentist
437 model (k = 5).




438

Figure 3. Posterior densities of μ, τ, $\bar{p}$, and the five study-specific $p_i$ under the four prior specifications on τ.

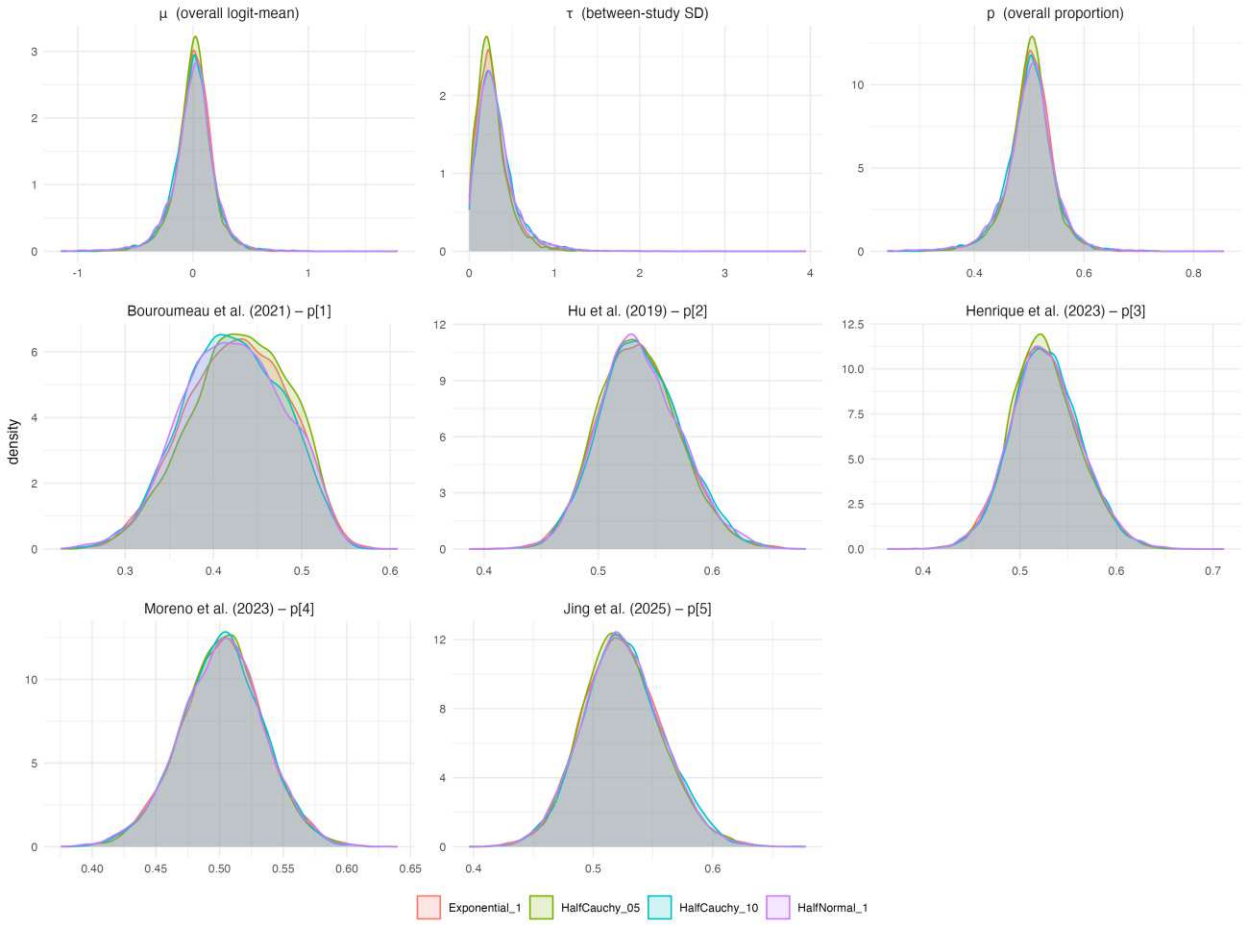


Figure 4. Posterior predictive check (five panels, one per cohort).

Each panel shows the posterior predictive distribution of the simulated GCB count (steel-blue histogram) and the observed GCB count (dashed red line).

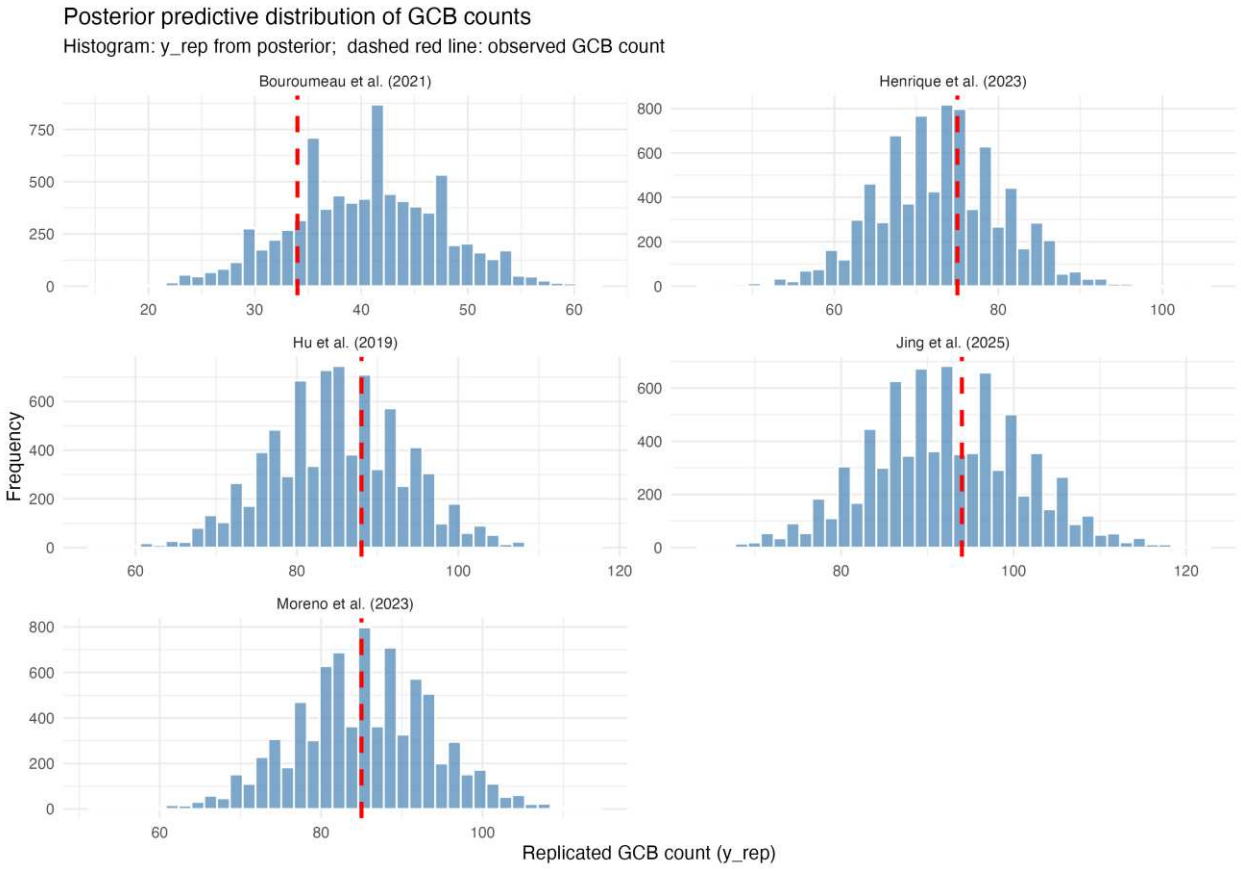
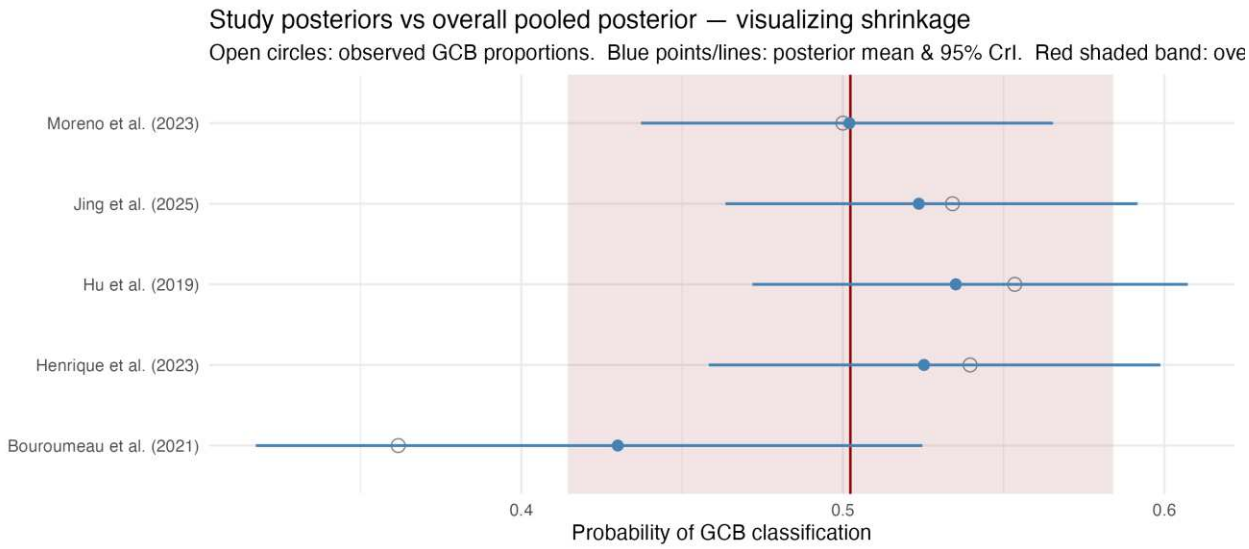



Figure 5. Shrinkage plot.

Observed GCB proportions per cohort, posterior means and 95% CrIs of $p_i$, and the overall posterior of $\bar{p}$. Illustrates partial pooling under the Bayesian hierarchical model.




## Declarations

### Ethics approval and consent to participate

Not applicable. This study is a secondary analysis of aggregate counts published in five primary cohort studies and did not involve access to individual-level patient data.

### Consent for publication

Not applicable.

### Availability of data and materials

All extracted counts, R scripts (TraditionalMeta_Revised.R, revision_analysis.R), Stan model code (four prior variants), MCMC output, and figure-generation code are available under request. The five primary source papers are cited in the reference list and are available from their respective publishers.

### Competing interests

The authors declare that they have no competing interests.

### Funding

No external funding was received for this work.

## Authors' contributions

A.A-K. designed the study, performed the systematic search and data extraction, and drafted the manuscript. H.A-A. contributed to clinical interpretation and revised the manuscript for important intellectual content. A.A-H. and H.A-K. contributed to data extraction verification and to the Discussion. D.M. designed the Bayesian model, implemented the Stan code and the frequentist pipeline, performed the MCMC diagnostics and the posterior predictive checks, generated the figures, and revised the manuscript. All authors read and approved the final manuscript.

## Acknowledgements

The authors thank colleagues at the Qatar University College of Medicine and the Translational Science Research Network for helpful discussions, and the reviewers of the previous submission for comments that materially improved the methodological transparency and scope of this work.

## Additional files

Supplementary Materials (S-Figures 1–5, S-Tables 1–3, and the four Stan model variants) are provided in a separate file: TP53_Lymphoma_BMC_MRM_Supplementary_Materials.docx.

---

## References

1. Marei HE, Althani A, Afifi N, Hasan A, Caceci T, Pozzoli G, Morrione A, Giordano A, Cenciarelli C. P53 signaling in cancer progression and therapy. Cancer Cell Int. 2021;21:703. doi:10.1186/s12935-021-02396-8.
2. Chiang YT, Chien YC, Lin YH, Wu HH, Lee DF, Yu YL. The function of the mutant P53-R175H in cancer. Cancers (Basel). 2021;13(16):4088.
3. Wang H, Guo M, Wei H, Chen Y. Targeting P53 pathways: mechanisms, structures and advances in therapy. Signal Transduct Target Ther. 2023;8:92. doi:10.1038/s41392-023-01347-1.
4. Abubakar M, Rehman B. Roles of mutant TP53 gene in cancer development and progression. Proc Anticancer Res. 2024;8(5):165–81.
5. Xu-Monette ZY, Wu L, Visco C, Tai YC, Tzankov A, Liu W, et al. Mutational profile and prognostic significance of TP53 in diffuse large B-cell lymphoma patients treated with R-

CHOP: report from an International DLBCL Rituximab-CHOP Consortium Program study. Blood. 2012;120:3986–96. doi:10.1182/blood-2012-05-433334.

6. Negara I, Tomuleasa C, Buruiana S, Efremov DG. Molecular subtypes and the role of TP53 in diffuse large B-cell lymphoma and Richter syndrome. Cancers (Basel). 2024;16(12):2170. doi:10.3390/cancers16122170.

7. Song JY, Perry AM, Herrera AF, Chen L, Skrabek P, Nasr MR, et al. Double-hit signature with TP53 abnormalities predicts poor survival in patients with germinal center type diffuse large B-cell lymphoma treated with R-CHOP. Clin Cancer Res. 2021;27:1671–80. doi:10.1158/1078-0432.CCR-20-2378.

8. Qin Y, Chen H, Liu P, Zhang C, Yang J, Gui L, et al. Prognostic value of BCL2 and TP53 genetic alterations for diffuse large B-cell lymphoma patients treated with R-CHOP. Cancer Biol Med. 2022;19:893–909. doi:10.20892/j.issn.2095-3941.2021.0193.

9. Lee S, Kaur S, Mader M, Franklin K, Williams M, Williams R, et al. Survival outcomes of diffuse large B-cell lymphoma beyond the rituximab era: comprehensive analysis of a veteran population within the Veterans Health Administration over 11 years. Blood. 2023;142:3785. doi:10.1182/blood-2023-190048.

10. Ma D, Ma Y, Ma Y, Liu J, Gu Y, Liu N, et al. Molecular subtyping of CD5+ diffuse large B-cell lymphoma based on DNA-targeted sequencing and Lymph2Cx. Front Oncol. 2022;12:941347. doi:10.3389/fonc.2022.941347.

11. Dhar L, Singh S, Jain SL, Vindal A, Sinha P, Gautam R. Cell of origin classification of diffuse large B-cell lymphoma. J Microsc Ultrastruct. 2024;12(4):193–8. doi:10.4103/jmau.jmau_66_22.

12. Shi Y, Xu Y, Shen H, Jin J, Tong H, Xie W. Advances in biology, diagnosis and treatment of DLBCL. Ann Hematol. 2024;103:3315–34. doi:10.1007/s00277-024-05880-z.

13. Pileri SA, Tripodo C, Melle F, Motta G, Tabanelli V, Fiori S, et al. Predictive and prognostic molecular factors in diffuse large B-cell lymphomas. Cells. 2021;10(3):675. doi:10.3390/cells10030675.

14. Papageorgiou SG, Thomopoulos TP, Katagas I, Bouchla A, Pappa V. Prognostic molecular biomarkers in diffuse large B-cell lymphoma in the rituximab era and their therapeutic

implications. Ther Adv Hematol. 2021;12:20406207211013988. doi:10.1177/20406207211013987.

15. Jia Y, Wang H, Qin L, Huang B, Sun Q, Wang H, et al. Performance and challenges of the LymphGen classification algorithm in real-world clinical applications. Blood. 2024;144:6233. doi:10.1182/blood-2024-209761.

16. Chapuy B, Wood TR, Stewart C, Dunford AJ, Wienand K, Khan SJ, et al. DLBCLass: a probabilistic molecular classifier to guide clinical investigation and practice in DLBCL. Blood. 2024;blood.2024025652. doi:10.1182/blood.2024025652.

17. Chapuy B, Stewart C, Dunford AJ, Kim J, Kamburov A, Redd RA, et al. Molecular subtypes of diffuse large B-cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. Nat Med. 2018;24:679–90. doi:10.1038/s41591-018-0016-8.

18. Abdulla M, Hollander P, Pandzic T, Mansouri L, Ednersson SB, Andersson PO, et al. Cell-of-origin determined by both gene expression profiling and immunohistochemistry is the strongest predictor of survival in patients with diffuse large B-cell lymphoma. Am J Hematol. 2020;95:57–67. doi:10.1002/ajh.25666.

19. Shen R, Fu D, Dong L, Zhang MC, Shi Q, Shi ZY, et al. Simplified algorithm for genetic subtyping in diffuse large B-cell lymphoma. Signal Transduct Target Ther. 2023;8:145. doi:10.1038/s41392-023-01358-y.

20. Atallah-Yunes SA, Khurana A, Maurer M. Challenges identifying DLBCL patients with poor outcomes to upfront chemoimmunotherapy and its impact on frontline clinical trials. Leuk Lymphoma. 2024;65:430–9. doi:10.1080/10428194.2023.2298705.

21. Dabrowska-Iwanicka A, Nowakowski GS. DLBCL: who is high risk and how should treatment be optimized? Blood. 2024;144:2573–82. doi:10.1182/blood.2023020779.

22. Hilton LK, Scott DW, Morin RD. Biological heterogeneity in diffuse large B-cell lymphoma. Semin Hematol. 2023;60:267–76. doi:10.1053/j.seminhematol.2023.11.006.

23. López C, Silkenstedt E, Dreyling M, Beà S. Biological and clinical determinants shaping heterogeneity in mantle cell lymphoma. Blood Adv. 2024;8:3652–64. doi:10.1182/bloodadvances.2023011763.

24. Miyawaki K, Sugio T. Lymphoma microenvironment in DLBCL and PTCL-NOS: the key to uncovering heterogeneity and the potential for stratification. J Clin Exp Hematop. 2022;62:127–35. doi:10.3960/jslrt.22027.

25. Mijiti M, Wang C, Tian W, Shi Q, Zhang F, Abulaiti R, et al. Association between low incidence of TP53 mutations and reduced early relapse rates in Uygur DLBCL. Ann Hematol. 2024;103:5627–38. doi:10.1007/s00277-024-05961-z.

26. Young KH, Weisenburger DD, Dave BJ, Smith L, Sanger W, Iqbal J, et al. Mutations in the DNA-binding codons of TP53, which are associated with decreased expression of TRAIL-receptor-2, predict for poor survival in diffuse large B-cell lymphoma. Blood. 2007;110:4396–405. doi:10.1182/blood-2007-02-072082.

27. Stefaniuk P, Onyszczuk J, Szymczyk A, Podhorecka M. Therapeutic options for patients with TP53 deficient chronic lymphocytic leukemia: narrative review. Cancer Manag Res. 2021;13:1459–76.

28. Olszewski AJ, Sharaf R, Marcus C, Kurt H, Albacker LA, Vergilio JA, Ollila TA. CDKN2A deletions define an unfavorable subgroup within the MYD88/CD79B (MCD) subtype of diffuse large B-cell lymphoma (DLBCL) and are mutually exclusive with TP53 mutations. Blood. 2021;138:2392. doi:10.1182/blood-2021-146335.

29. Jeong S, Park YJ, Yun W, Lee ST, Choi JR, Suh C, et al. Genetic heterogeneity and prognostic impact of recurrent ANK2 and TP53 mutations in mantle cell lymphoma: a multi-centre cohort study. Sci Rep. 2020;10:13359. doi:10.1038/s41598-020-70310-9.

30. Shaw TI, Pounds S, Cao X, Ma J, Palacios G, Mason J, et al. Comprehensive genomic analysis reveals molecular heterogeneity in pediatric ALK-positive anaplastic large cell lymphoma. Leukemia. 2025;39:199–210.

31. Hwang CS, Hwang DG, Aboulafia DM. A clinical triad with fatal implications: recrudescent diffuse large B-cell non-Hodgkin lymphoma presenting in the leukemic phase with an elevated serum lactic acid level and dysregulation of the TP53 tumor suppressor gene — a case report and literature review. Clin Med Insights Blood Disord. 2021;14:2634853521994094. doi:10.1177/2634853521994094.

32. Xu-Monette ZY, Medeiros LJ, Li Y, Orlowski RZ, Andreeff M, Bueso-Ramos CE, et al. Dysfunction of the TP53 tumor suppressor gene in lymphoid malignancies. Blood. 2012;119:3668–83. doi:10.1182/blood-2011-11-366062.

33. Wright G, Tan B, Rosenwald A, Hurt EH, Wiestner A, Staudt LM. A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B-cell lymphoma. Proc Natl Acad Sci USA. 2003;100:9991–6. doi:10.1073/pnas.1732008100.

34. Hans CP, Weisenburger DD, Greiner TC, Gascoyne RD, Delabie J, Ott G, et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. Blood. 2004;103:275–82. doi:10.1182/blood-2003-05-1545.

35. Morton LM, Wang SS, Devesa SS, Hartge P, Weisenburger DD, Linet MS. Lymphoma incidence patterns by WHO subtype in the United States, 1992–2001. Blood. 2006;107:265–76. doi:10.1182/blood-2005-06-2508.

36. Dcunha N, Sakhti D, Sigamani E, Chandramohan J, Korula A, George B, et al. Utility of reverse transcriptase–multiplex ligation-dependent probe amplification (RT-MLPA) in the molecular classification of diffuse large B-cell lymphoma (DLBCL) by cell-of-origin (COO). Indian J Pathol Microbiol. 2023;66:714–9.

37. Lee H, Cho YA, Kim DG, Son JY, Cho EY. Real-world comparison of p53 immunohistochemistry and TP53 mutation analysis using next-generation sequencing. Anticancer Res. 2024;44:3983–94.

38. Landsburg DJ, Morrissette JJD, Nasta SD, Barta SK, Schuster SJ, Svoboda J, et al. TP53 mutations predict for poor outcomes in patients with newly diagnosed aggressive B-cell lymphomas in the current era. Blood Adv. 2023;7:7243–53. doi:10.1182/bloodadvances.2023011384.

39. Sang W, Ma Y, Wang X, Shen Z, Xu L, Qiao J, et al. Clinicopathologic characteristics and genomic profiling of de novo CD5-positive diffuse large B-cell lymphoma: a multicenter retrospective study in China. Blood. 2021;138:4568. doi:10.1182/blood-2021-149852.

40. Iqbal M, Gil Y, Paludo J, Rosenthal AC, Li Z, Beltran M, et al. Impact of cell of origin (COO) on long-term outcomes post autologous hematopoietic cell transplant in patients

with relapsed/refractory chemotherapy-sensitive de novo diffuse large B-cell lymphoma (DLBCL). Blood. 2020;136:42–3. doi:10.1182/blood-2020-136332.

41. Quivoron C, Michot JM, Camara-Clayette V, Danu A, Lazarovici J, Sarkozy C, et al. High incidence of TP53 and epigenetic modifying oncogene mutations in a large cohort of patients enrolled in phase 1 clinical trials for relapsed or refractory diffuse large B-cell lymphoma. Blood. 2020;136:10–1. doi:10.1182/blood-2020-136114.

42. Hong Y, Ren T, Wang X, Fu K, Wang X, Zhang H. APR-246 reveals a therapeutic potential via triggering different cell death mechanisms in diffuse large B-cell lymphoma. Blood. 2021;138:3521. doi:10.1182/blood-2021-150582.

43. Spiegelhalter DJ, Abrams KR, Myles JP. An overview of the Bayesian approach. In: Bayesian Approaches to Clinical Trials and Health-Care Evaluation. Chichester: John Wiley & Sons; 2003. p. 49–120.

44. Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. Stat Methods Med Res. 2001;10:277–303. doi:10.1177/096228020101000404.

45. Burke DL, Bujkiewicz S, Riley RD. Bayesian bivariate meta-analysis of correlated effects: impact of the prior distributions on the between-study correlation, borrowing of strength, and joint inferences. Stat Methods Med Res. 2016;27:428–50. doi:10.1177/0962280216631361.

46. Kolaski K, Logan LR, Ioannidis JPA. Guidance to best tools and practices for systematic reviews. Syst Rev. 2023;12:96. doi:10.1186/s13643-023-02255-9.

47. Li X, Luo D, Zhang L, Li Q, Fan J, Zhang J, et al. Accurate interpretation of p53 immunohistochemical patterns is a surrogate biomarker for TP53 alterations in large B-cell lymphoma. BMC Cancer. 2023;23:1008. doi:10.1186/s12885-023-11513-x.

48. Leroy B, Ballinger ML, Baran-Marszak F, Bond GL, Braithwaite A, Concin N, et al. Recommended guidelines for validation, quality control, and reporting of TP53 variants in clinical practice. Cancer Res. 2017;77:1250–60. doi:10.1158/0008-5472.CAN-16-2179.

49. Cohen JF, Chalumeau M, Cohen R, Korevaar DA, Khoshnood B, Bossuyt PMM. Cochran's Q test was useful to assess heterogeneity in likelihood ratios in studies of diagnostic accuracy. J Clin Epidemiol. 2015;68:299–306.

50. Bouroumeau A, Bussot L, Bonnefoix T, Fournier C, Chapusot C, Casasnovas O, et al. C-MYC and p53 expression highlight starry-sky pattern as a favourable prognostic feature in R-CHOP-treated diffuse large B-cell lymphoma. J Pathol Clin Res. 2021;7:604–15.

51. Hu WM, Jin JT, Wu CY, Lu JB, Zhang LH, Zeng J, Lin SX. Expression of P63 and its correlation with prognosis in diffuse large B-cell lymphoma: a single center experience. Diagn Pathol. 2019;14(1):128. doi:10.1186/s13000-019-0880-7.

52. Filho EHCN, Zancheta SB, de Barros Silva PG, Rodríguez Burbano RM, Rabenhorst SHB. Prognostic impact of miR-125b and miR-155b and their relationship with MYC and TP53 in diffuse large B-cell lymphoma: cell-of-origin classification matters. J Clin Exp Hematop. 2023;63:164–72. doi:10.3960/jslrt.23009.

53. Moreno JCA, Bahmad HF, Aljamal AA, Delgado R, Salami A, Guillot C, et al. Prognostic significance of p53 and p63 in diffuse large B-cell lymphoma: a single-institution experience. Curr Oncol. 2023;30:1314–31.

54. Schmitz R, Wright GW, Huang DW, Johnson CA, Phelan JD, Wang JQ, et al. Genetics and pathogenesis of diffuse large B-cell lymphoma. N Engl J Med. 2018;378:1396–407.

55. Cerhan JR, Slager SL. Familial predisposition and genetic risk factors for lymphoma. Blood. 2015;126:2265–73. doi:10.1182/blood-2015-04-537498.

56. Jing Y, Yuan Y, Yu Y, Zhao H, Yang H, Wang X. p53/PD-L1 co-expression predicts poor prognosis in diffuse large B-cell lymphoma. Discov Oncol. 2025;16(1):1270. doi:10.1007/s12672-025-03062-5.

57. Vehtari A, Gelman A, Simpson D, Carpenter B, Bürkner PC. Rank-normalization, folding, and localization: an improved R-hat for assessing convergence of MCMC. Bayesian Anal. 2021;16(2):667–718. doi:10.1214/20-BA1221.

58. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21(11):1539–58. doi:10.1002/sim.1186.

59. Röver C. Bayesian random-effects meta-analysis using the bayesmeta R package. J Stat Softw. 2020;93(6):1–51. doi:10.18637/jss.v093.i06.

60. Friede T, Röver C, Wandel S, Neuenschwander B. Meta-analysis of two studies in the presence of heterogeneity with applications in rare diseases. Biom J. 2017;59(4):658–71. doi:10.1002/bimj.201500236.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TP53LymphomaBMCMRMSupplementaryMaterials.docx