Pap test slide limits appear intended to improve screening sensitivity by allowing cytotechnologists sufficient time to examine slides accurately. Almost any limit lower than the current 100 manual and 200/170 imaged slides in no less than 8 hours would be an improvement, but adjusting numbers alone is insufficient. Decision-makers should be aware of the following:

1. Quality means useful for its intended purpose. Once the intended purpose has been defined, contributing factors can be identified and measured. Anything that does not add value to a product or service from the standpoint of the external customer is waste.

2. All cytotechnologists screen Pap test slides differently.¹

3. There is no standardized approach to screening slides in print, is taught, or is practiced.²

4. Even if there were a screening SOP, errors would occur. Humans are human.

5. False negative Pap test results can be fatal. There are no definitive data re numbers annually. Whatever numbers exist are likely understated. Clinicians tend not to tell women who develop cervical cancer that their Pap tests may have been misread.

6. Screening a Pap test is a process by which a slide is moved successively in small steps to promote the likelihood that at least 1 recognizable abnormal cell will fall within the narrow field-of-vision of an alert professional observer long enough to be perceived, recognized, and interpreted.³

7. Screening time is interpreted differently by different labs, often in patient adverse ways—an hour is not an hour.⁴ “Instead of screening workload, the term ‘analytic productivity workload’ was used to refer to the analytic component of the test process. This consists of review of the clinical information, evaluation of the case, and recording of the results. This definition specifically excludes pre- and post-analytic processes such as specimen preparation, resolving clinical history discrepancies, QA review, and batch data entry into a laboratory information system.”³

8. Another way to determine what counts as screening time and what doesn’t is to refer to CPT coding language, for example: CPT 88142: Cytopathology, cervical or vaginal (any reporting system), collected in preservative fluid, automated thinlayer preparation; manual screening under physician supervision. In other words, if it can’t be billed, it’s not screening.

9. Not everyone understands the difference between QC and QA, or that QA now means quality assessment in the lexicon of CLIA ’88.

10. Error detection relies on the same process that caused the error in the first place: (re)screening by humans. Underfinding errors is the rule, not the exception.

11. The relationship between screening and rescreening can be characterized as a self-reinforcing negative feedback loop. The worse the rescreening, the better the screening appears to be.
12. To achieve perfection in screening, don’t detect errors. If rescreening identifies few errors, is it because the primary screening was that good, or the rescreening was that bad? Labs prefer the former possibility, I prefer the latter. There is little incentive to rescreen well.

13. True negatives will always be negative, regardless of how poorly a slide was screened, or whether it was screened at all. This realization introduces the concept of the false true negative: You got the right answer not because you’re good, but because you can’t be wrong.

14. In laboratories with a culture of high output, the false true negative can be self-delusional. For example, one laboratory reported finding zero screening errors by 33 of 65 cytotechnologists among 200,000 Pap tests rescreened over a 3-month period.

15. Publishing one’s false negative proportion (FNP) without also publishing the accompanying sensitivity is an exercise in futility. Ideally, FNP + sensitivity = 1. Therefore, FNP 0.05 implies a screening sensitivity of 0.95. If a lab’s screening sensitivity were 0.95, why rescreen?

16. Rescreening sensitivity is likely to be lower than that for primary screening. Requiring “qualified” individuals to review slides doesn’t guarantee quality. They may be “false positives.” When a lab thinks it knows its screening sensitivity, that value must be divided into the rescreening findings to account for the fact that rescreening sensitivity is not 100%.

17. It is doubtful that labs know how to calculate their screening sensitivity.

18. FNP 0.05 can be achieved when: 1) screening sensitivity is 90% and rescreening sensitivity is 50%, 2) screening sensitivity is 80% and rescreening sensitivity is 25%, and 3) screening sensitivity is 70% and rescreening sensitivity is 12.5%. These combinations are calculated using simple arithmetic. Using real data without understanding its limitations is misleading.

19. The FNP so-called floor of performance (0.05) is potentially dangerously misleading. It implies a screening sensitivity of 0.95. If that were true, all false negatives would be identified if, and only if, all initially NILM Pap tests were rescreened with 100% sensitivity.

20. Properly calculated FNP can be used to estimate the number of unidentified ASC-US, LSIL, and HSIL cases screened once and categorized initially as NILM.

21. Labs don’t know how to compare one cytotechnologist’s rescreening findings against those of the entire laboratory. As a result, cytotechnologists can be disciplined unfairly. Use chi square.

22. Labs don’t know how to simply determine whether cytotechnologist productivity is within CLIA limits.

23. Laboratories don’t know how to assess, or re-assess, each individual’s workload limit and adjust when necessary. They may do it, but it’s not reliable. Corrective actions fail because they do not address the root cause(s).

24. Errors can be acceptable or unacceptable. The former can occur in any laboratory (eg, few abnormal cells, well-differentiated cells that do not stand out). Unacceptable errors are blunders: well-preserved obviously abnormal cells everywhere.

25. Few cytoprofessionals know what vigilance (ie, sustained attention) means, what’s required to maintain it (eg, frequent breaks), and vigilance decrement’s role in screening errors. It’s universal and insidious.

26. High screening productivity means uninterrupted screening for long periods, which increases errors.

27. Rapid anything (eg, prescreening, review) doesn’t improve screening quality measurably, reproducibly, and sustainably. There’s no free lunch!
28. Proficiency testing is an artificial exercise unrelated to real world screening circumstances and errors, which are random.

29. CLIA expects better Interobserver agreement between cytotechnologists and pathologists on cell than the literature has demonstrated is possible among pathologists on tissue.

30. Discrepant true positives don’t kill women; unidentified false negatives do. CLIA’s emphasis in cytology is misplaced.  

31. Education is key to understanding the complexity of the screening process.  

32. If you recommend lowering the workload limits, don’t make the limit absolute. Future technologies, as yet unforeseen, may permit higher limits with acceptable sensitivity and specificity.

References


