

Source-Causal Control of Historical Context in Longitudinal Radiology Report Generation

Dmitry Lvov^{1,2} Ilya Pershin^{1,3}

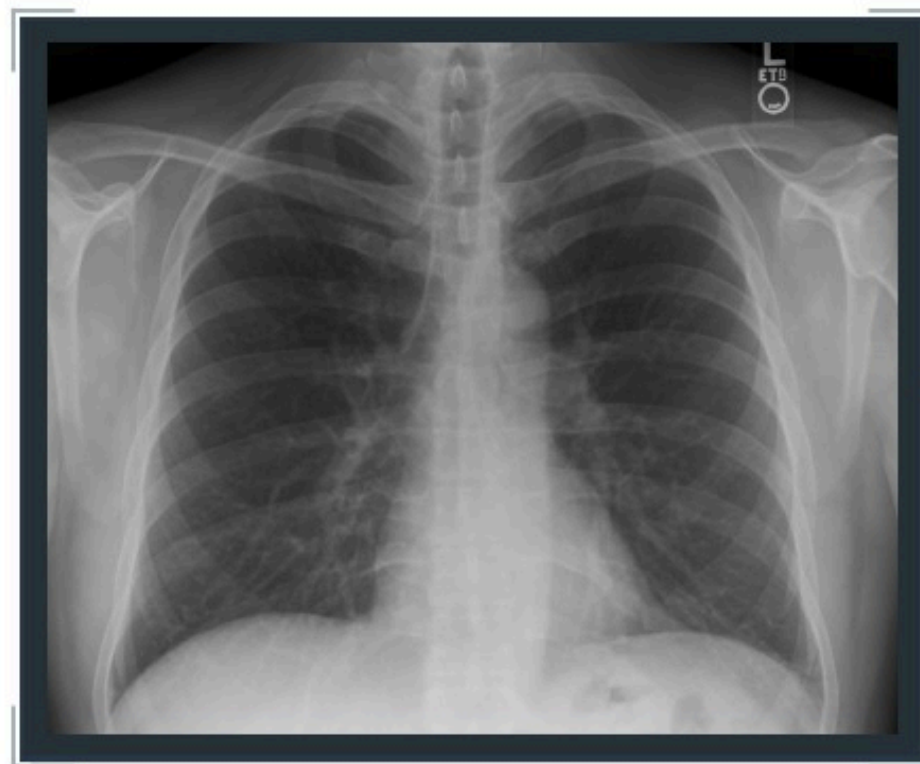
¹ Innopolis University ² ITMO University ³ Kazan Federal University

NeurIPS 2026

A plausible past can change today's report



Current chest X-ray



Same
generator

Current image only

No pleural effusion.

Prior report

Small pleural effusion.

+ Prior report

Small pleural effusion
persists.

Source interventions with the current X-ray fixed



18,716 study pairs

93,580 generated reports

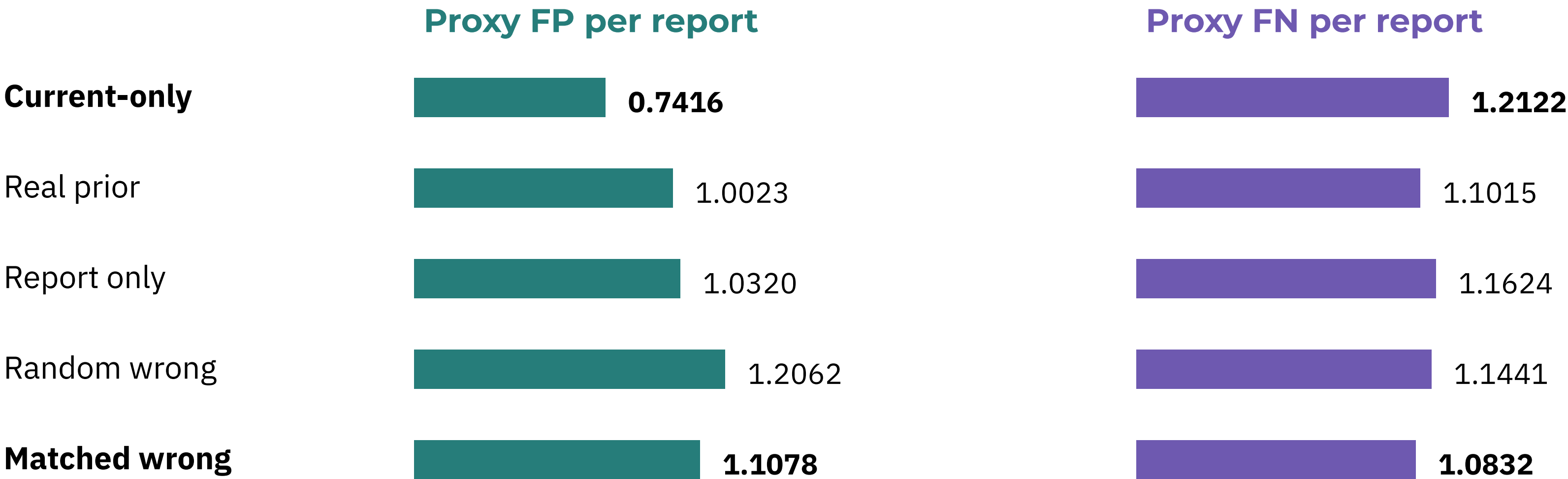
Historical source	Current X-ray	Historical input
Current-only	Fixed	None
Real prior	Fixed	Patient-aligned prior images and report
Report only	Fixed	Patient-aligned prior report
Random wrong	Fixed	Another subject’s history
Matched wrong	Fixed	Another subject, matched findings and report length

Paired comparisons isolate the effect of the historical source

Wrong history adds positives while increasing coverage



49.4% more proxy FP with matched wrong history



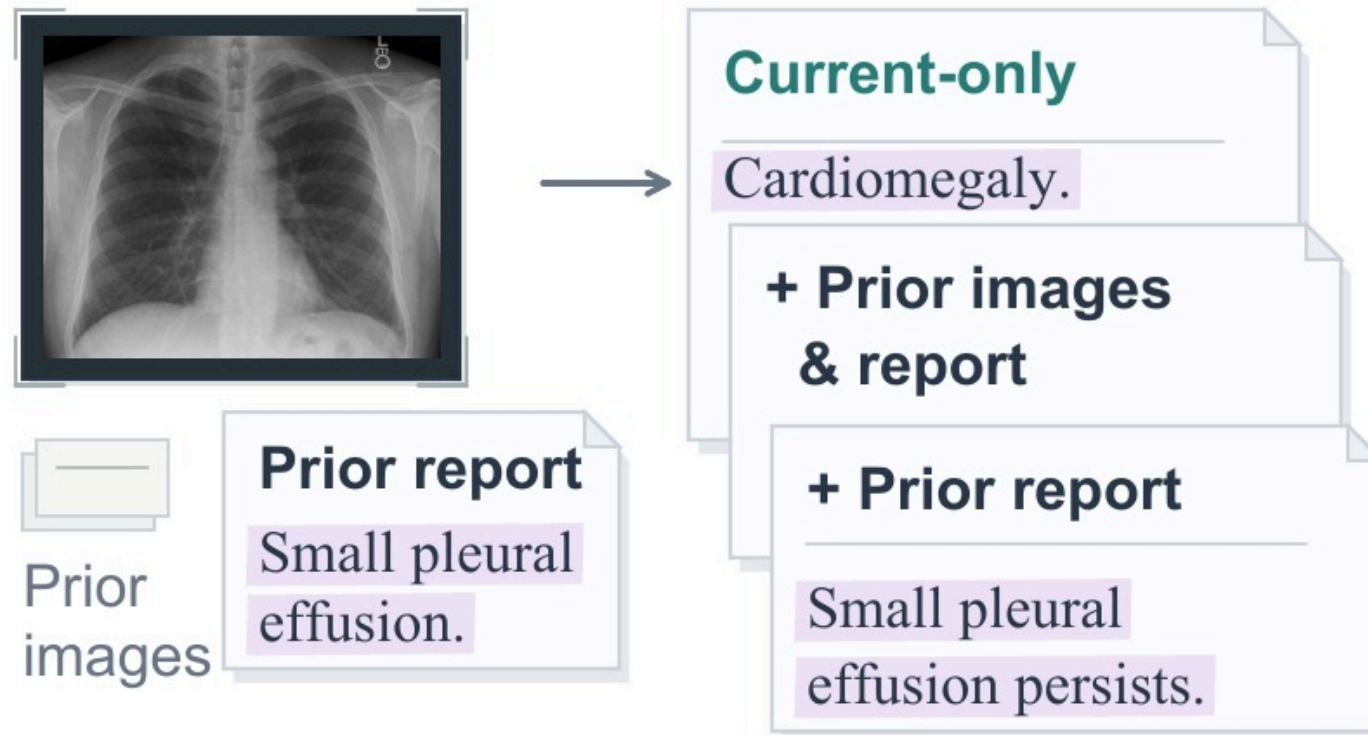
Matched wrong history: +0.3662 proxy FP and -0.1290 proxy FN per report

SCFR selects reports around a current-only anchor



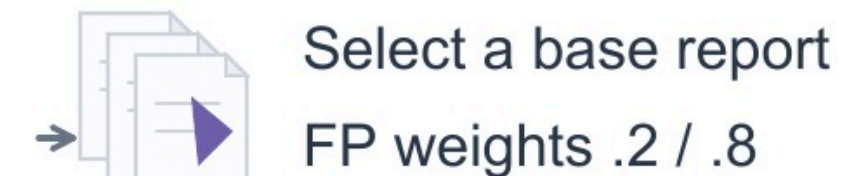
1 Generate three reports

Current CXR in every call



2 Construct proposals

Two fitted rankers



Two fitted finding verifiers

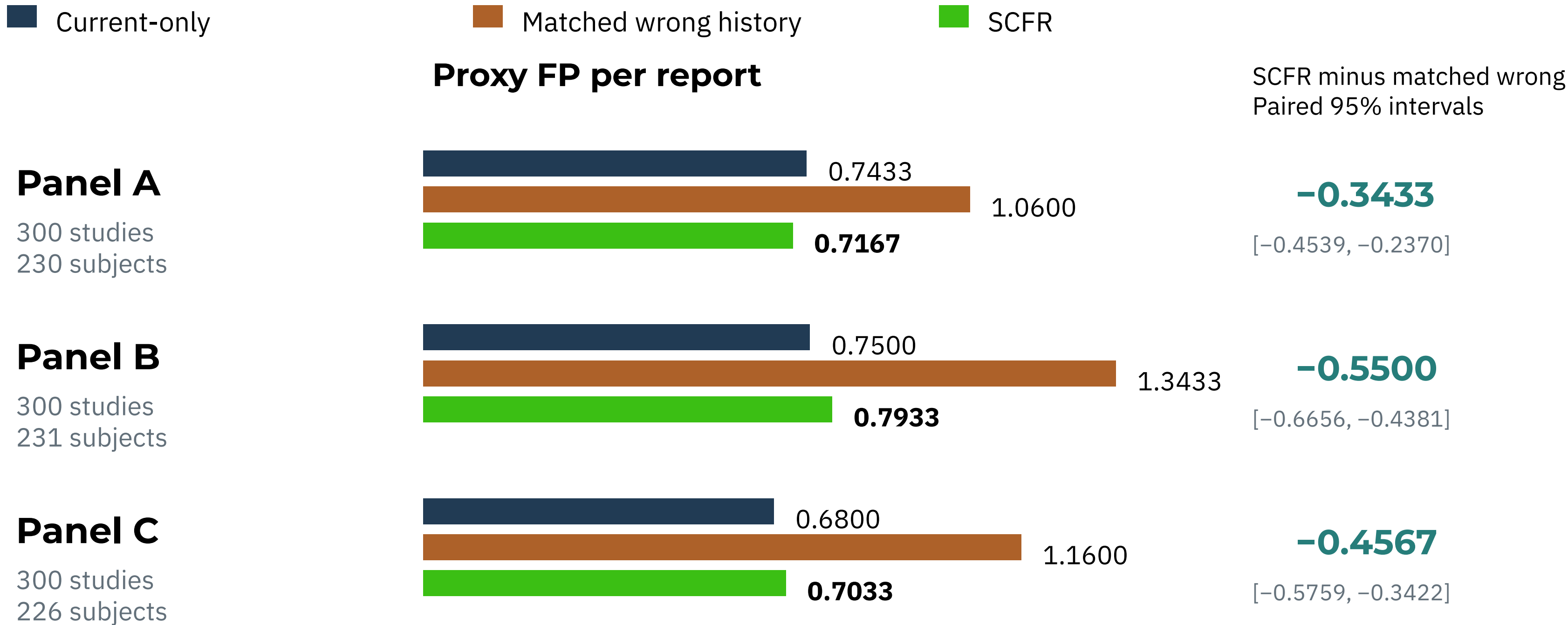
Filter + complete current-only



3 Select among five reports



SCFR removes most of the wrong-history FP excess



Holm-adjusted $p \leq 0.0012$ in all panels. Subject-cluster bootstrap.

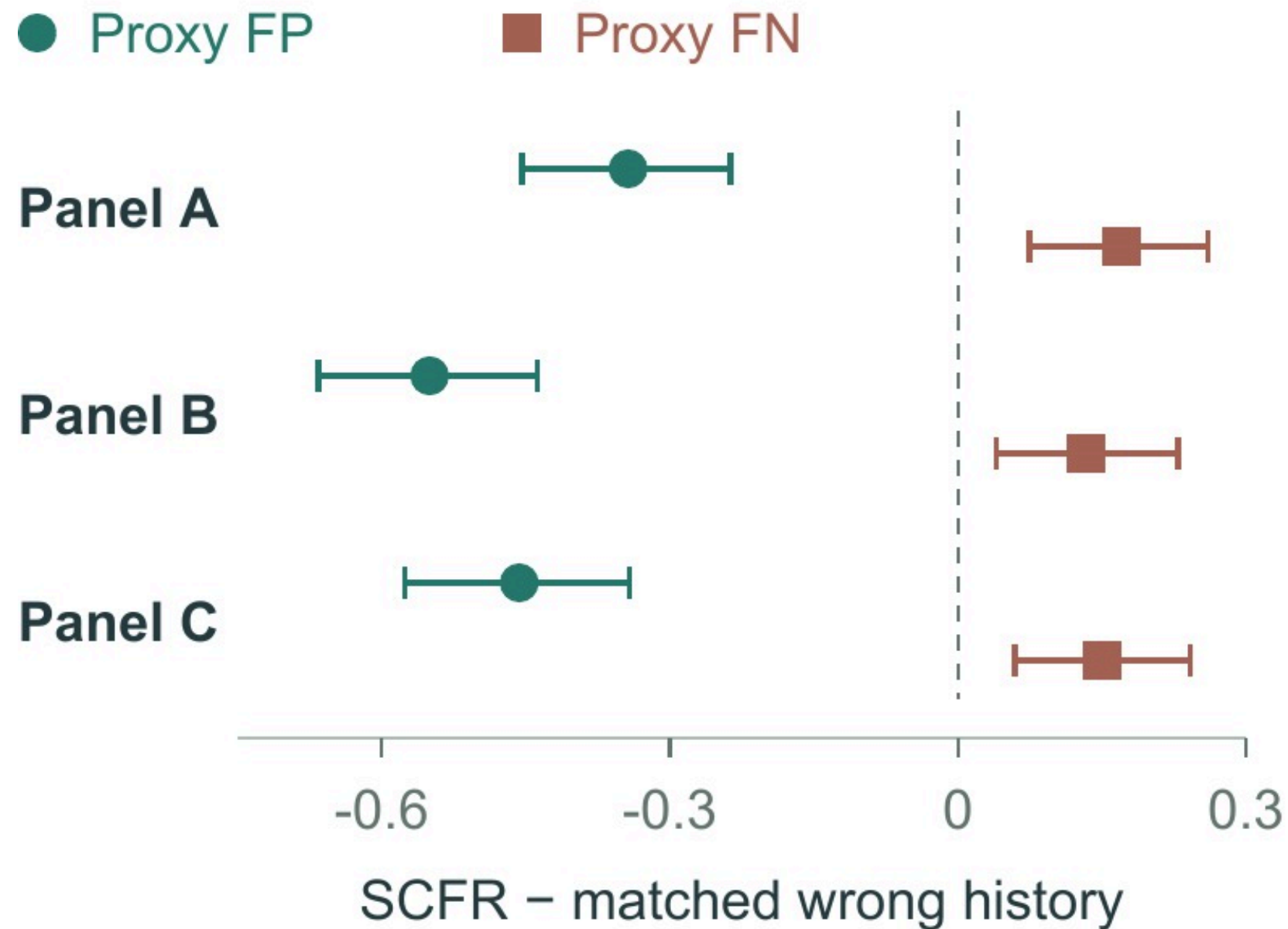
A evaluates the frozen selection. B/C reuse development cohorts.

Selective history use near current-only



139 / 900 evaluations select text distinct from current-only

(a) vs matched wrong history



(b) vs current-only

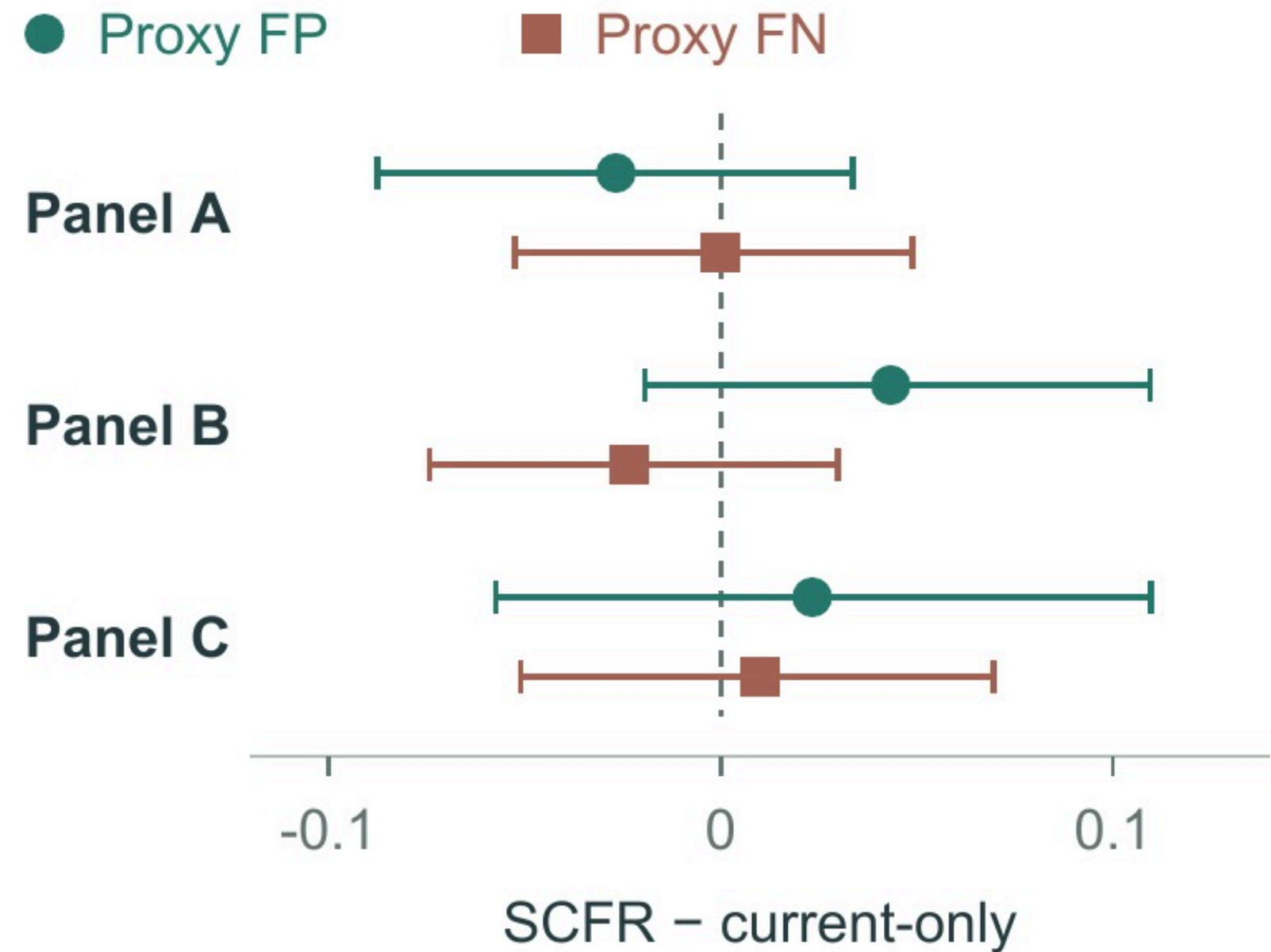


Image review calibrates the text proxy



60 challenge-stratified cases. Matched wrong vs real history: +33.8 pp (95% CI 5.6 to 63.6)

(a) Current-image support

60 cases · one radiologist

Unsupported share of resolved additions

Real history

7/18 resolved

Matched wrong

8/11 resolved

Random wrong

8/13 resolved



(b) Active-route adjudication

30 paired cases · one radiologist

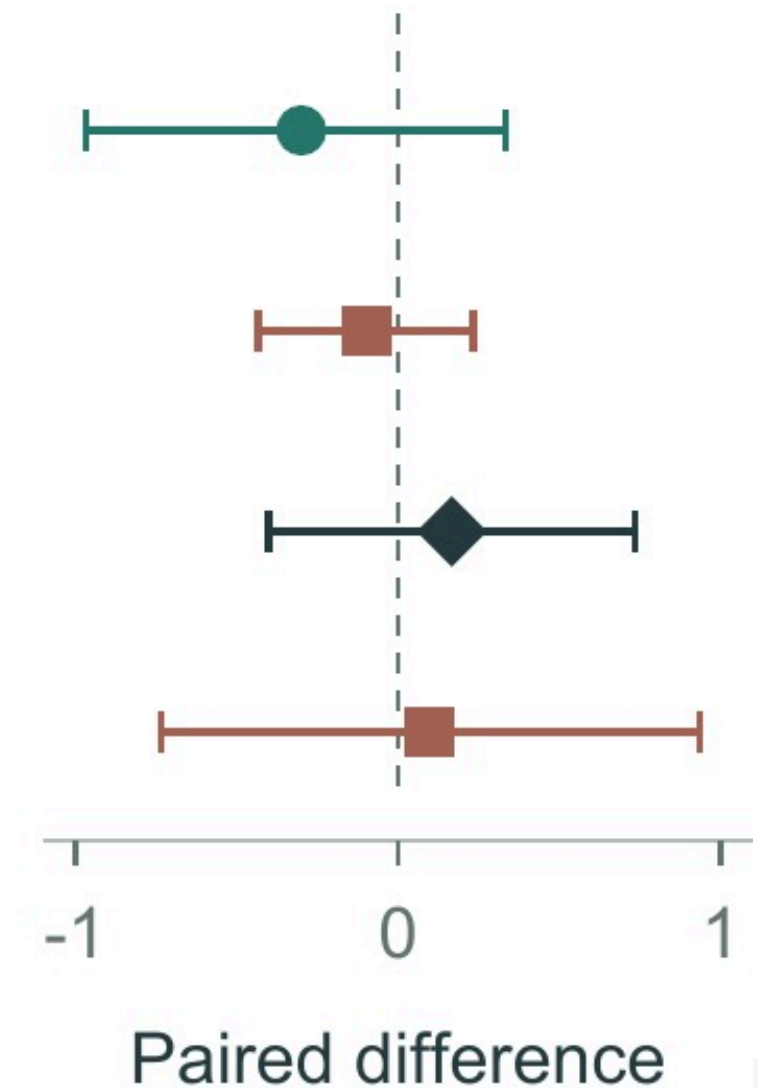
SCFR – current-only; items per case

Unsupported additions

Important omissions

Supported represented

Supported lost



Longitudinal content on selected history routes



30 cases where SCFR changes the report

Adjudicated items	Current-only	SCFR	Difference
Supported findings represented	151	156	+5
Supported findings lost	47	50	+3
Unsupported additions	26	17	−9
Important omissions	20	17	−3

4 correct temporal relations and 1 partially correct

True history informative in 25/30 cases.

Paired intervals include zero. Overall preference: 13 SCFR, 17 current-only (p = 0.585).

Source auditing and selective history use

Controlled source interventions expose historical influence on generated reports

SCFR reduces wrong-history proxy FP and selectively admits history-conditioned reports

Prospective multi-reader evaluation should assess clinical benefit

d.lvov@innopolis.ru github.com/LvovDmitro/SCFR

Retrospective MIMIC-CXR evidence. Direct-current effects are mixed.
One-reader image review.



Paper



Code



Google Scholar

Backup: filtering positive additions



Lower FP can come with lost coverage

Control	Finding F1	Proxy FP/report	Proxy FN/report
Current-only	0.4338	0.7416	1.2122
No positive additions	0.4267	0.6146	1.2922
Block prior-only additions	0.4109	0.7931	1.2541

Removing every addition lowers proxy FP and also removes useful coverage

Backup: generator-specific responses to history



Source sensitivity: matched wrong history minus current-only

Generator	Audit studies	WH Δ FP/report	WH Δ FN/report
MAIRA-2	18,716	+0.3662	−0.1290
RadInfer	400	−0.0950	−0.2175
Libra	306	−0.1830	+0.3824

Five-fold routing: SCFR minus current-only

Generator	Active / fallback folds	SCFR Δ F1	Δ FP/report	Δ FN/report
RadInfer	5 / 0	+0.1060	−0.2925	−0.1200
Libra	2 / 3	−0.0184	−0.0719	+0.0752