Potential for Economic Inefficiencies in Human-AI Hybrid Screening: An Analysis of the Sybil AI on NLST Data

Frank C. Schuller[1,2] & David Guo[1]

[1] MIT Jameel Clinic for Machine Learning in Health, Massachusetts Institute of Technology, Cambridge, MA, USA

[2] Initiative for Medical AI, The Hastings Center for Bioethics, Garrison, NY, USA

Correspondence: Frank C. Schuller, MIT Jameel Clinic for Machine Learning in Health, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA. Tel: 1-617-610-1361. E-mail: frank515@mit.edu

Received: June 6, 2026 Accepted: June 30, 2026 Online Published: July 5, 2026

doi:10.5430/ijba.v17n2p34 URL: https://doi.org/10.5430/ijba.v17n2p34

Abstract

Background: Lung cancer screening using low-dose chest CT (LDCT) can be enhanced by artificial intelligence algorithms such as Sybil. While autonomous AI models demonstrate superior predictive accuracy and cost-effectiveness compared to conventional radiologist evaluations, attempts to heuristically modify or augment these algorithms with non-complementary human methodologies can yield unintended economic and clinical consequences. **Methods:** The study incorporated a cost-effectiveness analysis utilizing an unpublished foundational dataset encompassing 11,424 participants from the National Lung Screening Trial (NLST), of whom 99 were diagnosed with lung cancer. The study evaluated three diagnostic workflows: Baseline (radiologist as sole reader), Sybil Cost-Saving (Sybil first reader; radiologist reviews positives), and Sybil Safeguard/Hybrid (Sybil first reader; radiologist reviews negatives). Performance metrics and total medical costs, based on the 2024 Medicare fee schedule, were compared. **Findings:** The Sybil Cost-Saving Strategy reduced average per-patient costs from $421 to $324, realizing a total cohort savings of $1,110,594. It identified seven additional true positives and 13 fewer false positives than the Baseline. Conversely, the Sybil Safeguard/Hybrid Strategy triggered a substantial surge of 561 false positives. This near-doubling of false positives erased the cost savings, increasing the overall total medical cost by $1,211,086 over the Sybil Cost-Saving strategy (an incremental cost of $173,012 per additionally identified true positive). **Conclusions:** Modifying autonomous AI algorithms through piecemeal human adjudication subverts economic efficiency and severely inflates false-positive rates due to divergent human-machine evaluation criteria. Policymakers and clinicians must carefully weigh the marginal sensitivity benefits against the exponential surge in false positives and profound opportunity costs.

Keywords: cost-effectiveness analysis, clinical AI, lung cancer screening, Sybil, false positives, health economics, human-AI teaming, diagnostic errors

1. Introduction

Sybil is an open-access artificial intelligence algorithm developed at MIT to predict future lung cancer risk from low-dose chest computed tomography. Sybil outperforms radiologists in assessing future risks of developing lung cancer (Ma et al., 2025). The underlying assumptions with the documented efficiency of Sybil would imply heightened welfare benefits for patients with the potential for early detection and economic advantages by treating cancer patients before the disease advances to the more critical stages III and IV.

This case study of Sybil investigates the potential for diseconomies and adverse ethical consequences from modifying the prescribed protocols of the original algorithm. Though this case study underscores the notable deleterious outcomes of veering from specified procedures, some other AI algorithms could, in theory, benefit from reworking. For some applications, augmenting computer programs with human interventions and physician participation could, in some instances, engender patient trust and strengthen efficacy without sacrificing economic gains or ethical principles by converting false negatives into true positives. Superficial recalibration of trained AI algorithm with inessential human mediation risks jeopardizing both economics and ethics.

Research by Philip Ma, MD and MS, a pulmonologist in Hong Kong, in an unpublished master's thesis (Ma et al., 2025), confirmed the economic gains from the use of Sybil relative to the conventional method of radiologists' reading of CT scans. Ma demonstrated cost effectiveness using a study cohort of 11,424 participants from the National Lung Screening Trial (NLST) dataset with actual medical costs, of whom 99 were diagnosed with lung cancer. The study reckoned that Sybil reduced costs by an average of $97 and increased incidences of true positives and true negatives and reduced the number of false positives and false negatives.

In his economic analysis, Ma experimented with an alternative methodology that synthesized the conventional radiologists' evaluation process with that of Sybil. The hybrid artifice reversed the cost savings of Sybil relative to the conventional approach by substantially expanding the volume of false positives. Despite the economic drawback of the hybrid approach, the co-joining of the conventional method with Sybil did fractionally increase the true positives by reducing the false negatives by the same amount. The novel approach did, however, uncover crucial economic hazards in the attempts to augment the efficiency of Sybil. Such forays with any other medical Al algorithm through piecemeal appending of non-complementary methodologies may significantly diminish the expected results.

The Ma research in this case study illustrates two potential, but interrelated, hazardous distortions of Al algorithms from attempts to improve on sensitivity or specificity or both. Firstly, tampering with the Al methodology may subvert economic efficiency through inadvertently incurring additional costs without compensating benefits. Secondly, human evaluations such as radiologists' readings may mismatch with the interpretation of the Al algorithm. This creates a fundamental divergence between human judgment and machine interpretation. With the Ma research, as a reference, this analysis examines how the misalignments between human and machine interpretations can impinge on economic efficiency without equal corresponding gains in cost savings or pronounced improvements in patient welfare.

2. Performance and Cost Comparison

This study utilized a previously collected, publicly available, and fully de-identified dataset from the NLST. Patient consent was not applicable. To properly contextualize the performance and economic metrics of screening strategies, this study evaluates three distinct workflows for lung cancer screening using low-dose CT (LDCT) scans. The Baseline Strategy is the conventional method where a radiologist serves as the sole reader to manually review and score all LDCT scans using the Lung-RADS system. In the Sybil Cost-Saving Strategy, the Sybil Al serves as the first reader. A human radiologist then acts as a second reader, but only reviews the scans that Sybil considers "positive". Finally, the Sybil Safeguard Strategy (Hybrid) also uses the Sybil AI as the first reader. However, a human radiologist serves as a second reader for scans that Sybil considers "negative" in order to catch potential false negatives. Radiologists do not review the scans that Sybil flags as positive. Table 1 summarizes the performance metrics (True Positive, False Positive, True Negative, False Negative) and cost differentials across these three workflows compared to the baseline.

Table 1. Performance and Cost Comparison

Strategy	True Positives (TP)	False Positives (FP)	True Negatives (TN)	False Negatives (FN)	Change in Total Cost vs. Baseline
Base case	47	758	10,567	52	Baseline reference
Sybil Cost-Saving	54	745	10,580	45	Saving: $988,309 - $1,232,878
Sybil Safeguard (Hybrid)	61	1,306	10,019	38	Additional Cost: $52,707 - $148,277

3. Summary Analysis of the Results

A comparison of the Base Case with Sybil indicates a performance difference. As Table 1 shows, Sybil scores a marginally higher ranking on the medical parameters for positive and negative predictors than the Base Case with only radiologists' assessments. Sybil labelled seven more true positives than in the Base Case with a corresponding seven fewer false negatives. For false positives, Sybil reduced the number by 13 compared with the Base Case, while increasing the number of true negatives by the same amount.

While Sybil exceeded the performance of the Base Case, the Al algorithm vastly outdid the radiologist in cost savings. The differential in average cost saving amounted to $1,110,594. The average cost per patient with Sybil shrank from $421 to $324. The average per patient saving amounted to $97. The cost reduction for culling out true positive patients showed even more substantial results of $641 per true positive patient. The per patient cost in the Base Case averaged $17,643, while with Sybil the per patient cost dwindled to $17,002 (Appendix A).

The Hybrid Case, in which radiologists assessed only the negative results from Sybil, outmatched the medical parameters of both the Base Case and Sybil. For true positives, the Hybrid Case identified seven more true positives than Sybil and 14 more than the Base Case with an equivalent seven and 14 reductions in false negatives, respectively. For true negatives and false positives, the difference between the Base Case and Sybil and the Hybrid Case dramatically shifted from incremental changes to almost doubling the volume of false positives and substantially decreasing the quantity of true negatives.

The Hybrid Case categorized 10,019 patients as true negative in comparison with 10,580 with Sybil and 10,567 in the Base Case. The combination of Sybil with radiologists evaluating the negative group added 561 patients to the number of false positives in Sybil alone. The bounce from 745 false positives with Sybil to 1,306 with Hybrid Case eradicated the cost savings with Sybil. This 561 surge in false positives from Sybil to the Hybrid Case amounts to a total cost approximately equal that of the: $665,632 for the Sybil Case and $1,487,321 for the Hybrid Case.

4. Discussion and Analysis

The Hybrid Case did convert an additional seven false negatives into true positives compared to the Sybil case. This refinement of the Sybil Case surely did benefit patients, especially those false negatives who might otherwise suffer unknowingly with lung cancer. Simultaneously, 561 false positive patients must endure the anxiety of believing that they are afflicted with cancer and are subjected to possible medical risks of further treatment such as invasive biopsies. The dilemma distils down to a trade-off between costs and benefits.

The crux of the quandary resurfaces with the conflict between ethics and economics. Some ethicists might argue that an early diagnosis of lung cancer for the seven converts justifies the additional expense of $1,211,086 or $173,012 incremental cost per patient despite layering on the potential risks and worries of the false positives. Thus, according to the argument, saving one life through early diagnosis of lung cancer warrants the outlay.

Humane economists might counter by maintaining the value of opportunity gains from investing in alternatives that save more than seven lives. The opportunity costs might exceed the value of rescuing the seven false negatives if the amount of additional costs were deployed in other areas such as breast cancer screening or innovative diagnostic tests. Meanwhile, rejecting the Hybrid strategy spares an additional 561 potential patients from the angst and perils of an interim false positive diagnosis.

The example of ad-hoc add-ons and other contrivances to Al algorithms as in the case of Sybil without a technical overhaul of the model highlights the potential complications. Adjustments to the algorithms tend to attempt to augment sensitivity, specificity, or both simultaneously. Ma's Hybrid Case ramped up the level of sensitivity, but with a financial cost that offset savings.

The Hybrid Case did extend sensitivity by gaining seven true positives which contracted false negatives by the same amount as the equations imply. The cross-bred model of Sybil and the Base Case, however, generated extreme results. The upshot of the correlation of the two parameters did conform to the mathematical interrelationship of sensitivity and specificity with an exaggerated number of false positives.

An inverse relation often exists between sensitivity and specificity. The increase in the sensitivity threshold, a higher accuracy rate which accurately identifying patients with a condition creates, produces an uptick in false positives. Conversely, upping the level of specificity, a higher accuracy in predicting those without a condition, spawns a bounce in false negatives.

The ratio of true positives, defined as the ratio of true positives to the sum of true positives plus false negatives, characterizes the level of sensitivity. A similar ratio marks the level of specificity: the ratio of true negatives to the sum of true negatives plus false positives. In the Ma study these ratios reflect the sensitivity and specificity levels under the three cases. The chart with the ratios of true positives and true negatives obscures the dramatic consequences of the Hybrid Case (Table 2).

Table 2. Detailed Performance Metrics

Strategy	Sensitivity	Specificity	True Positives (TP)	False Positives (FP)	True Negatives (TN)	False Negatives (FN)
Baseline	47.5%	93.3%	47	758	10,567	52
Sybil Cost-Saving	54.5%	93.4%	54	745	10,580	45
Sybil Safeguard (Hybrid)	61.6%	88.5%	61	1,306	10,019	38

The boost in sensitivity, which effectively edged the threshold to a higher percentage for designating members of the dataset as positive, behaved as the calculation would have predicted. True positives increased with a parallel decrease in false negatives. True negatives shrank while false positives ballooned. The upswing in false positives deviated from the relative magnitude from the Base to the Sybil Case.

The ratios are telling. In the Base Case, true negatives totaled 10 567 with 758 false positives. With Sybil, false positives lessened by 13 to 745 with a bump of 13 true negatives to 10,580. In the Base Case, radiologists found 14 true negatives for each false positive. (10,567/758).

With Sybil, false positives shaved off 13 for a total of 745 with an average of 14 true negatives for each false positive. (10,580/745). The Al algorithm transferred 13 false positives to the true negative category and seven false negatives to the true positive column. These relatively minor adjustments contributed approximately an overall cost savings ranging from $988,309 to $1,232,878.

The Hybrid Case posted substantially more changes relative to Sybil than from the Base Case to Sybil. With radiologists interpreting the negatives from Sybil in the Base Case, false positives swelled to almost double that of Sybil and the Base Case. False positives escalated to 1,306 compared with 745 with Sybil. The ratio of true negatives to false positives essentially halved with the increase in contrast with the absolute reduction in false positives from Base Case to Sybil. The precipitous spike in false positives effectively cancelled the cost saving gains from Sybil.

Why false positives in the Hybrid Case hurtled to almost twice that of Sybil portends an economic caution about injudicious modification of the AI algorithm's process with human interloping. The bloating of false positives in the Hybrid Case, while decimating cost savings and efficiencies, transferred only seven patients from false negative to true positive. This anomalous spurt of economically averse false positives in the Hybrid Case, which could imperil other medical Al algorithms with injudicious modification of the AI algorithm's process with human interloping, originates from two peculiarities of the Al algorithms.

The first concerns differences in data subsets within the AI algorithm. Again, the lung cancer study highlights the shifts in sample data being analyzed. The three variations segregated the aggregate dataset of (11,424) patients into three distinct subsets to be scrutinized under different parameters for evaluation.

Firstly, in the Base Case, radiologists manually review the entire dataset with the objective of classifying each patient image as positive or negative. The sequence in the Sybil Case changed along with the dataset to be manually reviewed. Radiologists reviewed only the positives that Sybil selected. The negatives were determined solely by Sybil, uninspected by radiologists.

While, in principle, the physicians appraised all positive patients in the Base Case, they were nonetheless surveying a different set of data in Sybil, a subset of the Base Case dataset. With Sybil, the radiologists classified seven more true negatives and seven less false negatives than in the Base Case. For whatever reason, the physicians, who implicitly agreed with Sybil's selection, overlooked those seven positives that either the algorithm or they reclassified from the Base Case.

A similar oversight happened in the Hybrid Case in relation to Sybil. In the Hybrid Case, radiologists inspected only the negative results designated by Sybil. As with Sybil, the physicians scanned all the images in the Base Case, including those judged to be negative in the Hybrid Case. These negatives, though a subset of the base dataset, form a sample different from that in the Base Case as the true and false positives from Sybil are excluded. The number of negatives in Sybil for physicians' readings totals 10,625 (10,580+45).

From this data subset, radiologists labeled 561 additional patients as false positives. This incremental increase in false positives, when combined with the false positives in Sybil, tops up at 1,306. Perhaps, knowing that the collection of images consisted of only negatives with the positives omitted, radiologists, being risk averse, aggressively flagged any suspicious or anomalous images as positives to avoid the possibility of being corrected by a machine (Columbia University, 2026).

Secondly, the disparity in false positives between cost-saving Sybil and the safeguard hybrid case suggests a difference in the evaluating process between the algorithm and radiologists. The difference between the number of false positives in the base case and the safeguard Sybil accentuates the distinct approaches. At the same time, the discrepancies reveal that, despite similarities in the number of false positives in the base case and the cost-saving Sybil, 758 and 745 respectively, the frequency of false positives nearly doubled to 1,306 in the safeguard Sybil case.

The process by radiologists is baffling. When radiologists sorted through the entire patient set to parse positives and negatives, the physicians classified 561 patients as false positives, while cost-saving Sybil shuffled 548 patients into the false positive trough. The two numbers, though similar in magnitude, are deceiving. The composition of the two classifications of false positives in the two methodologies consists of substantially different patients. From the dataset from the NLST, the analysis of patients determined only 197 patients in common in the base case, the cost-saving Sybil, and the safeguard hybrid. Table 3 displays evidence of substantial differences in the configurations of false positives in each of the three cases. Of the 758 and 745 false positives in the base case and the cost-saving Sybil, 561 and 548 respective patients were compiled in the two categories. The total of false positives in the base case and Sybil represents the two distinct groups of false positives added to the 197 patients in common. For the hybrid case, the total is comprised of the 197 patients in common plus the 561 and 548 false positives in the base case and Sybil respectively.

The independent groupings of false positives confirm that in the variations in the applications of Sybil, radiologists and the AI algorithm interpret the characteristics and the features of CT images significantly differently. Otherwise, the number of false positives would yield a greater ratio of false positives in common with the totals of the base and Sybil cases than the fraction of false positive patients. The divergences among the three AI approaches with heterogeneous outcomes implies distinctive and idiosyncratic modes of evaluating images.

The example of alternative evaluation methods with the Sybil algorithm warrants heeding the potential for subverting the potential of clinical AI. Indeed, some AI processes in their design may necessitate human intervention. Inchoate research on patient perceptions of medical AI affirms that trust in machine learning solidifies more rapidly and more thoroughly with human involvement and oversight than with sole reliance on an impersonal algorithm (Aguilar, 2026). The crucial aspect in integrating radiologists and other caregivers in the process centers on merging AI applications with human intervention with supplemental connections compatible with the objectives of AI technology. Inevitably, with any medical AI technology, humans will engage with the process in ways that enhance AI outcomes and predictions rather than competing directly with a trained algorithm.

Table 3. Non-Overlapping True Negative (TN)

Strategy	**True Positives (TP)**	**False Positives (FP)**	**Common FP**	**Doctor -only FP**	**Sybil-only FP**	**True Negatives (TN)**	**Common TN**	**Doctor-only TN**	**Sybil-only TN**	**False Negatives (FN)**
Baseline	47	758	197	561		10,567	10,019	548		52
Sybil Cost-Saving	54	745	197		548	10,580	10,019		561	45
Sybil Safeguard (Hybrid)	61	1,306	197	561	548	10,019	10,019			38

As Table 3 shows, the total number of non-overlapping True Negatives (TNs) is 1,109, comprising 548 Doctor-only TNs and 561 Sybil-only TNs (for the mathematical proof, see Appendix B).

The disparity in false positives between the Sybil and Hybrid Cases confirms that, at least in some instances, differences occur in respective compositions in the positive and negative categories. Even if all the false positives in the Sybil Case correspond one-to-one with the false positives, except for the 13 superfluous members in the Base Case, a surplus of false positives still accrues in the Hybrid Case. Theoretically, all the members in the false positives of the Base Case may populate true negatives in the Sybil or Hybrid Cases and vice versa.

Any similarities in magnitude of the categories of the three cases do not necessarily imply identical compositions. The Hybrid Case with 561 extra false positives than Sybil points to a potential difference in the compositions among the three cases. The patients in the Base Case's false positives may not appear in Sybil's. Then, some or all of the false positives in the Base Case may populate one or more other categories in the Sybil and Hybrid Cases. This potential for distinctive compositions in modified Al algorithms could prejudice results with sub-optimal outcomes relative to the algorithm's output alone.

The variation in sizes of categories, either positive or negative among the three cases, reveals that radiologists' interpretations may misalign with that of the algorithm. In the lung cancer study, radiologists cleaved off more false positives in the Hybrid Case than they did in the Base Case. If the radiologists concurred 100 percent with Sybil's criteria, the number of false positives would coincide exactly.

A comparison of overlapping patients between the Baseline and Sybil Cost-Saving strategies reveals a divergence in how AI and human practitioners segregate false positives. While both the AI algorithm and radiologists tend to identify true and false positives with comparable aggregate proficiency, their underlying evaluation criteria clearly misalign. Radiologists frequently deem specific morphological characteristics as positive that the AI algorithm correctly dismisses as negative, and vice versa.

This human-AI divergence introduces economic and ethical questions (Schuller & Hyun, 2025). While the Hybrid strategy's identification of seven additional true-positive cases yields undeniable clinical benefit for those individuals, it does so at a systemic cost. Tampering with AI algorithms by imposing human judgement to address untargeted patients may increase the false-positive rate and drive-up total costs. From a macroeconomic health perspective, the opportunity costs of this expenditure are severe. The capital absorbed by re-screening and testing errant false positives could alternatively be directed to re-screening and testing programs for breast cancer that could save multiple women, identifying patients with other curable diseases, or purchasing medical innovations to treat other maladies.

5. Limitations

A notable limitation of this retrospective study is that the dataset does not permit a granular, case-by-case determination of whether radiologists successfully marked specific true positives that the Sybil algorithm missed, and vice versa. Consequently, while the aggregate divergence in false-positive segregation between human and machine evaluators is evident, the exact feature-level discrepancies in true-positive identification remain an area requiring future prospective investigation.

6. Conclusion

Human intervention isn't inherently bad, only when dysfunctional. The temptation to "soup up" Al algorithms with modification that merge human judgement into the computation can skew the integrity of outcomes to a sub-optimal level. The skewness can offset economic gains, present intractable ethical concerns, and foment worry and anxiety as well as medical risks to patients. This viewpoint recognizes that some Al models could warrant modifications. No algorithm can achieve perfection. Yet, if practitioners do intervene incorporating human decisions, they must ensure against the possibility of creating biased subsets and recognize that humans and machines may operate with different criteria that detract from the benefits of Al.

Authors' contributions

Frank C. Schuller: Conceptualization, Investigation, Methodology, Formal analysis, Validation, Visualization, Writing - original draft, Writing - review & editing.

David Guo: Conceptualization, Investigation, Methodology, Data curation, Formal analysis, Validation, Visualization, Writing - review & editing.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Competing interests

We declare no competing interests.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of Sciedu Press.

The journal and publisher adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The dataset from the National Lung Screening Trial (NLST) is available upon request from the National Cancer Institute. The analytical code and specific subsets used in this study will be made available upon reasonable request to the corresponding author without additional restrictions

Data sharing statement

No additional data are available.

References

Aguilar, M. (2026). The AI push in health care is deepening medicine's trust crisis. STAT. Retrieved from https://www.statnews.com/2026/03/23/ai-use-distrust-american-medicine/

Columbia University. (2026) When Clinical Judgment and AI Diverge: The Clinical, Legal, and Ethical Considerations [Video]. YouTube. Retrieved from https://www.youtube.com/watch?v=PWOPatb631w

Ma, P., *et al.*. (2025). *Lung cancer screening with and without the open-access Sybil algorithm: Cost-effectiveness analysis of the National Lung Screening Trial.* Hong Kong.

Schuller, F. C., & Hyun, I. (2025). AI-driven diagnostics need an adaptive regulatory approach to benefit patients. *Journal of Medical Artificial Intelligence, 8*(27). Retrieved from https://jmai.amegroups.org/article/view/9756

Appendix A. Detailed Cost Computations

The cost computations are based on a study cohort of 11,424 participants. Total Cost Savings Differential is $1,110,594, calculated as the average of the savings range lower bound ($988,309) and upper bound ($1,232,878).

- The average savings per participant is $97 (total savings of $1,110,594 divided by 11,424 participants).
- The Average Cost Per Patient is $421 for the Baseline strategy (average of $3,788,272 and $5,841,473 divided by 11,424) and $324 for the Sybil strategy (average of $2,799,963 and $4,608,595 divided by 11,424).
- The cost reduction for True Positive (TP) patients is $641 per TP patient .
- The average cost per TP patient in the Baseline is $17,643 (average screening of $192.00 plus average diagnostic workup of $17,451) , while the Sybil average cost per TP patient is $17,002 (average screening of $139 plus average diagnostic workup of $16,863).

Appendix B. Mathematical Proof of Non-Overlapping True Negatives

- Non-overlapping TN = Doctor-only TNs (548) + Sybil-only TNs (561) = 1,109
- Doctor-Only TN (548) is identical to Sybil-Only FP (548) by definition. The following is the proof. Both strictly apply only to a healthy patient.

Direction 1: If Doctor-Only TN → Then Sybil-Only FP

1. Doctor-Only TN: By definition, the patient is healthy, and the Doctor correctly identified them as negative.
2. The "Only" Constraint: Since the Doctor is the *only* one correct, Sybil must have been incorrect.
3. Sybil's Error: The only way to be incorrect about a healthy patient is to mark them positive.
4. Conclusion: A healthy patient marked positive is a False Positive (FP). Since the Doctor was negative, this error is Sybil-Only.

Direction 2: If Sybil-Only FP → Then Doctor-Only TN

1. Sybil-Only FP: By definition, the patient is healthy, and Sybil incorrectly marked them as positive.
2. The "Only" Constraint: Since Sybil is the *only* one with the False Positive error, the Doctor did *not* mark them positive.
3. Doctor's Success: If the Doctor did not mark a healthy patient positive, they must have marked them negative.
4. Conclusion: A healthy patient marked negative is a True Negative (TN). Since Sybil was positive, this success is Doctor-Only.