Confusion accounting

ORIGINAL ARTICLE

Cell-free DNA Analysis for Noninvasive Examination of Trisomy

Mary E. Norton, M.D., Bo Jacobsson, M.D., Ph.D., Geeta K. Swamy, M.D., Louise C. Laurent, M.D., Ph.D., Angela C. Ranzini, M.D., Herb Brar, M.D., Mark W. Tomlinson, M.D., Leonardo Pereira, M.D., M.C.R., Jean L. Spitz, M.P.H., Desiree Hollemon, M.S.N., M.P.H., Howard Cuckle, D.Phil., M.B.A., Thomas J. Musci, M.D., et al.

Article	Figures/Media		Metrics

26 References 460 Citing Articles Letters 3 Comments

Abstract

Related Articles

The confusion matrix

TRUTH

N
ok

P
trisomy

TEST

— negative test

	N ok	P trisomy
— negative test	TN	FN
+ positive test	FP	TP

+ positive test

The confusion matrix

TRUTH

	N ok	P trisomy
TEST − negative test	TN	FN
+ positive test	FP	TP

TRUTH

	N	P	
− negative test	a	b	a+b
+ positive test	c	d	c+d
	a+c	a+d	a+b+c+d

This table has four numbers in it; one of the numbers a+b+c+d is total sample size; another number (b+d)/(a+b+c+d) is the class imbalance.

The confusion matrix

TRUTH

	N ok	P trisomy
TEST		
− negative test	TN	FN
+ positive test	FP	TP

TRUTH

	N ok	P trisomy	
− negative test	15,794	0	15,794
+ positive test	9	38	47
	15,803	38	15,841

Binary classification squared: confusion matrices

- Accuracy: proportion of examples labeled correctly
- We need to know true labels to calculate accuracy
- For binary classification, we can call the two labels 'yes' and 'no' (for example, disease prediction)
- **true positives (TP):** predicted yes (they have the disease), and they do have the disease.
- **true negatives (TN):** predicted no, and they don't have the disease
- **false positives (FP):** predicted yes, but they don't actually have the disease
- **false negatives (FN):** predicted no, but they actually do have the disease

Rates computed from the confusion matrix

- **Accuracy**: (TP+TN)/total = (15794+38)/ 15841 (how often is the classifier correct overall). Crude accuracy weights TPR and TNR by true-class-proportions at evaluation time.
- **Misclassification rate**: (FP+FN)/total = (0+9)/15,841 (how often is it incorrect; 1-Accuracy).
- **True Positive Rate**: TP/ Actual YES (also known as sensitivity or recall)
- **False Positive Rate:** FP/= truly negative
- **True Negative Rate**: TN/Actual NO   (also known as specificity)
- **Precision=Positive Predictive Value**: TP/Predicted YES=100/110=0.91 (when it predicts YES, how often is it correct)

Good?

- There's a 0 false negatives. Out of 38 patients with trisomy in the cohort, all got correct (positive) test results.
- **The maximum-likelihood answer "FNR = 0 !!!" is not the most reasonable answer**; Bayesian approaches will "regularize" 0/38, biasing it a small distance away from 0 depending on how strong an assumption you're willing to make.
- The predictive value of the tests is asymmetrical; a negative test is overwhelmingly likely to be correct. A positive test is likely to be correct, but not overwhelmingly so (FPR)
- These particular technologies, in the mid 2010s, blew away the standard of care in terms of accuracy (for trisomy detection)

The confusion matrix

TRUTH

TRUTH IN DENOMINATOR

$$TPR = TP/P = TP / (TP + FN) = SENSITIVITY$$

$$TNR = TN/N = TN / (TN + FP) = SPECIFICITY$$

N
ok

P
trisomy

TEST

−
negative
test

	TN	FN
	FP	TP

+
positive
test

Ariosa® DIAGNOSTICS

Harmony™ PRENATAL TEST

The confusion matrix
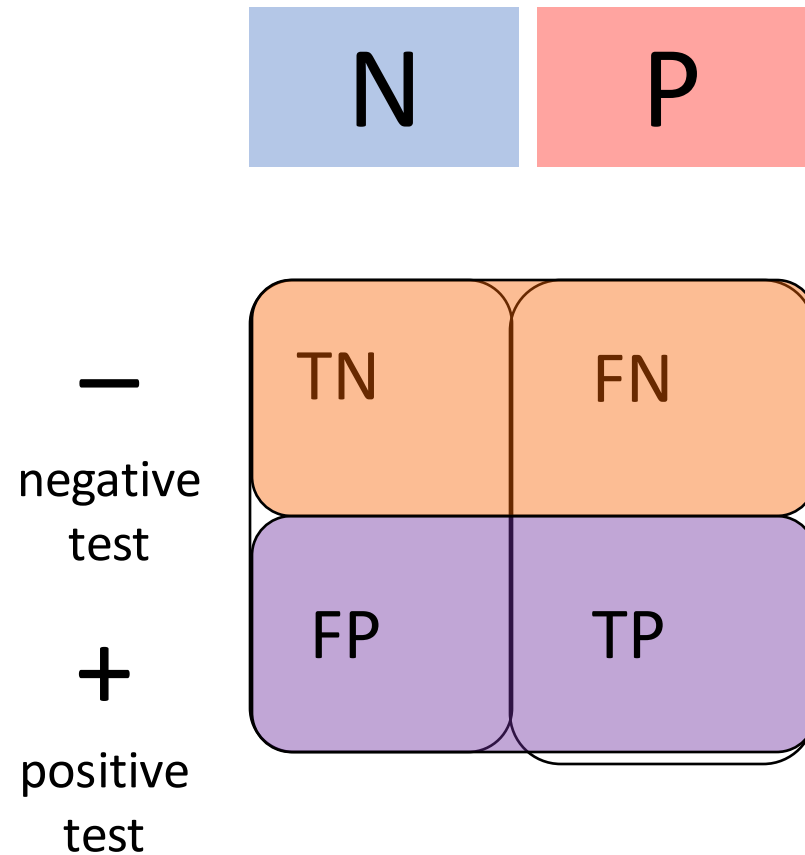
TRUTH

TRUTH IN DENOMINATOR

N
ok

P
trisomy

$\text{TPR} = \text{TP/P} = \text{TP} / (\text{TP} + \text{FN}) = \text{SENSITIVITY}$

$\text{TNR} = \text{TN/N} = \text{TN} / (\text{TN} + \text{FP}) = \text{SPECIFICITY}$

−

negative
test

TN
15,794

FN
0

15,794

+

positive
test

FP
9

TP
38

47

15,803

38

15,841

When you are evaluating a test (or an algorithm) you must Know The Truth.

N	P

"Ground truth" or
"Gold standard"

—

negative
test

+

positive
test

TN	FN
FP	TP

TRUTH IN DENOMINATOR

TPR = TP/P = TP / (TP + FN) = SENSITIVITY

TNR = TN/N = TN / (TN + FP) = SPECIFICITY

When you are evaluating a test (or an algorithm) you must Know The Truth.
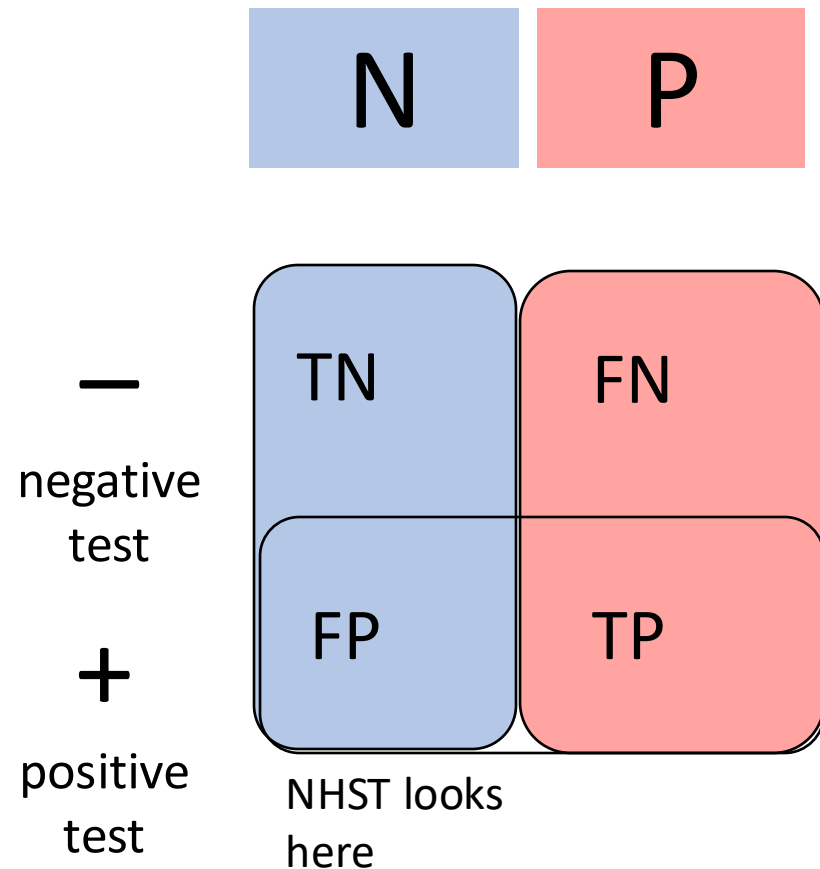
	N	P

"Ground truth" or "Gold standard"

—
negative test

+
positive test

	N (TN, FP)	P (FN, TP)
—	TN	FN
+	FP	TP

NHST looks here

TRUTH IN DENOMINATOR

TPR = TP/P = TP / (TP + FN) = SENSITIVITY

TNR = TN/N = TN / (TN + FP) = SPECIFICITY

SENSITIVITY AKA RECALL

When you are heading home from the doctor's office, TPR and TNR are not the ratios you need.

N	P

TRUTH IN DENOMINATOR

TPR = TP/P = TP / (TP + FN) = SENSITIVITY

TNR = TN/N = TN / (TN + FP) = SPECIFICITY

— negative test

TN	FN
FP	TP

+ positive test

TEST IN DENOMINATOR

NPV = TN/Ntest = TN / (TN + FN)

PPV = TP/Ptest = TP / (TP + FP) = PRECISION

The guide from Sensitivity, Specificity to PPV and NPV is Bayes Rule + a necessary estimate of class balance P(positive)

$$\text{PPV} = \text{P(Positive | + Test)} = \frac{\text{P (Test + | Positive) P(Positive)}}{\text{P (Test +)}}$$

$$\text{PPV} = \text{P(Positive | + Test)} = \frac{\text{P (Test + | Positive) P(Positive)}}{\text{P (Test + | Positive) P (Positive)} + \text{P (Test + | Negative) P (Negative}}$$

The guide from Sensitivity, Specificity to PPV and NPV is Bayes Rule + a necessary estimate of class balance P(positive)

$$PPV = P(\text{ Positive } | + \text{ Test }) = \frac{P(\text{ Test } + | \text{ Positive }) P(\text{Positive})}{P(\text{Test } +)}$$

$$PPV = P(\text{ Positive } | + \text{ Test }) = \frac{P(\text{ Test } + | \text{ Positive }) P(\text{Positive})}{P(\text{Test } + | \text{ Positive}) P(\text{Positive}) + P(\text{Test } + | \text{ Negative}) P(\text{Negative}}$$

$$NPV = P(\text{ Positive } | - \text{ Test }) = \frac{P(\text{ Test } - | \text{ Positive }) P(\text{Positive})}{P(\text{Test } -)}$$

$$NPV = P(\text{ Positive } | - \text{ Test }) = \frac{P(\text{ Test } - | \text{ Positive }) P(\text{Positive})}{P(\text{Test } - | \text{ Positive}) P(\text{Positive}) + P(\text{Test } - | \text{ Negative}) P(\text{Negative})}$$

And of course, overall accuracy

ACCURACY = P(positive) * SENSITIVITY + (1- P(positive)) * SPECIFICITY

The conditions for the experiment which reveals SENSITIVITY and SPECIFICITY are not generally the conditions where the test / classifier will be deployed.

(They aren't if the experiment was designed well!)

Overall accuracy of 99.5% for a useless cancer screening that classifies everyone as OK. You can't communicate a classifier's accuracy with one or two numbers!

What do wee need to optimize?

- Total accuracy ? TP + TN ??

- Some kind of average of precision and recall?

Oftentimes I have a knob that permits a range of values for TPR and TNR. How to choose the value of that knob?

- Equal numbers of False positives and False negatives ??

They Trusted Their Prenatal Test. They Didn't Know the Industry Is an Unregulated "Wild West."

by Anna Clark, Adriana Gallardo, Jenny Deam and Mariam Elba

Dec. 6, 2022, 6 a.m. EST

https://www.propublica.org/article/how-prenatal-screenings-have-escaped-regulation

As regulators stay on the sideline, a growing industry expands its reach but leaves some pregnant patients feeling misled and heartbroken.

Key Findings

•While upwards of half of all pregnant women get noninvasive prenatal screening tests, or NIPTs, the tests are not regulated by the U.S. Food and Drug Administration. "This is a Wild West scenario," said one expert.

•After fierce industry backlash, the FDA retreated on oversight of lab tests, including NIPTs. Not pushing back more, a former agency official said, "remains one of my greatest regrets."

•Experts say the screenings were sold before they were appropriately tested. Companies downplay "inconvenient truths" in the research, said one doctor.

•Marketing materials have sometimes pitched the tests as providing far more certainty than they actually do. The statistical nuances of the test aren't easy to parse for patients and even some doctors and nurses.

•While patients have been left confused and sometimes shattered, executives profit. Last year, the compensation package for the head of one lab was over $23 million.

Prenatal Test.
he Industry Is
Vild West."

n and Mariam Elba

lica

As regulators stay on the sideline, a growing industry expands its reach but leaves some pregnant patients feeling misled and heartbroken.

They Trusted Their Prenatal Test. They Didn't Know the Industry Is an Unregulated "Wild West."

by Anna Clark, Adriana Gallardo, Jenny Deam and Mariam Elba

Dec. 6, 2022, 6 a.m. EST

lica

As regulators stay on the sideline, a growing industry expands its reach but leaves some pregnant patients feeling misled and heartbroken.

Journalists can't recruit cohorts of 20,000 pregnant people and karyotype every one. They can raise the alarm that the study may be falsely optimistic (or that the technology's accuracy has changed as commercial optimization has progressed) but they haven't got the firepower to Count The Cases.

They Trusted Their Prenatal Test. They Didn't Know the Industry Is an Unregulated "Wild West."

by Anna Clark, Adriana Gallardo, Jenny Deam and Mariam Elba

Dec. 6, 2022, 6 a.m. EST

lica

As designed, it's not a fair fight.

As regulators stay o̶ leaves some pregnant patients feeling misled and heartbroken.