



Step 1 - Phase I as uploaded
n = 851 prostate core biopsy whole-slide images from 60 patients

- 22 cores per canonical exclude-list

Excluded: 19 duplicate rescans (mostly case c17; also c45, c48, c5, c56, c59, c60)

Excluded: 2 accidentally uploaded IHC stain slides (c11_s13, c35_s4)

Excluded: 1 slide where the Paige website did not run (c17_s18)

Step 2 - Phase I analytical cohort
n = 829 cores from 60 patients (canonical Phase I denominator; per paige-prostate-exclude-list.xlsx)

- 19 cores at the data-cleaning stage (but +6 duplicate-rescan cores that slipped through)

Step 3 - Phase II RDS
n = 810 cores re-read by all 4 pathologists with and without AI assistance (used for AI-effect / IHC / inter-rater analyses)

- 672 cores without parseable Gleason from every interpreter

Step 4 - Inter-rater complete-cases
n = 138 / 810 cores (parseable Gleason from every interpreter; AI - Report - Reference - P1-P4 x no-AI / with-AI) used for Fleiss / Light kappa and AI-effect-on-pathologist Grade Group analyses

Paige Prostate categorization (suspicious vs not suspicious)

Excluded: 17 fully-read but dropped from Phase II RDS (data-cleaning)

Excluded: 4 partial Phase II reads (1-6 of 8 reader cells filled)

Excluded: 1 with 7-of-8 cells, not in RDS (c17_s20)

Concordant with original pathology report
n = 800 / 829 (96.5%; Cohen κ = 0.909)

Discordant with original pathology report
n = 29 / 829 (3.5%)

AI = benign - Report = benign
n = 601 (72.5%)

AI = malignant - Report = malignant
n = 199 (24.0%)

AI suspicious for adenocarcinoma - Report benign
n = 28 / 829 (3.4%)

AI benign - Report malignant
n = 1 / 829 (0.1%)

After expert re-evaluation with IHC (reference diagnosis on the 28 cores):

Adenocarcinoma confirmed by reference review
n = 1 (0.1%)

Confirmed benign
n = 18 (2.2%)

ASAP
n = 1 (0.1%)

Adenocarcinoma confirmed
n = 9 (1.1%)
minute foci, low-grade Gleason patterns; additional tumor present in other cores of the same case