Quick Intro to Sensitivity, Specificity, PPV, and NPV

Zachary Binney, PhD

April 4, 2020

Georgia State University | SCHOOL OF PUBLIC HEALTH

Screening Tests: Validity – Result Combinations

- 4 possible combinations of truth (gold standard) (diseased/nondiseased) and test results (positive/negative)
  - **True positives (a):** screen positive + have disease **(GOOD!)**
  - **False positives (b):** screen positive + don't have disease **(BAD!)**
  - **False negatives (c):** screen negative + have disease **(BAD!)**
  - **True negatives (d):** screen negative + don't have disease **(GOOD!)**

		Condition According to Gold Standard			
		Present	Absent	Total	
Test result	Positive	a = True positives	b = False positives	a + b	Predictive value (+) $\frac{a}{a+b}$
	Negative	c = False negatives	d = True negatives	c + d	Predictive value (–) $\frac{d}{c+d}$
	Total	a + c	b + d	**Grand total a + b + c + d**	
		Sensitivity $\frac{a}{a+c}$	Specificity $\frac{d}{b+d}$		

Screening Tests: Validity – Metrics

- 4 numbers important to a test's validity:
  - **Sensitivity (Se)**
  - **Specificity (Sp)**
  - **Positive Predictive Value (PPV)** or **Predictive Value (+)**
  - **Negative Predictive Value (NPV)** or **Predictive Value (-)**
- **Higher is better for all numbers!**

		Condition According to Gold Standard			
		Present	Absent	Total	
Test result	Positive	a = True positives	b = False positives	a + b	Predictive value (+) $\frac{a}{a+b}$
	Negative	c = False negatives	d = True negatives	c + d	Predictive value (–) $\frac{d}{c+d}$
	Total	a + c	b + d	**Grand total a + b + c + d**	
		Sensitivity $\frac{a}{a+c}$	Specificity $\frac{d}{b+d}$		

Screening Tests: Validity – Sensitivity and Specificity

- **Sensitivity (Se):** the ability of the test to correctly identify individuals who actually have the disease
  - Math: $\frac{a}{a+c} = \frac{True\ Positives}{Total\ Who\ Have\ the\ Disease}$
    - Ranges 0-100%
  - English: "What percent of sick people did the test catch?"

		Condition According to Gold Standard			
		Present	Absent	Total	
Test result	Positive	a = True positives	b = False positives	a + b	Predictive value (+) $\frac{a}{a+b}$
	Negative	c = False negatives	d = True negatives	c + d	Predictive value (–) $\frac{d}{c+d}$
	Total	a + c	b + d	**Grand total a + b + c + d**	
		Sensitivity $\frac{a}{a+c}$	Specificity $\frac{d}{b+d}$		

Screening Tests: Validity – Sensitivity and Specificity

- **Specificity (Sp):** The ability of the test to identify **non-diseased** individuals who actually do not have the disease
  - Math: $\frac{d}{b+d} = \frac{True\ Negatives}{Total\ Who\ Don't\ Have\ the\ Disease}$
    - Ranges 0-100%
  - English: "What percent of non-sick people did the test correctly screen out?"

		Condition According to Gold Standard			
		Present	Absent	Total	
Test result	Positive	a = True positives	b = False positives	a + b	Predictive value (+) $\frac{a}{a+b}$
	Negative	c = False negatives	d = True negatives	c + d	Predictive value (–) $\frac{d}{c+d}$
	Total	a + c	b + d	**Grand total a + b + c + d**	
		Sensitivity $\frac{a}{a+c}$	Specificity $\frac{d}{b+d}$		

Screening Tests: Validity – Sensitivity and Specificity

- Both Se & Sp important, but which is more important varies from test to test
  - Bad **Sensitivity**: lots of false negatives
    - Especially bad when the costs of missing a case are high (e.g. you're fine, go home)
  - Bad **Specificity**: lots of false positives
    - Especially bad when the costs of misidentifying someone as + are high (e.g. you're immune when you're not)
- **Se & Sp often a tradeoff** – improving one may require worsening the other. Do you need to be able to ID every case vs. are false positives extremely bad?

		Condition According to Gold Standard			
		Present	Absent	Total	
Test result	Positive	a = True positives	b = False positives	a + b	Predictive value (+) $\frac{a}{a+b}$
	Negative	c = False negatives	d = True negatives	c + d	Predictive value (–) $\frac{d}{c+d}$
	Total	a + c	b + d	**Grand total a + b + c + d**	
		Sensitivity $\frac{a}{a+c}$	Specificity $\frac{d}{b+d}$		

Screening Tests: Validity – Sensitivity and Specificity

- Example:
  - 1,000 people
  - 100 (10%) have disease

Test Results	**Gold Standard Result / "True" Disease Status**		**Total**
	Diseased	**Not Diseased**	
Positive	a = 80	b = 100	**180**
Negative	c = 20	d = 800	**820**
Total	**a + c = 100**	**b + d = 900**	**1,000**

Screening Tests: Validity – Sensitivity and Specificity

- **Sensitivity:** percent of people with disease who test positive
  - 100 have disease, 80 of them test positive
  - **Sensitivity** = 80/100 = 80%

Test Results	**Gold Standard Result / "True" Disease Status**		**Total**
	Diseased	**Not Diseased**	
Positive	**a = 80**	b = 100	**180**
Negative	c = 20	d = 800	**820**
Total	**a + c = 100**	**b + d = 900**	**1,000**

Screening Tests: Validity – Sensitivity and Specificity

- **Specificity:** percent of people with*out* disease who test negative
  - 900 don't have disease, 800 of them test negative
  - **Specificity** = 800/900 = 89%

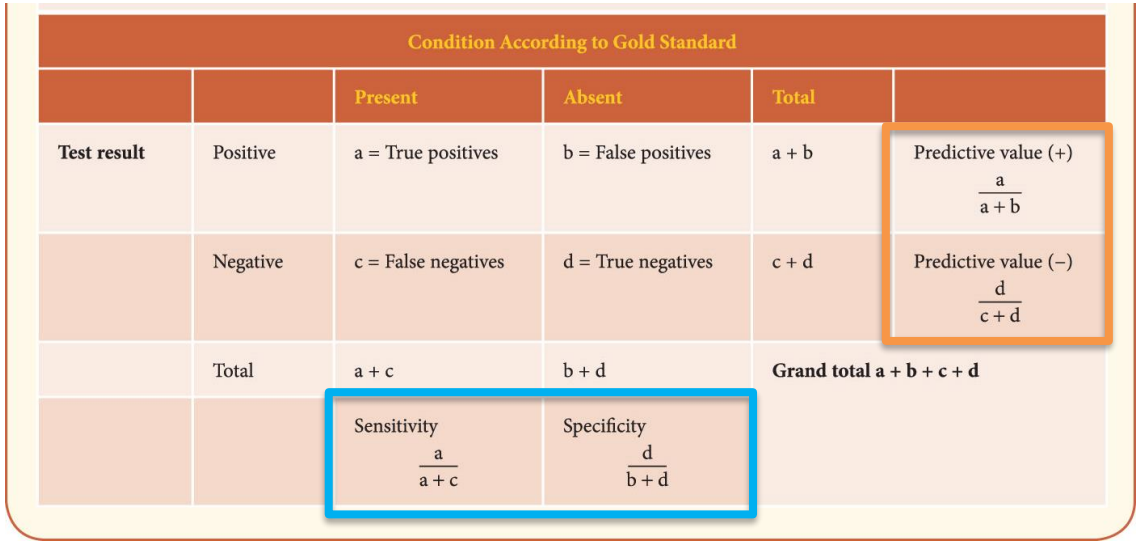
Test Results	**Gold Standard Result / "True" Disease Status**		**Total**
	Diseased	**Not Diseased**	
Positive	a = 80	b = 100	**180**
Negative	c = 20	**d = 800**	**820**
Total	**a + c = 100**	**b + d = 900**	**1,000**

Screening Tests: Validity – Metrics

- Imagine you go into a doctor's office and get a blood test for HIV
  - **Sensitivity**: If you have HIV, how likely is the test to be positive?
  - **Specificity**: If you don't have HIV, how likely is the test to be negative?
  - How useful are these questions?
    - Why would we do the test if we already knew you were or weren't sick?
- Flip it around to get more relevant questions:
  - **If you test positive, how likely is it you have HIV?** – This is **PPV**
  - **If you test negative, how likely is it you *don't* have HIV?** – This is **NPV**

Screening Tests: Validity – Metrics

- 4 numbers important to a test's validity:
  - **Sensitivity (Se)**
  - **Specificity (Sp)**
  - **Positive Predictive Value (PPV)** or **Predictive Value (+)**
  - **Negative Predictive Value (NPV)** or **Predictive Value (-)**
- **Higher is better for all numbers!**



		Condition According to Gold Standard			
		Present	Absent	Total	
Test result	Positive	a = True positives	b = False positives	a + b	Predictive value (+) $\frac{a}{a+b}$
	Negative	c = False negatives	d = True negatives	c + d	Predictive value (–) $\frac{d}{c+d}$
	Total	a + c	b + d	**Grand total a + b + c + d**	
		Sensitivity $\frac{a}{a+c}$	Specificity $\frac{d}{b+d}$		

Screening Tests: Validity – PPV and NPV

- **PPV or PV+:** proportion of patients who test positive who actually have the disease
  - Math: $\frac{a}{a+b} = \frac{True\ Positives}{Total\ Who\ Test\ Positive}$
    - Ranges 0-100%
  - English: "If you test positive, how likely is it that you're really sick (or immune in the case of an antibody test)?"

		Condition According to Gold Standard			
		Present	Absent	Total	
Test result	Positive	a = True positives	b = False positives	a + b	Predictive value (+) $\frac{a}{a+b}$
	Negative	c = False negatives	d = True negatives	c + d	Predictive value (–) $\frac{d}{c+d}$
	Total	a + c	b + d	**Grand total a + b + c + d**	
		Sensitivity $\frac{a}{a+c}$	Specificity $\frac{d}{b+d}$		

Screening Tests: Validity – PPV and NPV

- **NPV or PV-:** proportion of patients who test negative who do not have the disease
  - Math: $\frac{d}{c+d} = \frac{True\ Negatives}{Total\ Who\ Test\ Negative}$
    - Ranges 0-100%
  - English: "If you test negative, how likely is it that you're really not sick (or not immune in the case of an antibody test)?"

		Condition According to Gold Standard			
		Present	Absent	Total	
Test result	Positive	a = True positives	b = False positives	a + b	Predictive value (+) $\frac{a}{a+b}$
	Negative	c = False negatives	d = True negatives	c + d	Predictive value (–) $\frac{d}{c+d}$
	Total	a + c	b + d	**Grand total a + b + c + d**	
		Sensitivity $\frac{a}{a+c}$	Specificity $\frac{d}{b+d}$		

Screening Tests: Validity – PPV and NPV

- Example:
  - 1,000 people
  - 100 (10%) have disease

Test Results	**Gold Standard Result / "True" Disease Status**		**Total**
	Diseased	**Not Diseased**	
Positive	a = 80	b = 100	**a + b = 180**
Negative	c = 20	d = 800	**c + d = 820**
Total	**100**	**900**	**1,000**

Screening Tests: Validity – PPV and NPV

- **PPV:** probability of a positive test having disease
  - 180 test positive, 80 of them have disease
  - **PPV** = 80/180 = 44% (remember **sensitivity** was much higher at 80%)

Test Results	**Gold Standard Result / "True" Disease Status**		**Total**
	Diseased	**Not Diseased**	
Positive	**a = 80**	b = 100	**a + b = 180**
Negative	c = 20	d = 800	**c + d = 820**
Total	**100**	**900**	**1,000**

Screening Tests: Validity – PPV and NPV

- **NPV:** probability of a negative test *not* having disease
  - 820 test negative, 800 of them don't have disease
  - **NPV** = 800/820 = 98% (remember **specificity** was also high at 89%)

Test Results	**Gold Standard Result / "True" Disease Status**		**Total**
	Diseased	**Not Diseased**	
Positive	a = 80	b = 100	**a + b = 180**
Negative	c = 20	**d = 800**	**c + d = 820**
Total	**100**	**900**	**1,000**

Screening Tests: Validity – PPV and NPV

- Example: let's summarize...
  - **Se** = 80% **Sp**= 89%
  - **PPV** = 44% **NPV** = 98%

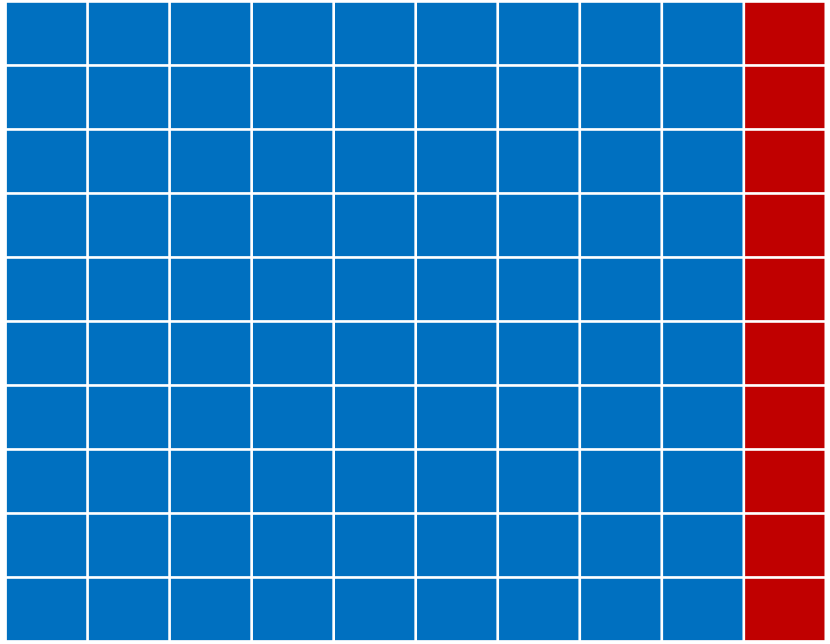
Common pattern for rare diseases: **test is right most of the time** (high Se & Sp), but **a single positive test doesn't mean that much** (low PPV)

Test Results	**Gold Standard Result / "True" Disease Status**		**Total**
	Diseased	**Not Diseased**	
Positive	a = 80	b = 100	**a + b = 180**
Negative	c = 20	d = 800	**c + d = 820**
Total	**100**	**900**	**1,000**

Screening Tests: PPV/NPV and Prevalence

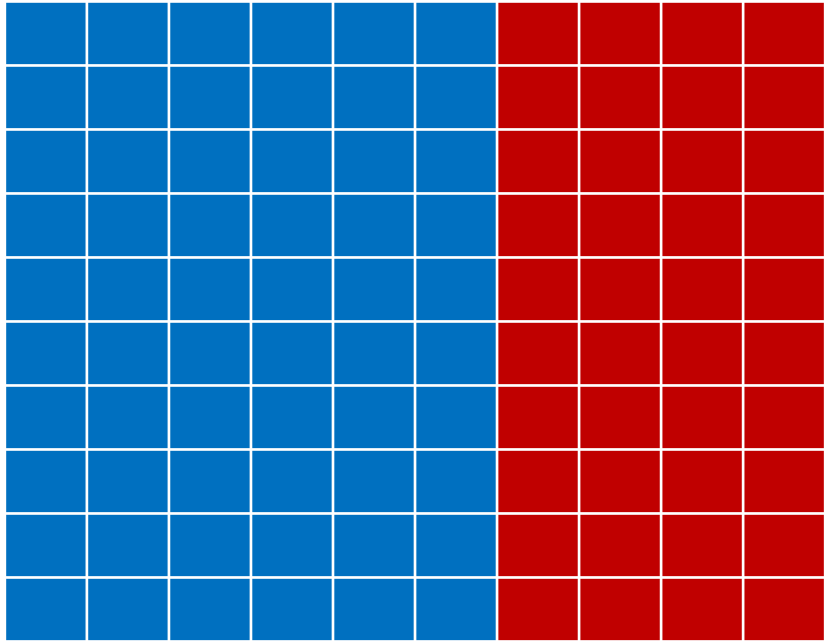
- **Sensitivity and Specificity**: **fixed characteristic of test**; the same for every population tested
- **PPV and NPV**: different for every population tested
  - Depends on disease prevalence!
- Consider two groups of 100 people with different prevalence of diabetes

Prevalence = 10%



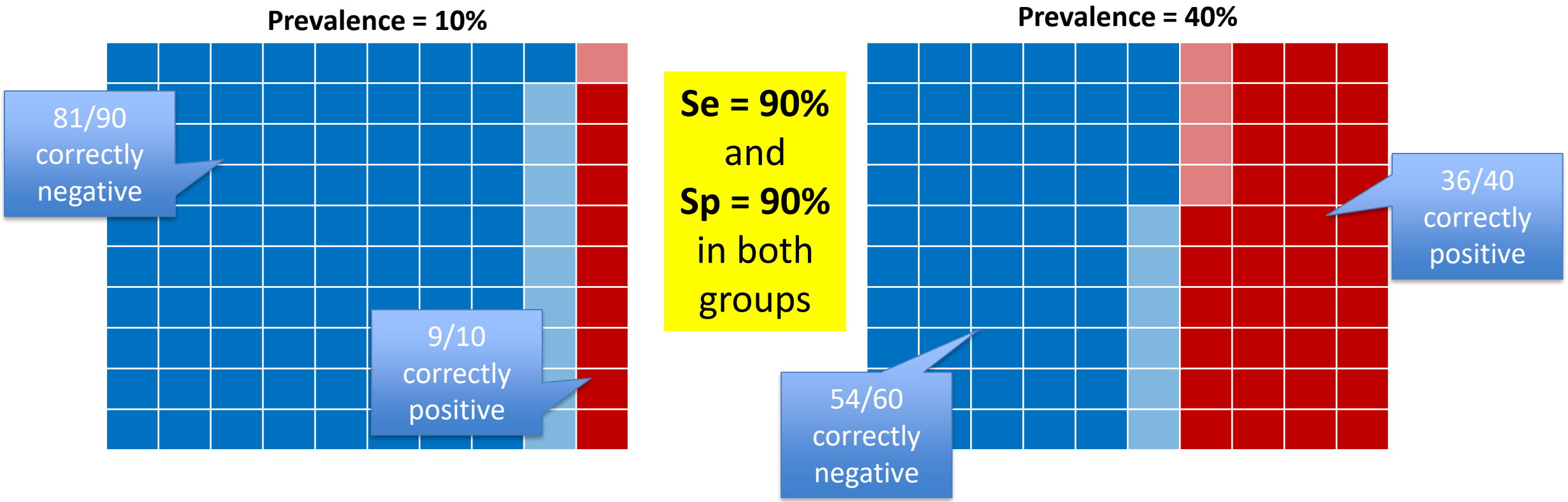
Red = diabetes

Prevalence = 40%



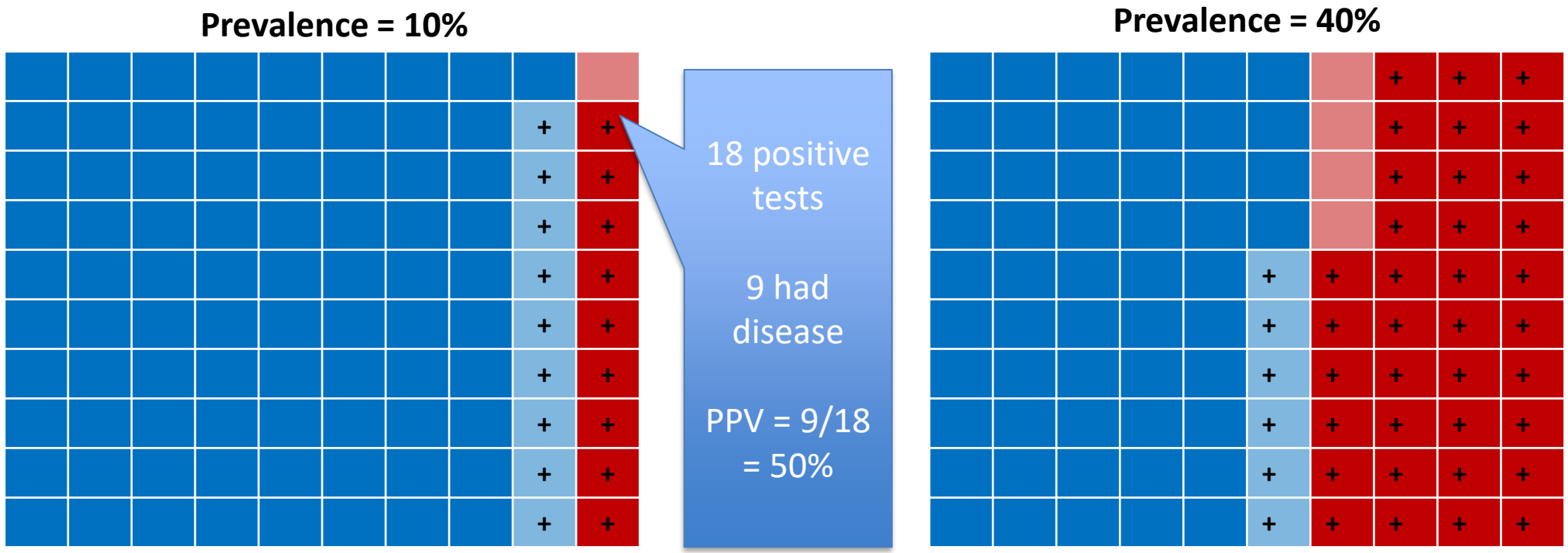
Screening Tests: PPV/NPV and Prevalence

- Say we have a test that correctly identifies diabetics 90% of the time (Se = 90%) and non-diabetics also 90% of the time (Sp =90%)
  - Correct identifications are brightly-colored



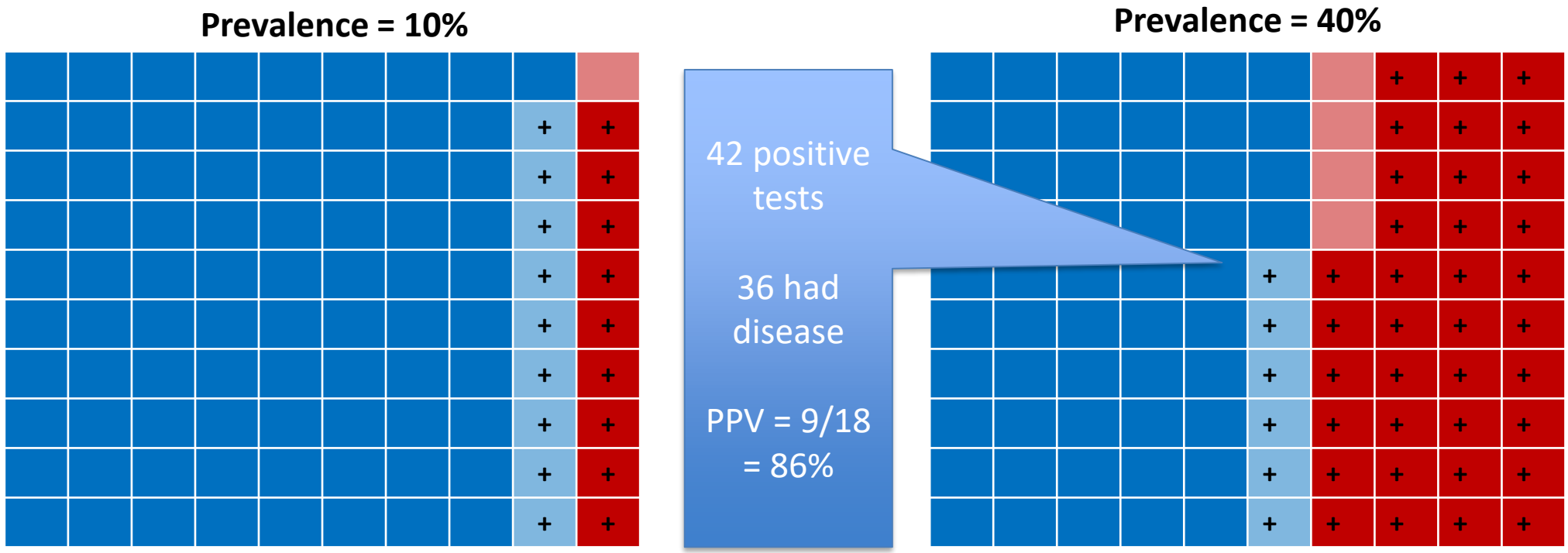
Screening Tests: PPV/NPV and Prevalence

- What is the PPV in both groups?
  - "+" = tested positive, regardless of true status



Screening Tests: PPV/NPV and Prevalence

- What is the PPV in both groups?
  - "+" = tested positive, regardless of true status



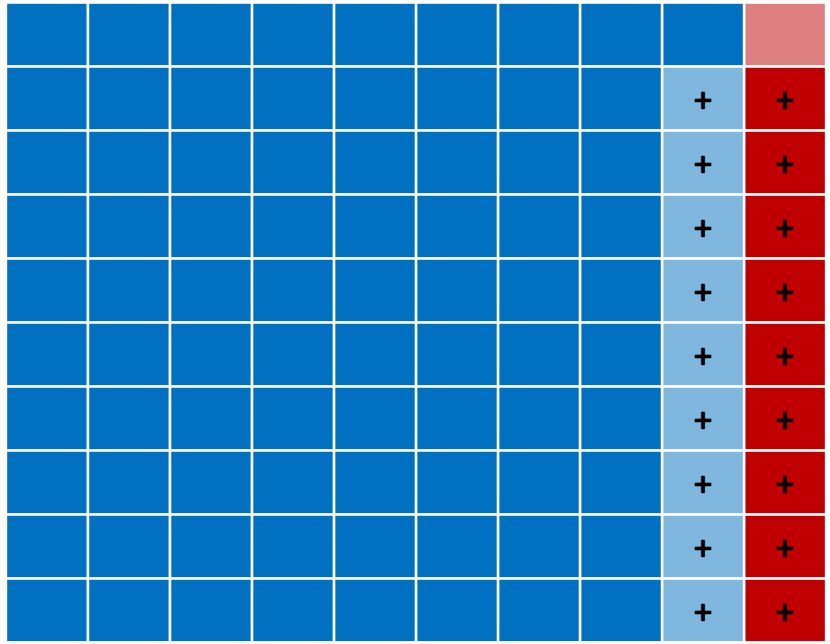
Screening Tests: PPV/NPV and Prevalence

- What is the PPV in both groups?
  - Same test: Se = 90%, Sp = 90%

Regardless of your test result you're just more likely to have diabetes in the higher-risk group...period!

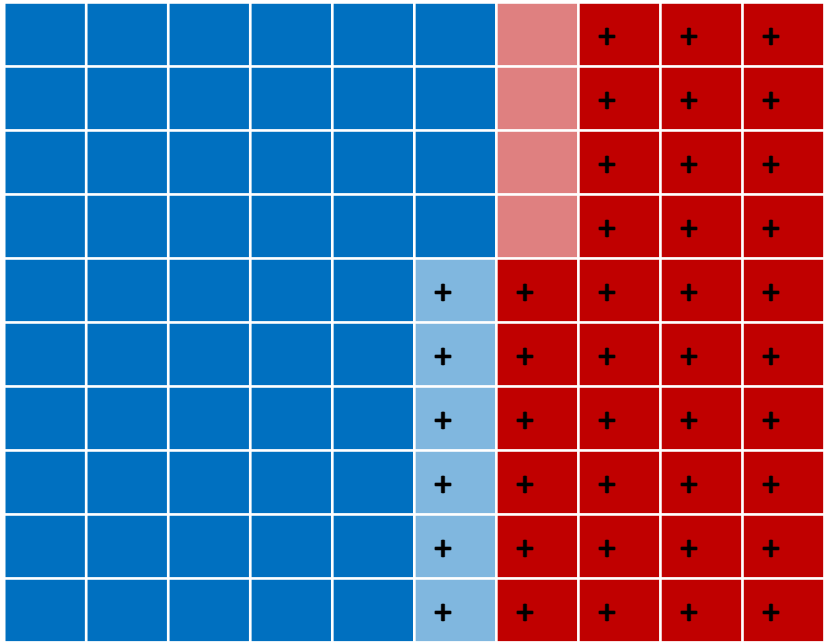
PPV = 50%

Prevalence = 10%



PPV = 86%

Prevalence = 40%



Screening Tests: PPV/NPV and Prevalence

- Here's a more mathematical example

EXAMPLE: SENSITIVITY = 99%, SPECIFICITY = 95%					
Disease Prevalence	**Test Results**	**Sick**	**Not Sick**	**Totals**	**Positive Predictive Value**
1%	+	99	495	594	$\frac{99}{594} = 17\%$
	–	1	9,405	9,406	
	Totals	100	9,900	10,000	

Screening Tests: PPV/NPV and Prevalence

- Here's a more mathematical example

EXAMPLE: SENSITIVITY = 99%, SPECIFICITY = 95%					
Disease Prevalence	**Test Results**	**Sick**	**Not Sick**	**Totals**	**Positive Predictive Value**
1%	+	99	495	594	$\frac{99}{594} = 17\%$
	–	1	9,405	9,406	
	Totals	100	9,900	10,000	
5%	+	495	475	970	$\frac{495}{970} = 51\%$
	–	5	9,025	9,030	
	Totals	500	9,500	10,000	

Screening Tests: PPV/NPV and Prevalence

- Similar but opposite relationship for prevalence and NPV

- **Higher prevalence → higher PPV, lower NPV**
  - Screening programs are "more efficient" (higher % of positive tests are diseased) when directed to high-risk target population
  - Does **NOT** mean screening programs should *only* ever be targeted to high-risk groups
    - Is "ruling out" or "ruling in" a disease more important?