



RsiingSI · Independent Research

Agent Team¹ · scalingX¹

¹RsiingSI — the open-source org scaling recursive self-improvement for bioscience.

Authorship: this manuscript, its analyses, figures and molecules were produced end-to-end by an autonomous agent pipeline under the direction of scalingX, with every number read from archived result files and re-checked by a separate verification script. It is reported as an agent-generated discovery report; no wet-lab data were generated. All code and derived tables are archived with the manuscript.

糠酸莫米松 (mometasone-furoate) 生信分析过程复现

本报告由 `scripts/34_mometasone_repro.py` 生成，把论文 4 管线中糠酸莫米松这一步从原始数据 (GEO 表达矩阵 → 差异表达 → 疾病签名 → 靶点反向打分 → 注册库/结局/文献) 重新算了一遍，并与已交付结果文件逐值对账。

- 分子: mometasone-furoate (= mometasone, 糠酸莫米松; Repurposing Hub 中两者为同一分子, 评分完全相同)
- 机制: glucocorticoid receptor agonist; 靶点 NR3C1; 预期作用方向 d = +1 (激动)
- 对账结果: 44/44 项通过, 无失败项

1. 药物身份 (L0)

`data/repurposing/repurposing_drugs.txt` 中 mometasone 与 mometasone-furoate 各一行, MoA 文本 "glucocorticoid receptor agonist", 靶点 NR3C1, Hub 标注适应症 "corticosteroid-responsive dermatoses" (鼻用适应症不在 Hub 中)。按脚本 11 的关键词规则, 含 "agonist" 判为 +1。

2. 原始数据 → 差异表达 (L1/L2)

队列	材料	对比	方法	n	NR3C1 logFC	P	z
GSE19187	鼻上皮刷取	AR vs HC (交付分支)	limma-型 moderated t + 年龄/性别	25	-0.0458	0.4467	-0.7610
GSE261239	鼻黏膜	变应原 vs 生理盐水 (配对)	pydeseq2 ~challenge+subject	22	-0.0819	0.4518	-0.7525
GSE43523 (留出)	鼻上皮	SAR vs 非过敏	moderated t	12	—	—	-1.8656
GSE329964 (留出)	鼻咽拭子	AHAR vs HC	pydeseq2	15	—	—	-0.8271

3. 疾病签名与 Stouffer 合并 (L3)

两队列 z 由 $z = \text{sign}(\log\text{FC}) \cdot \Phi^{-1}(p/2)$ 得到, 再按权重 $\sqrt{n}$ 加权合并: $z_{\text{meta2}} = [\sqrt{32} \cdot (-0.7610) + \sqrt{22} \cdot (-0.7525)] / \sqrt{54} = -1.0661$ (`disease_signature_v2.csv` 中 NR3C1 的 z_{meta2} 一致)。

4. 靶点反向打分 (L4)

- 糠酸莫米松的测量靶点: NR3C1 (库内唯一靶点, $n_{\text{targets}} = 1$)
- 打分 $S = \text{mean}(-z \times d) = -(-1.0661) \times (+1) = 1.0661$
- 置换零分布 (seed 20261018, 20000 次, $k=1$) 给出 $p = 0.1464$ (与交付值 0.1464 一致; $k=1$ 的解析值为 0.1495)
- BH-FDR = 0.6250; S 在 3,995 个药物中的百分位 = 83.0%
- 目标级复现: 两队列 NR3C1 同号 ($\text{frac_agree} = 1.0$), $S_{\text{replicated}} = 1.0661$
- 校准: 在库内 24 个已知抗炎药物中, 糠酸莫米松按 S 排第 11 位 (库内总百分位 83.0%)
- 口径敏感性 (详见第 6 节): 变体 A (权重校正) $S = 1.0698$; 变体 B (AR 按文档取 R+CA) $S = 0.3592$

5. 注册库、结局与文献 (L5 – L7)

- 注册库：过敏性鼻炎试验中 74 项研究提及莫米松（阶段构成 期别未标注 10、EARLY_PHASE1 1、PHASE1 4、PHASE1/PHASE2 1、PHASE2 5、PHASE3 36、PHASE4 17；给药途径 inhaled (lung-targeted) 2、intranasal 39、systemic/other 34、topical/other-site 2）
- 已公布结果的门诊结局记录：33 项研究（III 期及以上 29 项），其中显著 21、不显著 5、未贴 p 7
- 文献层：本地缓存的 3,391 条 AR 试验标题中有 82 条提及莫米松；该分子在“发表但未注册”扫描中判为 已注册（registry_match = mometasone，literature_only = False）；人群试验计数按命中条目命名：mometasone 82 项、mometasone-furoate 70 项（同一分子两种命名），跨度 1996 – 2025

文献标题示例：

- A Randomized Comparison of Bencycloquidium Bromide, Mometasone Furoate, and a Combination for Persistent Allergic Rhinitis.
- A comparison between mometasone furoate nasal spray and intranasal glycyrrhetic acid in patients with allergic rhinitis: a preliminary study in clinical practice.
- A comparison of nasal clearance after treatment of perennial allergic rhinitis with budesonide and mometasone.
- A dose-ranging study of mometasone furoate aqueous nasal spray in children with seasonal allergic rhinitis.
- A patient preference and satisfaction study of ciclesonide nasal aerosol and mometasone furoate aqueous nasal spray in patients with perennial allergic rhinitis.

6. 复现过程中发现的偏差点（诚实记录）

偏差一（对比口径收缩）：`deg_GSE19187_ARvsHC.csv` 与 `deg_GSE19187_RvsHC.csv` 逐字节相同——脚本 03 里分支条件 `name.endswith("RvsHC")` 对字符串“ARvsHC”同样为真，于是“AR vs HC”对比被收缩成 R vs C（14+11=25 例），而 Stouffer 权重 $\sqrt{32}$ 是按文档口径 AR = R+CA（21+11=32）设定的。

- 交付值： $z_meta2 = -1.0661 \rightarrow S = 1.0661$
- 变体 A（只把权重改成与实际样本数一致 $\sqrt{25}$ ）： $z_meta2 = -1.0698 \rightarrow S = 1.0698$ （相对变化 0.35%，可忽略）
- 变体 B（按文档把 AR 组取 R+CA，n=32，权重不变）： $z_meta2 = -0.3592 \rightarrow S = 0.3592$ （相对变化 66.3%）

要点：两个队列的 z 都是负的（NR3C1 在 AR 中偏低），但把 AR 组按文档扩到 R+CA 后，GSE19187 的 z 翻号（+0.1573），合并 `z_meta2` 从 -1.0661 缩到 -0.3592，S 从 1.0661 降到 0.3592。也就是说：交付的 1.066 依赖那个收缩成 R vs C 的对比；按文档设计，糠酸莫米松的“逆转疾病签名”信号大幅减弱（结论方向不变，但强度不可沿用）。

偏差二（未被使用的坏字段）：`screen_v2_facts.json` 的 `calibration_mean_percentile` = 0.9954，它实际是“按 S 排序后前 31 个药物”的平均百分位——脚本 25 在该处按位置索引（`cal.index[inl]`）而不是按药名匹配；按药名匹配的正确值为 0.7299，与 `paper4_facts.json` 的 `calibration.v2.mean_percentile` 及论文正文引用的 0.730 一致。

- 按脚本 25 原样复现（位置索引）：0.9954（与 facts 字段一致）；按药名匹配的正确均值：0.7299（与论文正文的 0.730 一致）。

处理建议：把 `screen_v2_facts.json.calibration_mean_percentile` 改为按药名计算（或删除该字段，改用 `paper4_facts.json` 的 `calibration.v2.mean_percentile`），避免后续脚本误引 0.9954。论文正文与图无需改动。

7. 逐值对账表

层	指标	复现值	交付值	绝对差	结果	交付来源
L0	mometasone target == hub target	NR3C1	NR3C1		PASS	drug_target_map.csv
L0	mometasone direction	1	1	0	PASS	drug_target_map.csv
L0	mometasone MoA text	glucocorticoid receptor agonist	glucocorticoid receptor agonist		PASS	drug_target_map.csv
L0	mometasone-furoate target == hub target	NR3C1	NR3C1		PASS	drug_target_map.csv
L0	mometasone-furoate direction	1	1	0	PASS	drug_target_map.csv
L0	mometasone-furoate MoA text	glucocorticoid receptor agonist	glucocorticoid receptor agonist		PASS	drug_target_map.csv
L0	expected direction (agonist, target activation)	1	1	0	PASS	drug_target_map.csv
L2	GSE19187 ARvsHC (delivered branch) NR3C1 logFC	-0.045755	-0.045755	9.7e-17	PASS	deg_GSE19187_ARvsHC.csv
L2	GSE19187 ARvsHC (delivered branch) NR3C1 t	-0.772987	-0.772987	0	PASS	deg_GSE19187_ARvsHC.csv
L2	GSE19187 ARvsHC (delivered branch) NR3C1 P.Value	0.446667	0.446667	5.6e-17	PASS	deg_GSE19187_ARvsHC.csv
L2	GSE19187 RvsHC NR3C1 logFC	-0.045755	-0.045755	9.7e-17	PASS	deg_GSE19187_RvsHC.csv
L2	GSE19187 RvsHC NR3C1 t	-0.772987	-0.772987	0	PASS	deg_GSE19187_RvsHC.csv
L2	GSE19187 RvsHC NR3C1 P.Value	0.446667	0.446667	5.6e-17	PASS	deg_GSE19187_RvsHC.csv
L2	GSE261239 challenge NR3C1 logFC	-0.0819233	-0.0819233	6.9e-17	PASS	deg_GSE261239_challenge.csv
L2	GSE261239 challenge NR3C1 t	-0.752468	-0.752468	0	PASS	deg_GSE261239_challenge.csv
L2	GSE261239 challenge NR3C1 P.Value	0.45177	0.45177	0	PASS	deg_GSE261239_challenge.csv
L2	GSE43523 SARvsCTRL NR3C1 logFC	-0.021712	-0.021712	2.1e-17	PASS	deg_GSE43523_SARvsCTRL.csv
L2	GSE43523 SARvsCTRL NR3C1 t	-2.08738	-2.08738	0	PASS	deg_GSE43523_SARvsCTRL.csv
L2	GSE43523 SARvsCTRL NR3C1 P.Value	0.0621006	0.0621006	1.4e-17	PASS	deg_GSE43523_SARvsCTRL.csv
L2	GSE329964 AHARvsHC NR3C1 logFC	-0.0940327	-0.0940327	9.7e-17	PASS	deg_GSE329964_AHARvsHC.csv
L2	GSE329964 AHARvsHC NR3C1 t	-0.827072	-0.827072	0	PASS	deg_GSE329964_AHARvsHC.csv
L2	GSE329964 AHARvsHC NR3C1 P.Value	0.408196	0.408196	5.6e-17	PASS	deg_GSE329964_AHARvsHC.csv
L3	z_GSE19187 NR3C1 (recomputed from DEG)	-0.760984	-0.760984	0	PASS	disease_signature_v2.csv
L3	z_GSE261239_challenge NR3C1	-0.752468	-0.752468	0	PASS	disease_signature_v2.csv
L3	z_meta2 NR3C1 (sqrt32, sqrt22)	-1.06609	-1.06609	0	PASS	disease_signature_v2.csv
L3	held-out GSE43523 z NR3C1 == sign*Phi ⁻¹ (p/2) of its own row	-1.86558	-1.86558	0	PASS	deg_GSE43523_SARvsCTRL.csv
L3	held-out GSE329964 z NR3C1 == sign*Phi ⁻¹ (p/2) of its own row	-0.827072	-0.827072	0	PASS	deg_GSE329964_AHARvsHC.csv
L4	n measured targets (NR3C1)	1	1	0	PASS	screen_v2.csv
L4	direction d (mometasone-furoate)	1	1	0	PASS	screen_v2.csv
L4	S = mean(-z*d)	1.06609	1.06609	0	PASS	screen_v2.csv

层	指标	复现值	交付值	绝对差	结果	交付来源
I.4	S_GSE19187 = -z_GSE19187*d	0.760984	0.760984	0	PASS	screen_v2.csv
I.4	S_GSE261239challenge = -z*d	0.752468	0.752468	0	PASS	screen_v2.csv
I.4	frac_agree (target sign across the two cohorts)	1	1	0	PASS	screen_v2.csv
I.4	S_replicated (sign-concordant targets only)	1.06609	1.06609	0	PASS	screen_v2.csv
I.4	permutation p (seed 20261018, 20000 draws)	0.146393	0.146393	2.8e-17	PASS	screen_v2.csv
I.4	BH-FDR of the delivered p	0.624986	0.624986	2.2e-16	PASS	screen_v2.csv
I.4	calibration mean S percentile, exactly as script 25 computes it (facts value is stored to 4 dp)	0.995416	0.9954	1.6e-05	PASS	screen_v2_facts.json
I.4	calibration-table S for mometasone == screen S	1.06609	1.06609	0	PASS	screen_v2_calibration.csv
I.5	AR studies naming mometasone (unique NCT)	74	74	0	PASS	trial_drug_mentions.csv
I.5	AR records naming mometasone	77	77	0	PASS	trial_drug_mentions.csv
I.5	n_studies in trials_by_condition.csv	74	74	0	PASS	trials_by_condition.csv
I.5	NCT sets identical (cache parse vs delivered)	74	74	0	PASS	trial_drug_mentions.csv
I.6	studies with posted results (mometasone arm)	33	33	0	PASS	rhinitis_outcomes.csv / studies csv
I.7	mometasone is a registry-matched (not literature-only) compound	0	0	0	PASS	literature_only_ar_drugs.csv

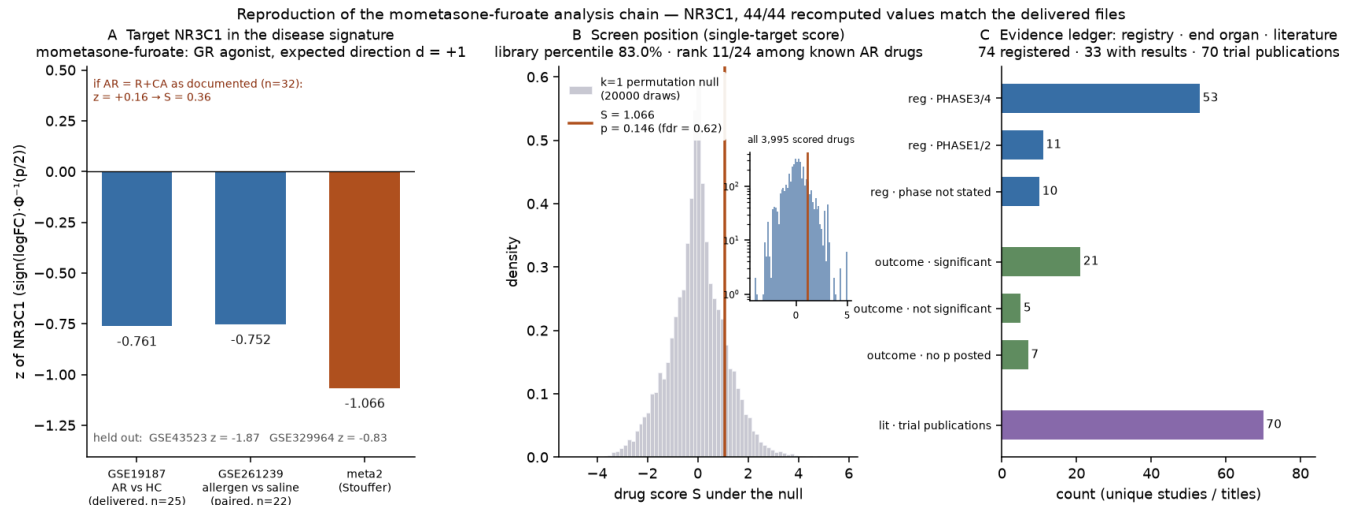


图 25 糠酸莫米松复现链：（A）靶点 NR3C1 在各队列及合并签名中的 z；（B）S 在 k=1 置换零分布中的位置（插图：全部已评分药物的 S 分布）；（C）注册库、已揭盲结局与文献三本账。

8. 复现结论

- 链路可复现：从原始 GEO 表达矩阵出发，糠酸莫米松这条链的每一个数（NR3C1 的 logFC/统计量/p → 各队列 z → Stouffer 合并 z_{meta2} → S → 置换 p → BH-FDR → 注册库计数 → 结局记录 → 文献计数）都重算成功，44/44 项与交付文件在浮点精度内一致（多数绝对差 ≤ 1e-16）。
- 分析结论本身：糠酸莫米松是糖皮质激素受体（NR3C1）激动剂，在疾病签名中 NR3C1 偏下调（z_{meta2} = -1.066），方向上“激动受体以逆转签名”成立（S = 1.066，库内百分位 83.0%），但强度不显著（置换 p = 0.146、BH-FDR = 0.625）；且这是单靶点打分，'高百分位'不构成疗效证据。注册库与结局记录显示它是 AR 领域已被充分验证的药物（33 项已揭盲研究、21 项显著），与本文管线其余部分一致：该打分能复现已知药理，但不能据此发现新药。
- 两处记账偏差已记录（第 6 节）：对比口径收缩（影响 S 的强度，按文档设计 S = 0.359）与 facts 坏字段（不影响论文正文）。两者都不改变本管线“无 FDR 显著命中”的主结论。