

The Algorithm Is Not the Supervisor

Five useful warnings about AI, decision support, and the difference between a convincing output and a controlled clinical process.

Wednesday, September 16,
2026

LIS/informatics | AI governance | Decision support | Patient safety

FIVE SIGNALS WORTH YOUR BENCH TIME

01 / AI / SOFTWARE SAFETY

Code that runs is not code that is safe

Finding. A laboratory-medicine commentary examines 'vibe coding': using generative AI to produce clinical software without enough engineering expertise to recognize hidden defects. The central warning is that polished, syntactically correct output can create an illusion of competence.

Why it matters. Any code influencing routing, calculations, flags, autoverification, or result release is part of the quality system. It needs requirements, version control, test cases, peer review, change control, rollback, and monitored deployment.

Limits. Opinion paper rather than outcome research. Its risk argument is persuasive but does not quantify failure rates across tools or laboratory settings.

Evidence label: Peer-reviewed commentary. No primary clinical cohort.

02 / AI / GOVERNANCE

Governance has to arrive before the model does

Finding. A 2026 review organizes the opportunities, pitfalls, ethics, and clinical-governance obligations surrounding AI in laboratory medicine, emphasizing that technical performance is only one part of safe use.

Why it matters. Ownership, intended use, bias review, privacy, human oversight, downtime behavior, auditability, and post-deployment monitoring must be assigned before a model can influence patient work.

Limits. Review-level synthesis; it does not validate a single governance framework or provide comparative implementation outcomes. Full-text access and disclosure details should be checked.

Evidence label: Peer-reviewed review. Publisher DOI verified; full text may require subscription.

03 / PRECISION ONCOLOGY / EQA

Decision-support software may need ring testing too

Finding. An international pilot used 20 synthetic genomic datasets - 10 tumor-normal and 10 tumor-only - to compare precision-oncology decision-support tools and the manual interpretation layered around them.

Why it matters. External quality assessment should not stop at the variant call. Knowledge bases, software versions, filters, evidence grading, and human interpretation can all change the final recommendation.

Limits. Pilot scope, synthetic cases, and a targeted-panel context limit conclusions about live cases and broader platforms. Participation and tool-selection effects may understate real-world variation.

Evidence label: Peer-reviewed international pilot study. Publisher DOI verified.

Explainable ML can support a thrombocytopenia differential - after validation

Finding. A retrospective study of 387 children with newly diagnosed thrombocytopenia compared six algorithms for distinguishing primary immune thrombocytopenia from connective-tissue-disease-related thrombocytopenia and used SHAP values to examine feature contributions.

Why it matters. Routine data may help prioritize confirmatory work or follow-up, particularly when early phenotypes overlap. Explainability is useful for review, but it is not a substitute for clinical validity.

Limits. Single-center, retrospective data; a small CTD-related subgroup; internal train-test splitting; and no prospective impact analysis. The model should not be used as a diagnostic endpoint.

Evidence label: Peer-reviewed diagnostic modeling study. Full text may require subscription.

Routine chemistry can triage myeloma suspicion, not close the case

Finding. Researchers used 15 routine demographic and laboratory variables from 422 consecutive hematology referrals to develop models separating multiple myeloma from clinical mimickers, with a held-out test set, calibration measures, and decision-curve analysis.

Why it matters. A transparent triage tool could help direct scarce specialized testing or expedite high-risk referrals. Its safe role is prioritization - never replacement of electrophoresis, immunofixation, free-light-chain, marrow, imaging, and clinical assessment.

Limits. Single-center retrospective cohort with internal validation only. Referral prevalence affects predictive values, and model drift is likely across sites and analyzers.

Evidence label: Peer-reviewed primary research; open-access publisher page verified.

Verified source trail

1. De Bruyne S, Rommes S, Fiers T, et al.. Vibe coding and the illusion of competence: why laboratory medicine specialists shouldn't trust their AI-generated code for clinical deployment. Clinical Chemistry and Laboratory Medicine (CCLM). 2026. DOI: 10.1515/cclm-2026-0887.
2. Lippi G, Mattiuzzi C, Henry BM. Opportunities, pitfalls, ethics, and clinical governance of AI in laboratory medicine. LabMed Discovery. 2026. DOI: 10.1016/j.lmd.2026.100145.
3. Nygaard V, Zhao S, Tamborero D, et al.. Assessing clinical decision support system tools in precision oncology: piloting ring testing. ESMO Real World Data and Digital Oncology. 2026. DOI: 10.1016/j.esmorw.2026.100731.
4. Kang F, Yan M, Yue Y, et al.. Development of a Machine Learning Algorithm for Differential Diagnosis Between Primary Immune Thrombocytopenia and Connective Tissue Disease-Related Thrombocytopenia in Pediatric Patients. Journal of Clinical Laboratory Analysis. 2026. DOI: 10.1002/jcla.70337.
5. Erdoğdu B, Gül D, Aylı Blr, et al.. Routine Laboratory-Based Machine Learning for Discriminating Multiple Myeloma from Clinical Mimickers: Development and Internal Validation of a Diagnostic Prediction Model. Diagnostics. 2026. DOI: 10.3390/diagnostics16182928.

VERIFICATION	DISCLOSURE & USE
Each DOI was checked against current Crossref metadata and its DOI resolution path. Source titles and publication year were matched before inclusion. Previously published Editions 001-013 were screened for duplicate DOI, canonical source, title, and substantially similar story.	Funding and conflict statements are source-specific and may sit behind full-text access. Readers should review the publisher disclosure before policy adoption. Named instruments, reagents, software, and manufacturers are reported because they define method scope; inclusion is not endorsement.

Editorial scope: professional education and quality-improvement intelligence. This newsletter does not replace local validation, accreditation requirements, manufacturer instructions, institutional policy, or patient-specific clinical judgment.