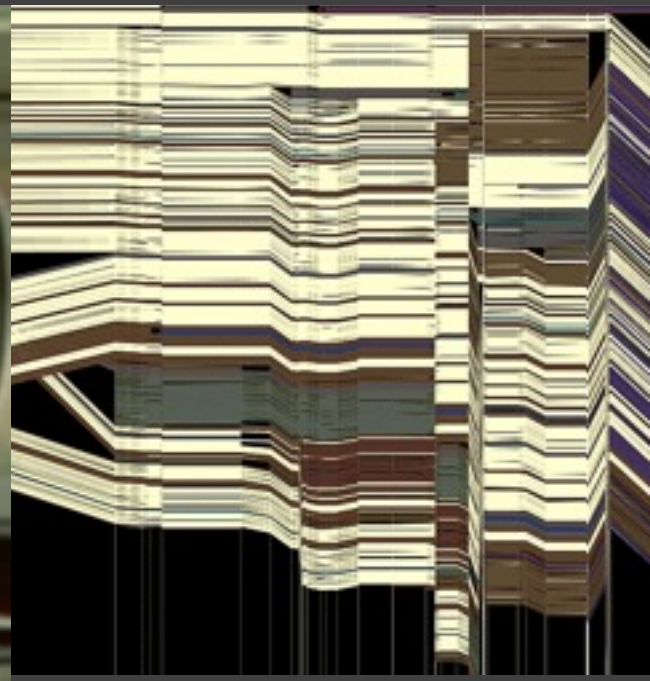
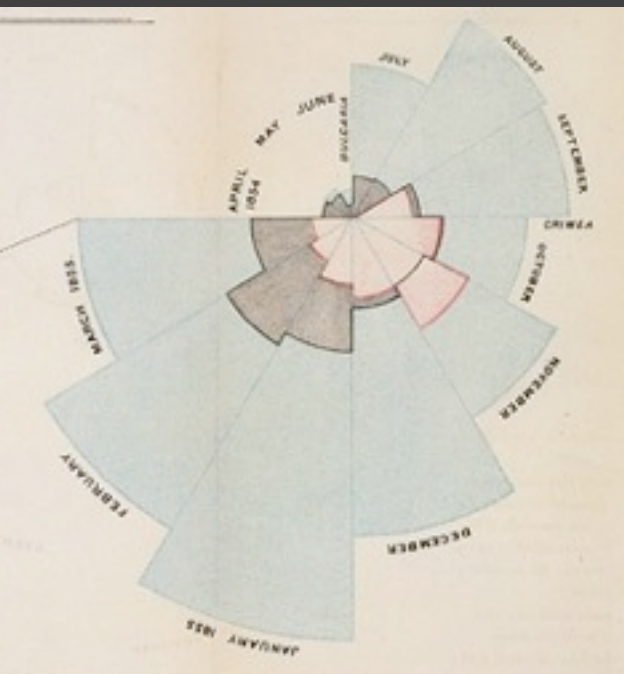


CSE512 :: 14 Jan 2014

Visualization Design



Jeffrey Heer University of Washington

Last Time: Data and Image Models

The Big Picture

task

questions & hypotheses
intended audience

data

physical type
 int, float, etc.
abstract type
 nominal, ordinal, etc.

domain

metadata
semantics
conceptual model

processing
algorithms

mapping
visual encoding
visual metaphor

image

visual channel
perception



Nominal, Ordinal and Quantitative

N - Nominal (labels)

- Operations: =, \neq

O - Ordered (rank-ordered, sorted)

- Operations: =, \neq , $<$, $>$

Q - Interval (location of zero arbitrary)

- Operations: =, \neq , $<$, $>$, -
- Can measure distances or spans

Q - Ratio (zero fixed)

- Operations: =, \neq , $<$, $>$, -, $\%$
- Can measure ratios or proportions

Visual Encoding Variables

Position

Size

Value

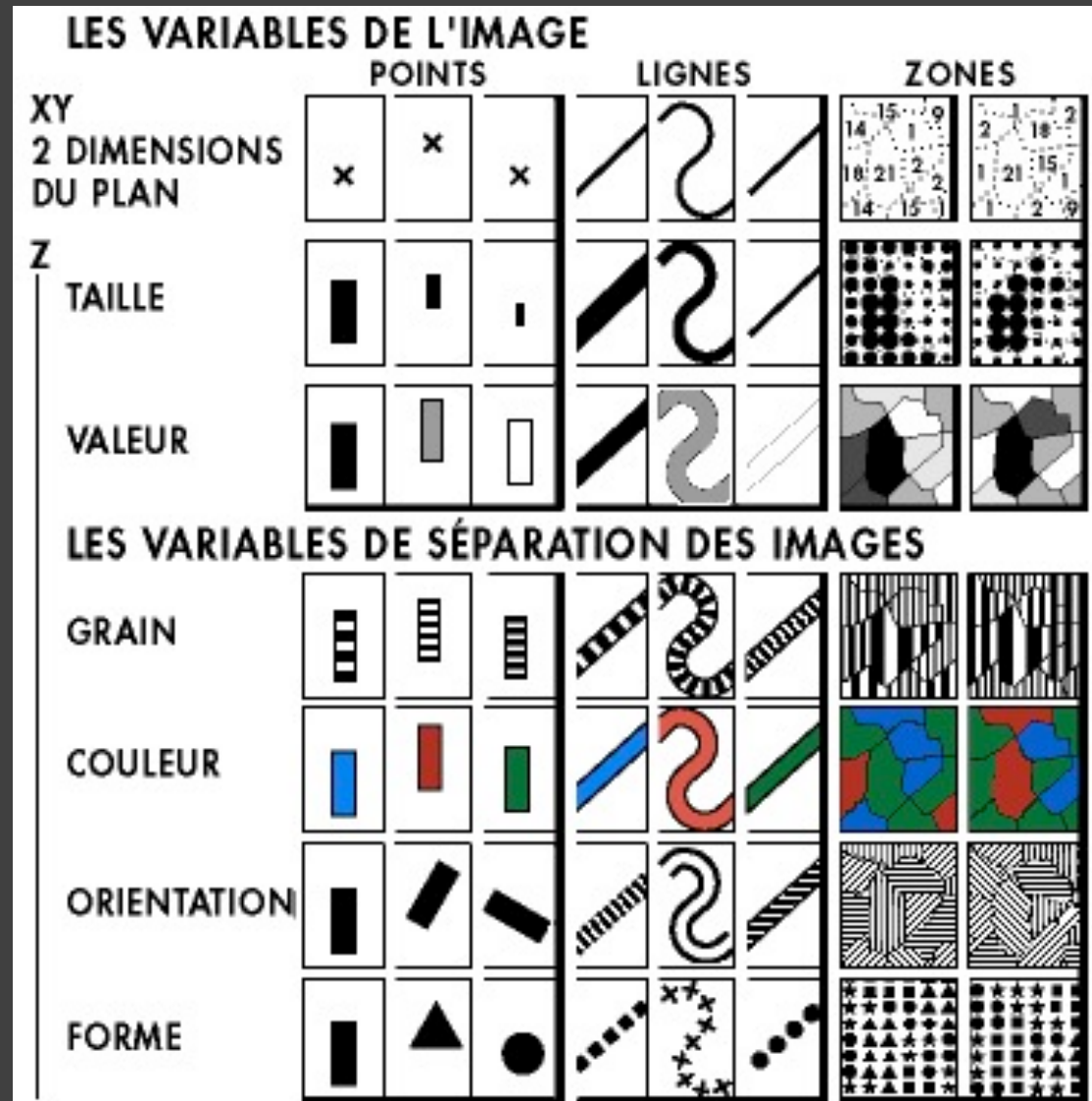
Texture

Color

Orientation

Shape

Others?



Formalizing Design

(Mackinlay 1986)

Choosing Visual Encodings

Challenge:

Assume 8 visual encodings and n data attributes. We would like to pick the “best” encoding among a combinatorial set of possibilities with size $(n+1)^8$

Principle of Consistency:

The properties of the image (visual variables) should match the properties of the data.

Principle of Importance Ordering:

Encode the most important information in the most effective way.

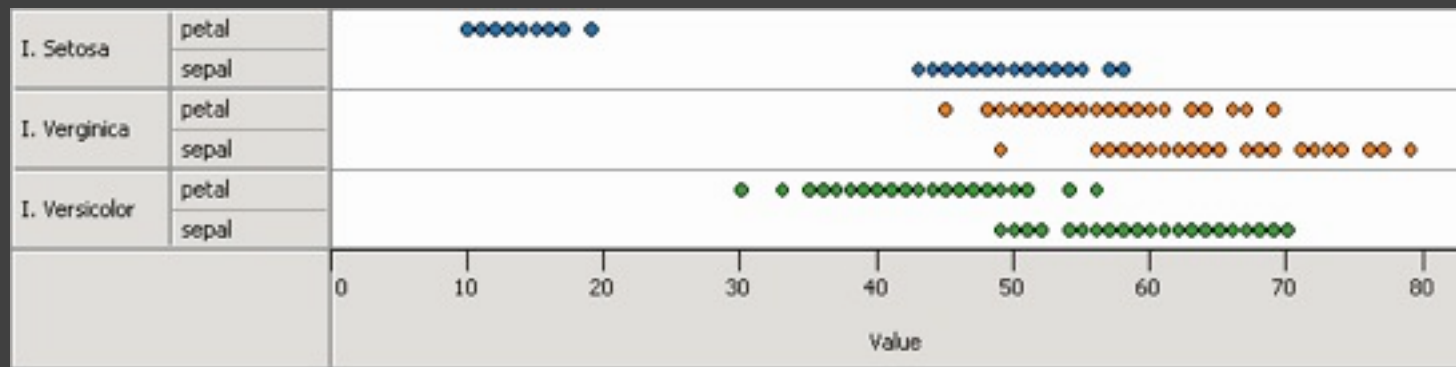
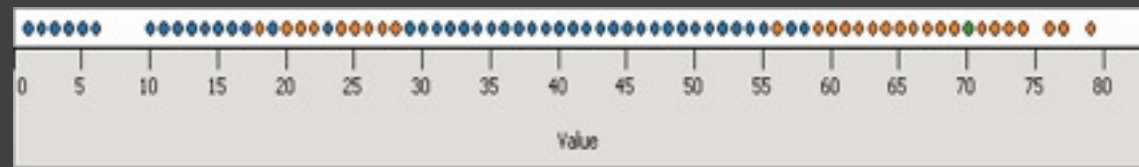
Design Criteria (Mackinlay)

Expressiveness

A set of facts is expressible in a visual language if the sentences (i.e. the visualizations) in the language express all the facts in the set of data, and only the facts in the data.

Cannot express the facts

A one-to-many ($1 \rightarrow N$) relation cannot be expressed in a single horizontal dot plot because multiple tuples are mapped to the same position



Expresses facts not in the data

A length is interpreted as a quantitative value;
∴ Length of bar says something untrue about N data

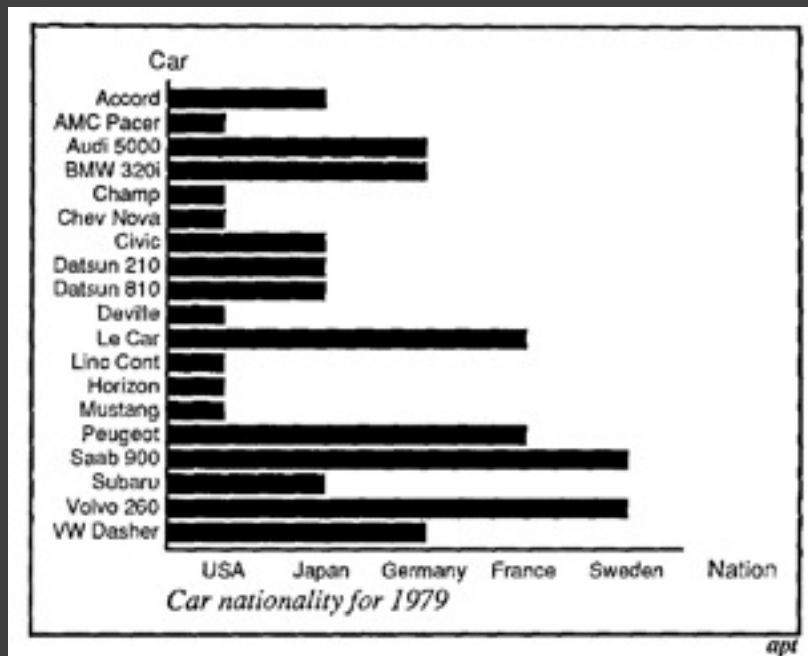


Fig. 11. Incorrect use of a bar chart for the *Nation* relation. The lengths of the bars suggest an ordering on the vertical axis, as if the USA cars were longer or better than the other cars, which is not true for the *Nation* relation.

[Mackinlay, APT, 1986]

Design Criteria (Mackinlay)

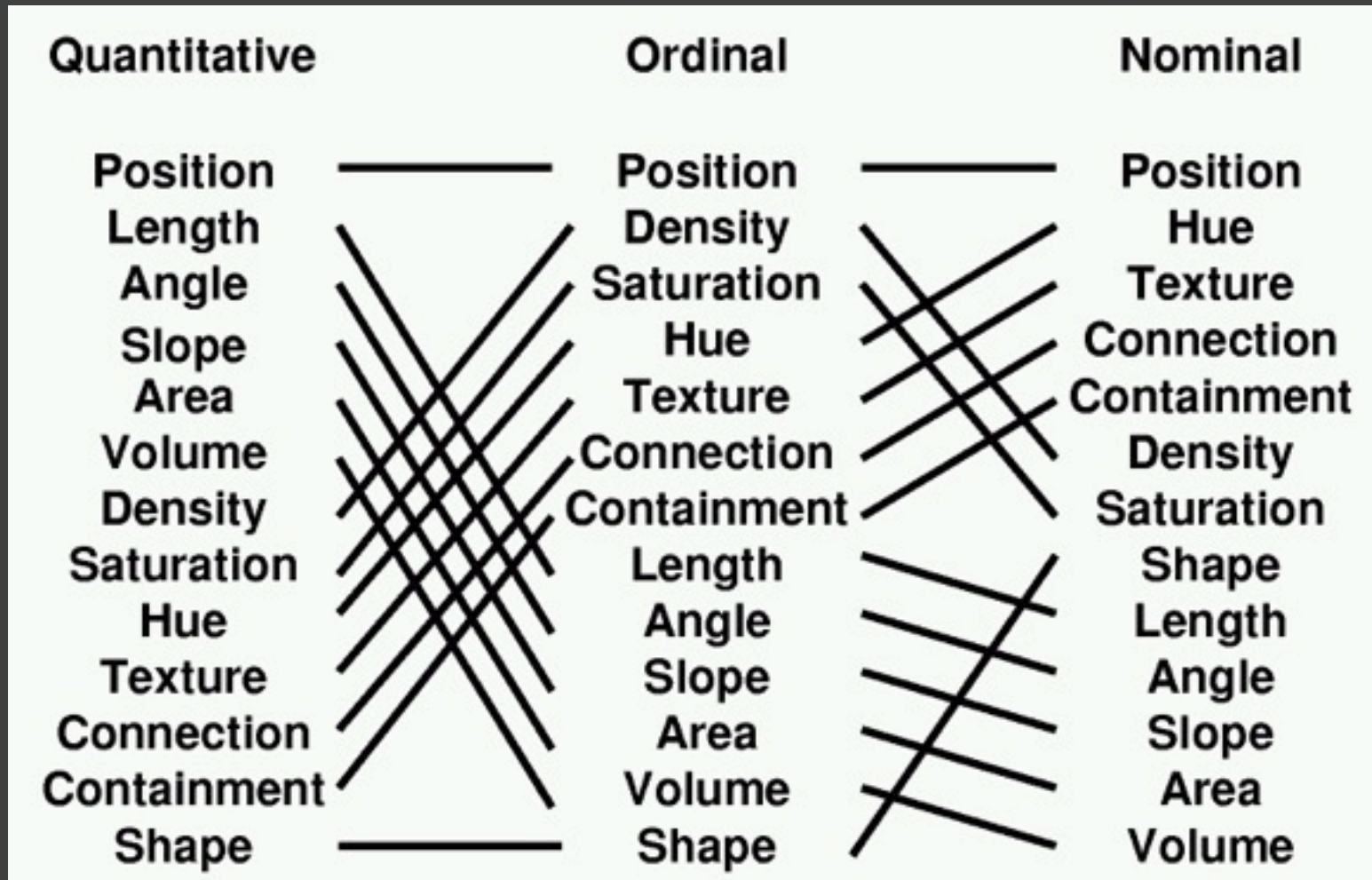
Expressiveness

A set of facts is expressible in a visual language if the sentences (i.e. the visualizations) in the language express all the facts in the set of data, and only the facts in the data.

Effectiveness

A visualization is more effective than another visualization if the information conveyed by one visualization is more readily perceived than the information in the other visualization.

Mackinlay's Ranking



Conjectured *effectiveness* of the encoding

Mackinlay's Design Algorithm

User formally specifies data model and type

- Additional input: ordered list of data variables to show

APT searches over design space

- Tests expressiveness of each visual encoding
- Generates specification for encodings that pass test
- Tests perceptual effectiveness of resulting image

Outputs the “most effective” visualization

Limitations

Does not cover many visualization techniques

- Bertin and others discuss networks, maps, diagrams
- Does not consider 3D, animation, illustration, photography, ...

Does not model interaction

Does not consider semantic data types / conventions

Assignment 1 Review

Design Considerations

Title, labels, legend, captions, source!

Expressiveness and Effectiveness

Avoid unexpressive marks (lines? bars? gradients?)

Use perceptually effective encodings

Don't distract: faint gridlines, pastel highlights/fills

The “elimination diet” approach – start minimal

Support comparison and pattern perception

Between elements, to a reference line, or to totals

Design Considerations

Group / sort data by meaningful dimensions

Transform data (e.g., invert, log, normalize)

Are model choices (regression lines) appropriate?

Reduce cognitive overhead

Minimize visual search, minimize ambiguity

-> Avoid legend lookups if direct labeling works

-> Avoid color mappings with indiscernible colors

Be consistent! **Visual inferences** should consistently support **data inferences**

In-Class Review Rubric

Expressiveness

- Prioritizes important information / Avoids false inferences
- Consistent visual mappings (e.g., respect color mappings)
- Make encodings *meaningful* rather than arbitrary

Effectiveness

- Facilitates accurate decoding / Minimizes cognitive overhead
- Highlight elements of primary interest

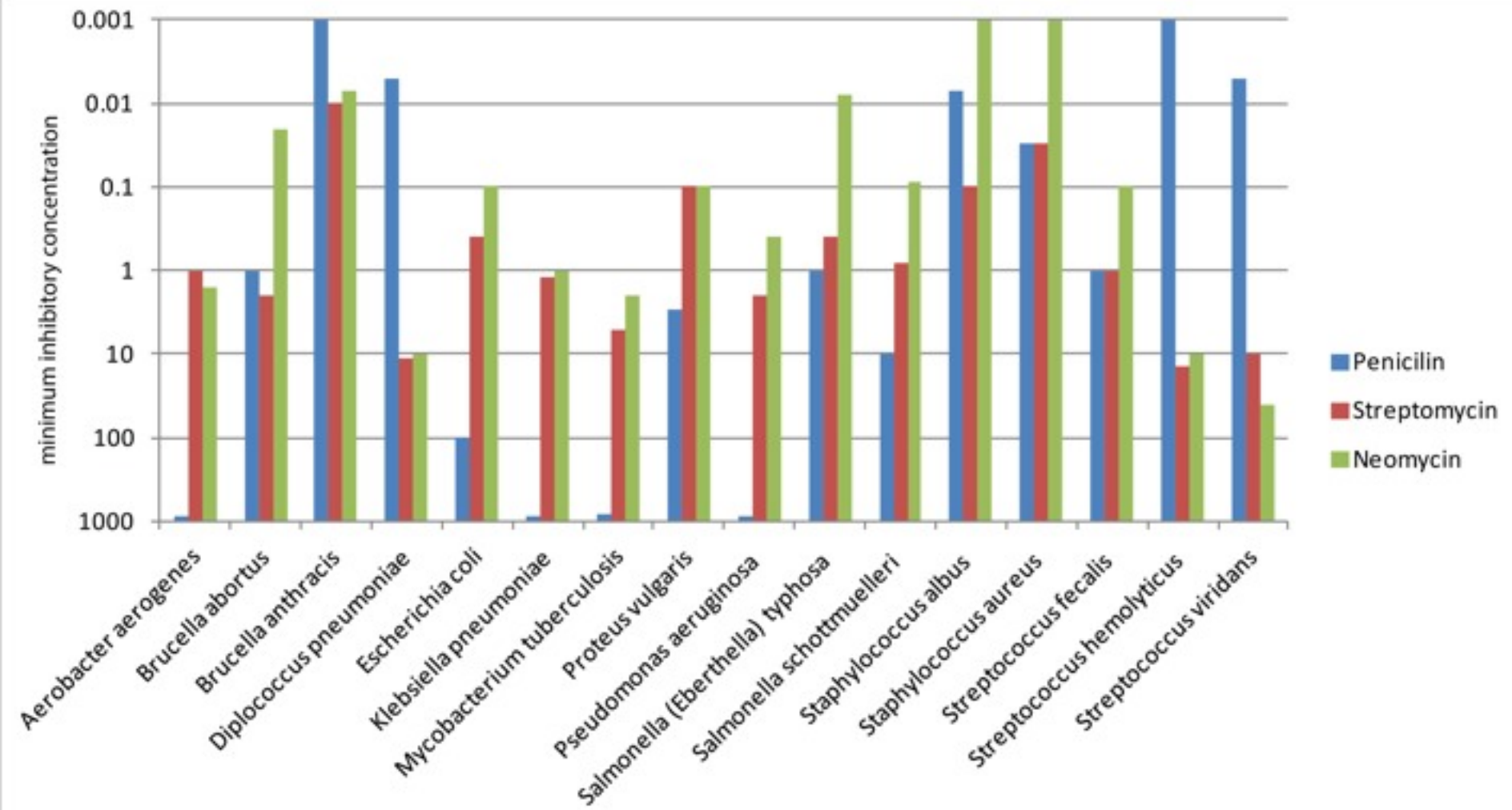
Grouping / Sorting

Data Transformation

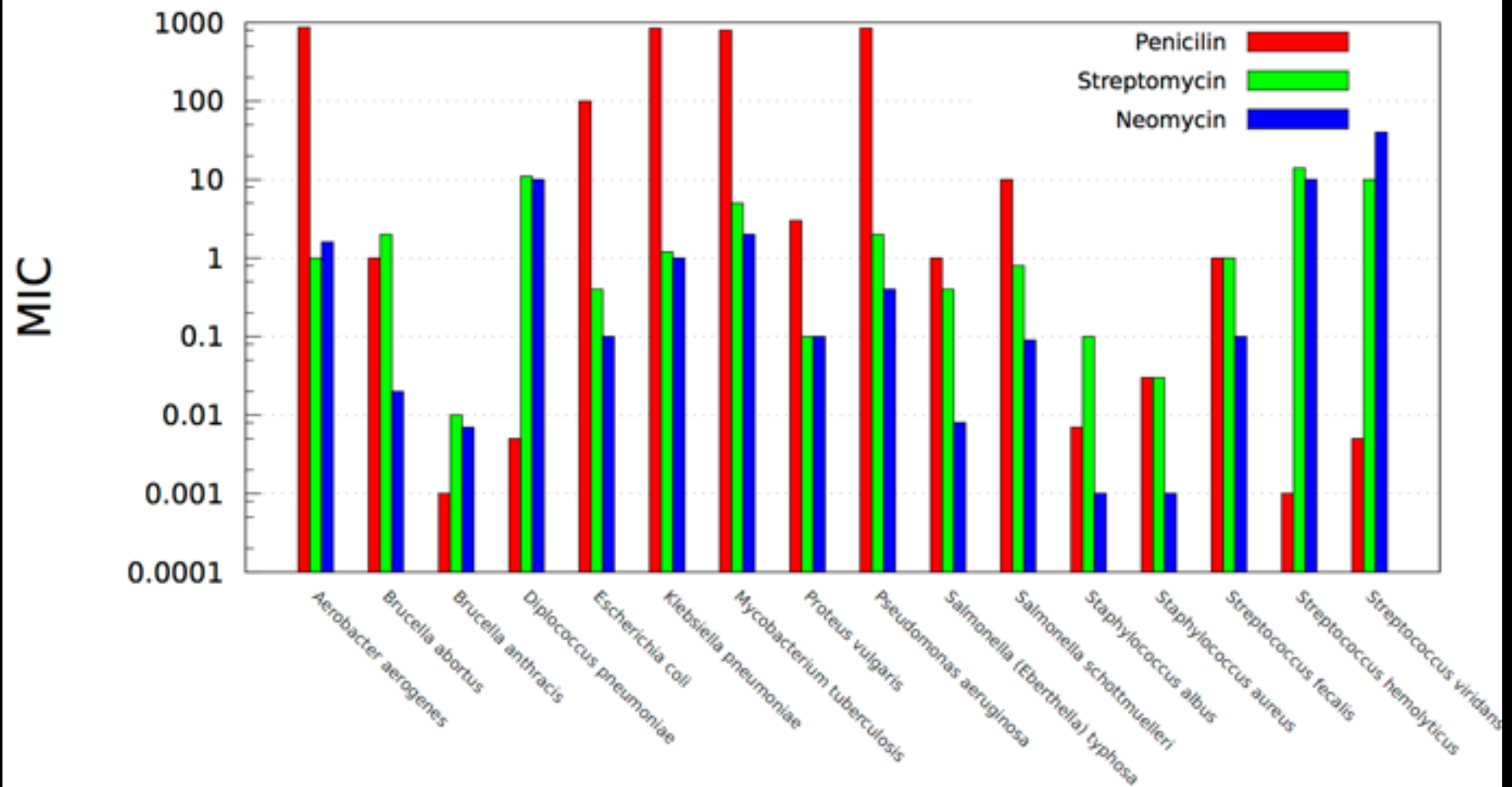
Non-Data Elements

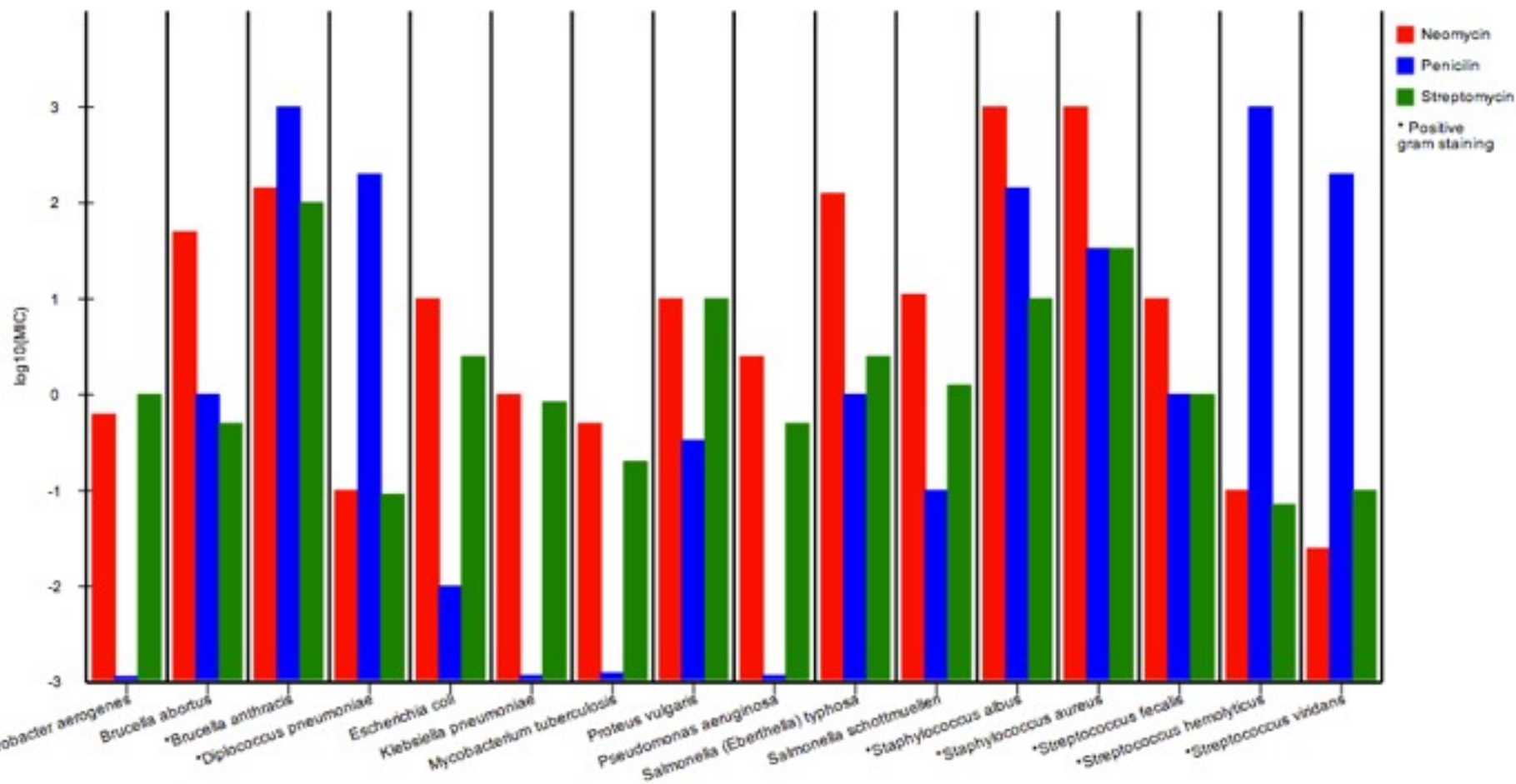
- Descriptive: Title, Label, Caption, Data Source, Annotations
- Reference: Gridlines, Legend

Bar Charts

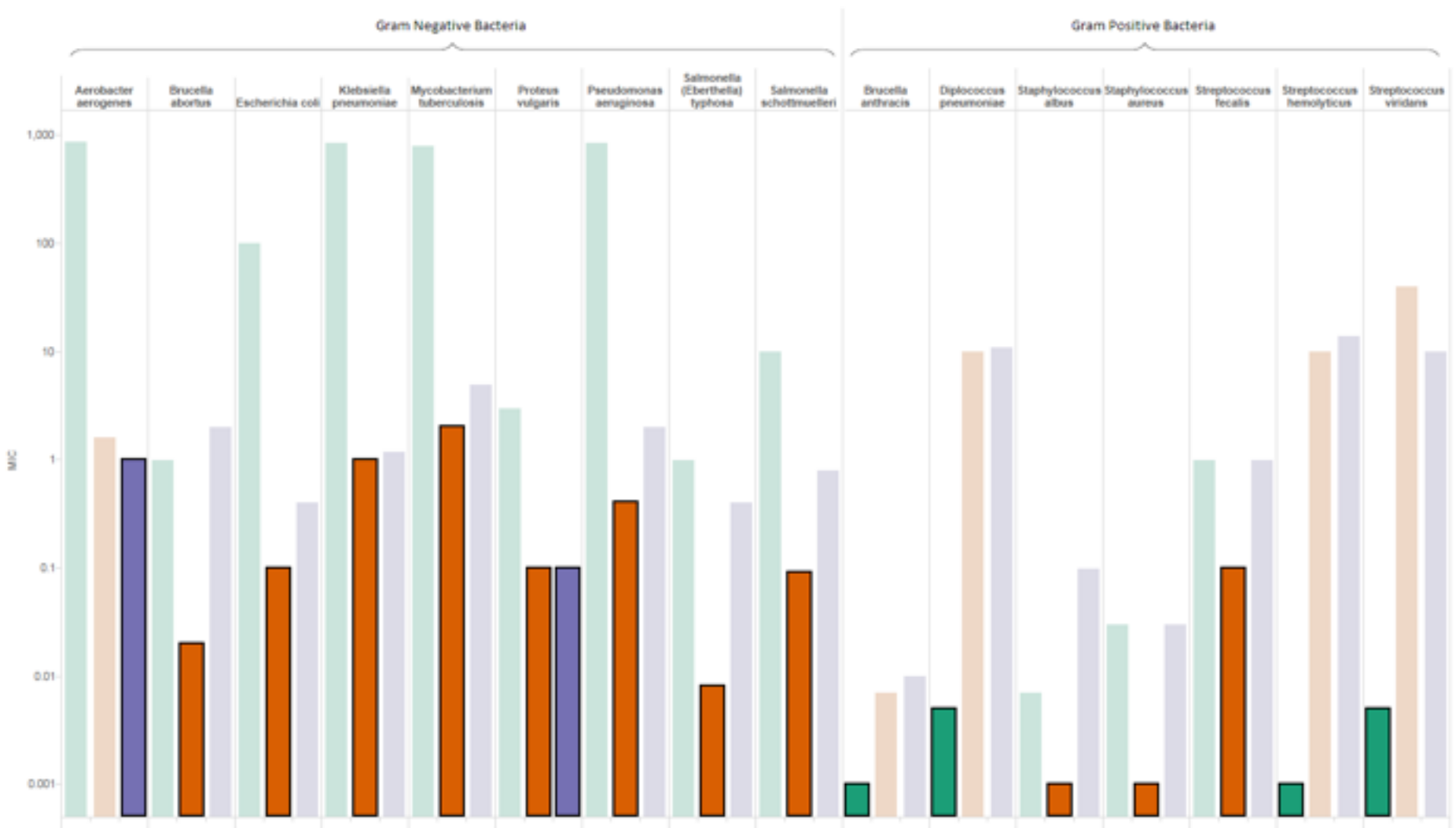


MIC for Penicilin, Streptomycin and Neomycin on 16 bacteria



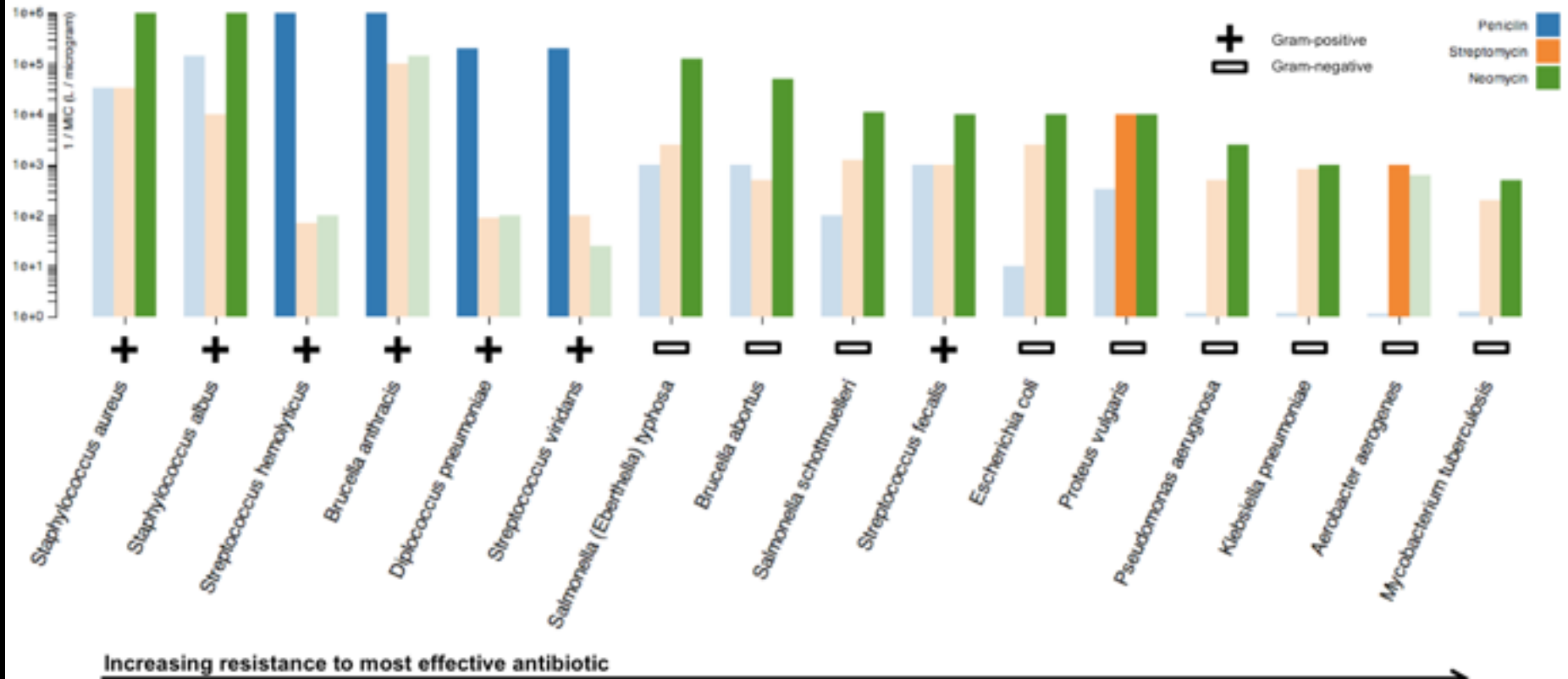


Efficacy of Penicillin (■), Neomycin (■) and Streptomycin (■) against various bacteria

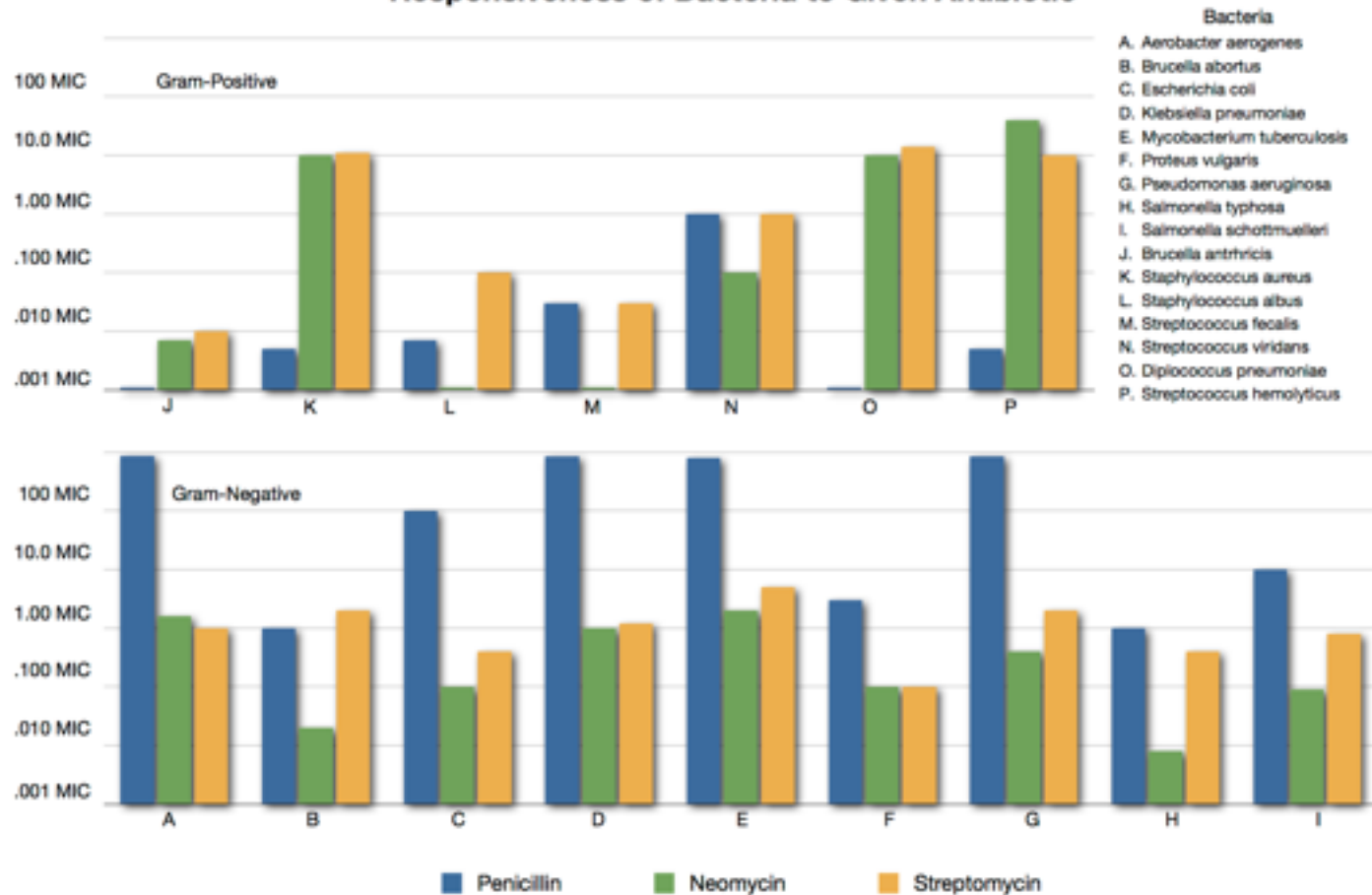


MICs (note log scale) of Penicillin, Neomycin and Streptomycin are shown for various strains of bacteria. Most effective antibiotic is highlighted for each bacterium.

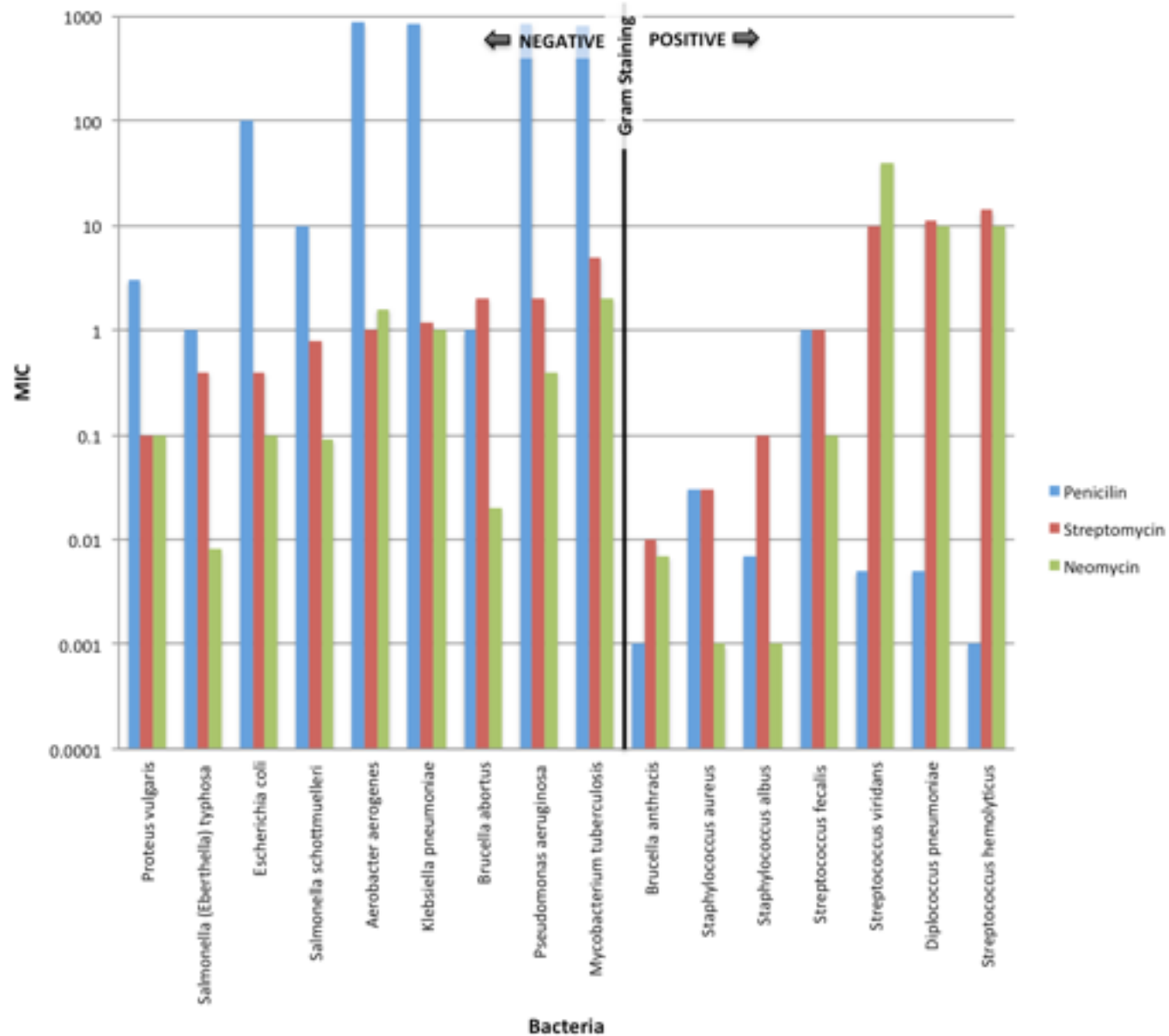
Antibiotic Effectiveness



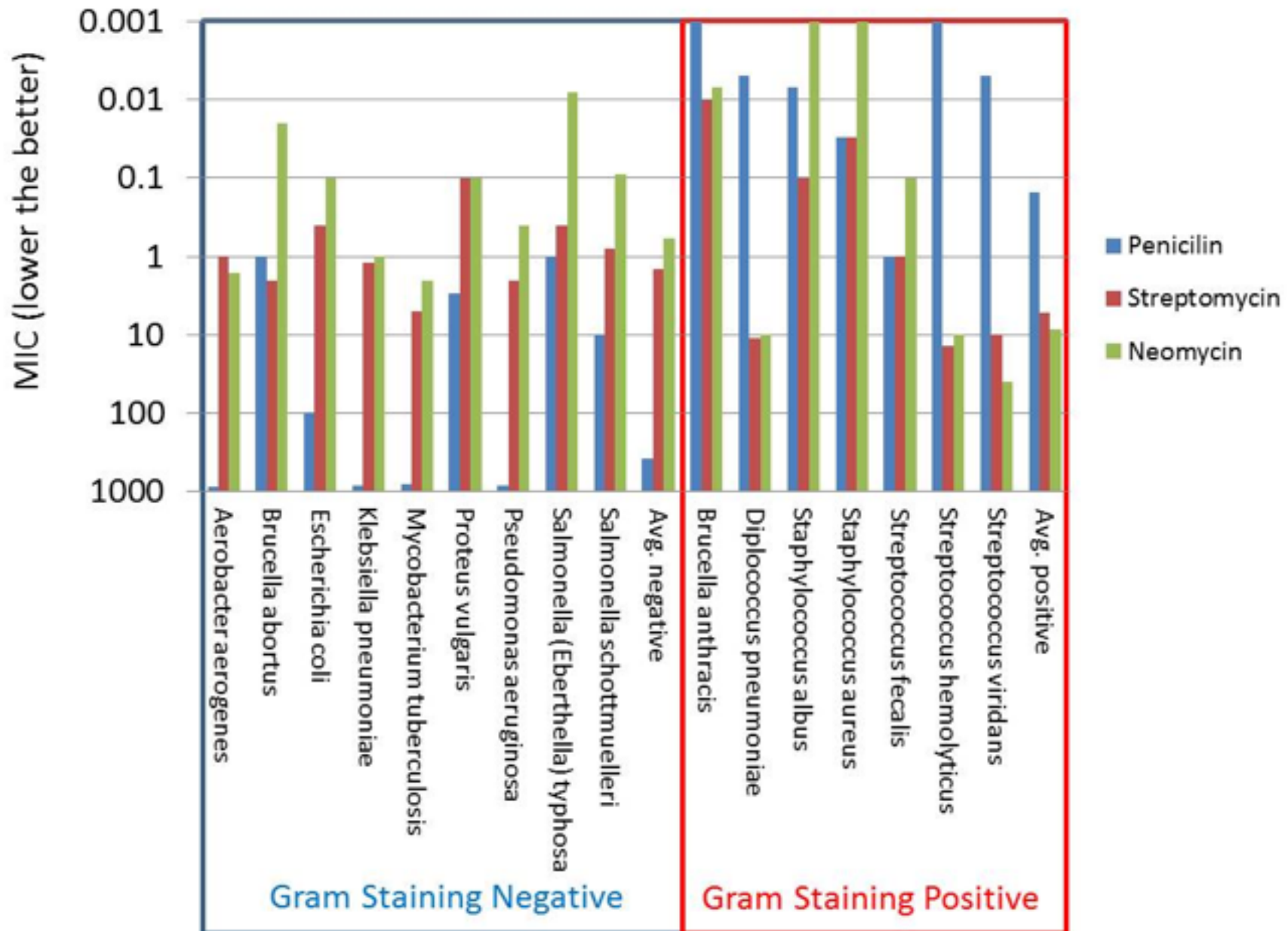
Responsiveness of Bacteria to Given Antibiotic



Effectiveness of Antibiotics on Bacteria



Performance of Antibiotics on 16 bacteria

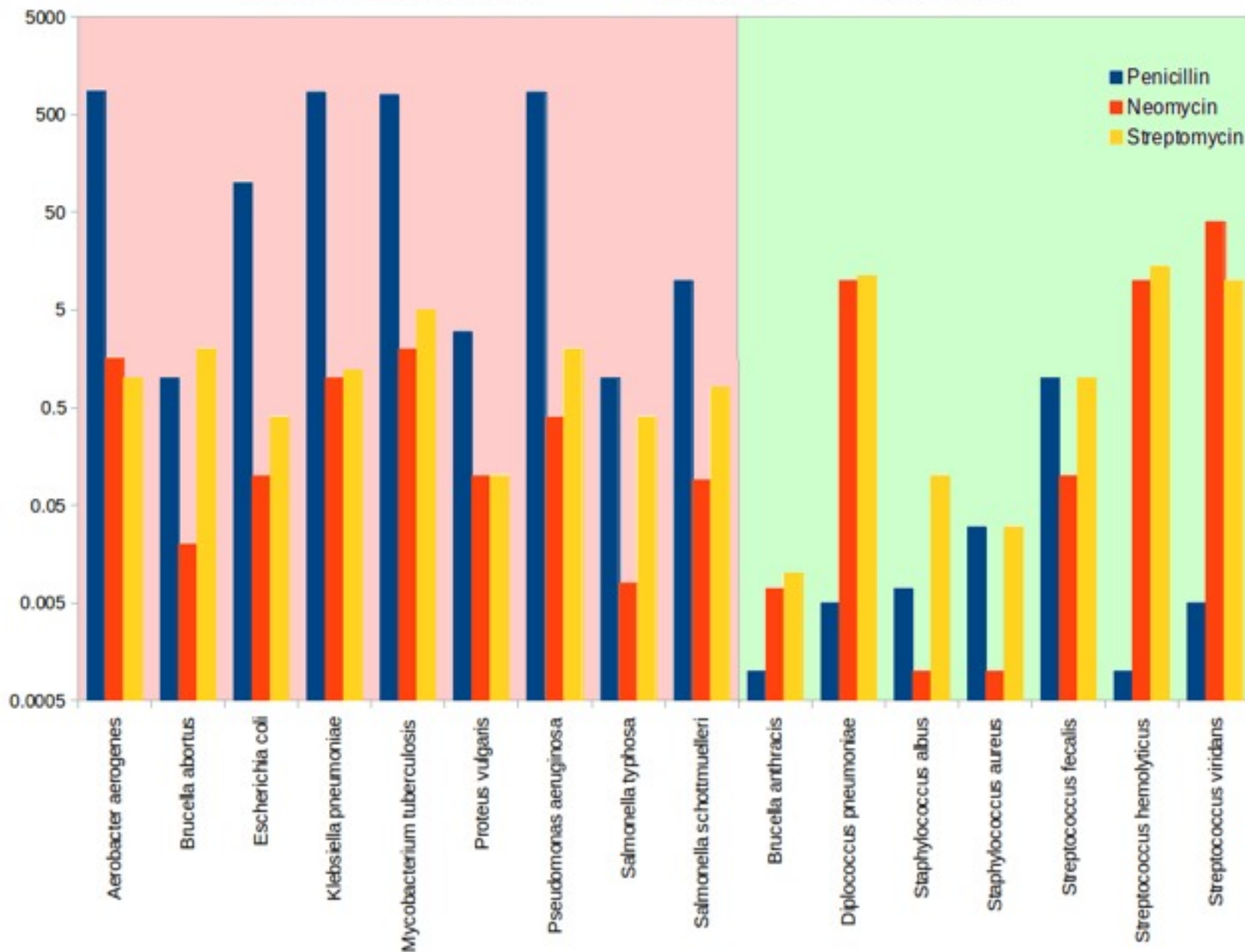


Performance of Wonder Drugs

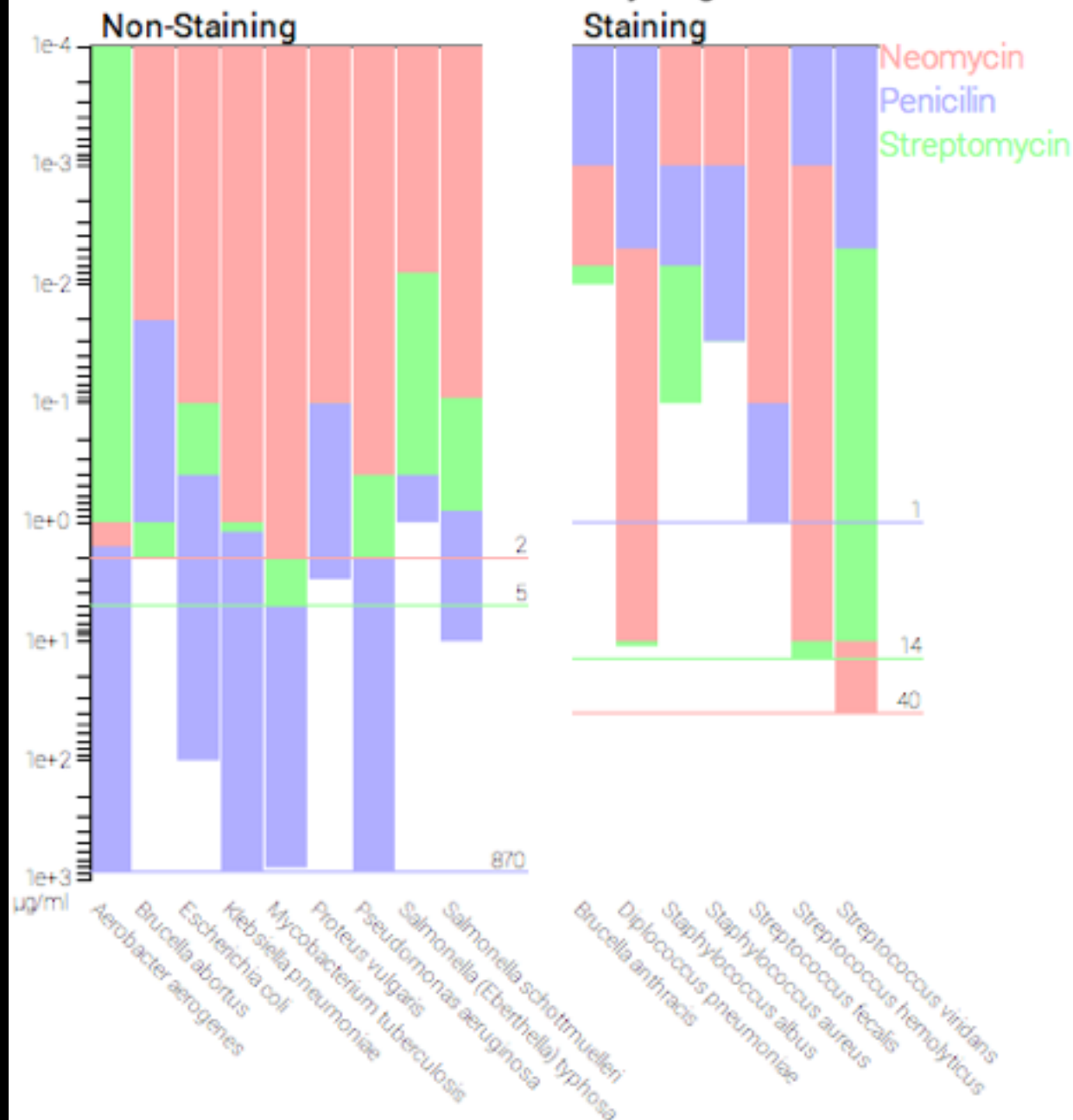
Gram Positive

Gram Negative

MIC (Concentration required to neutralize bacteria)



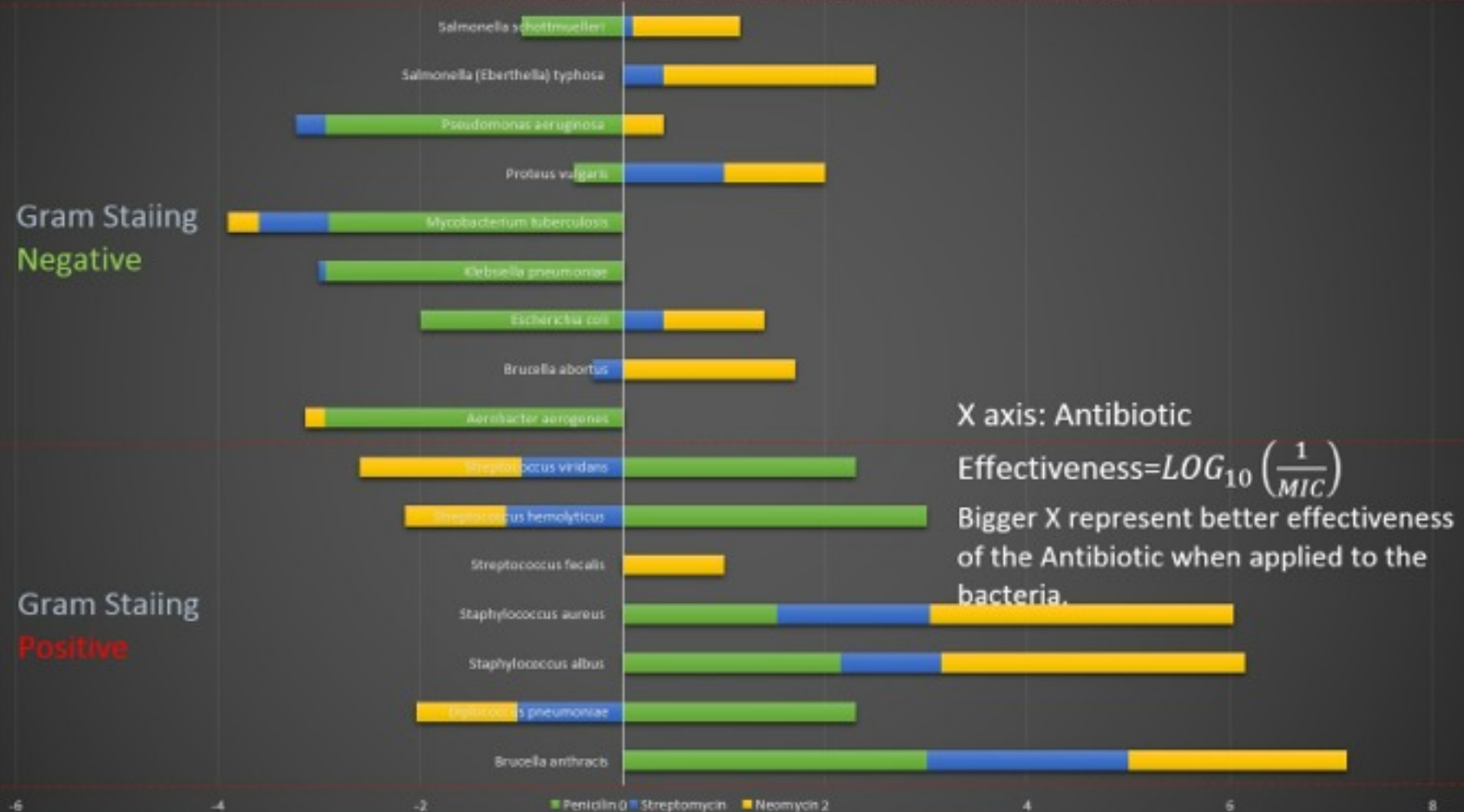
Concentrations of Antibiotics necessary to fight common bacteria

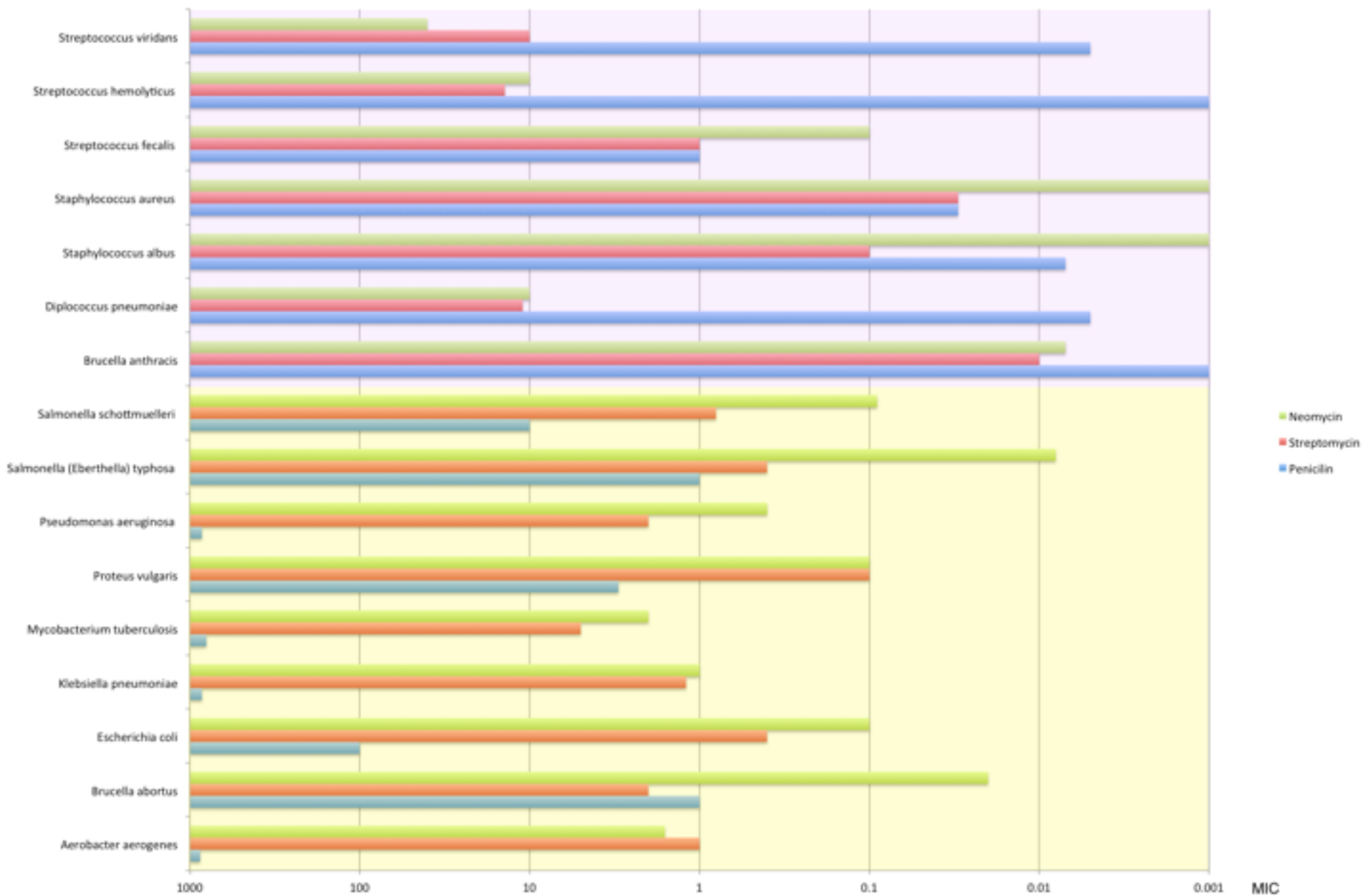


Antibiotic Effectiveness Plot

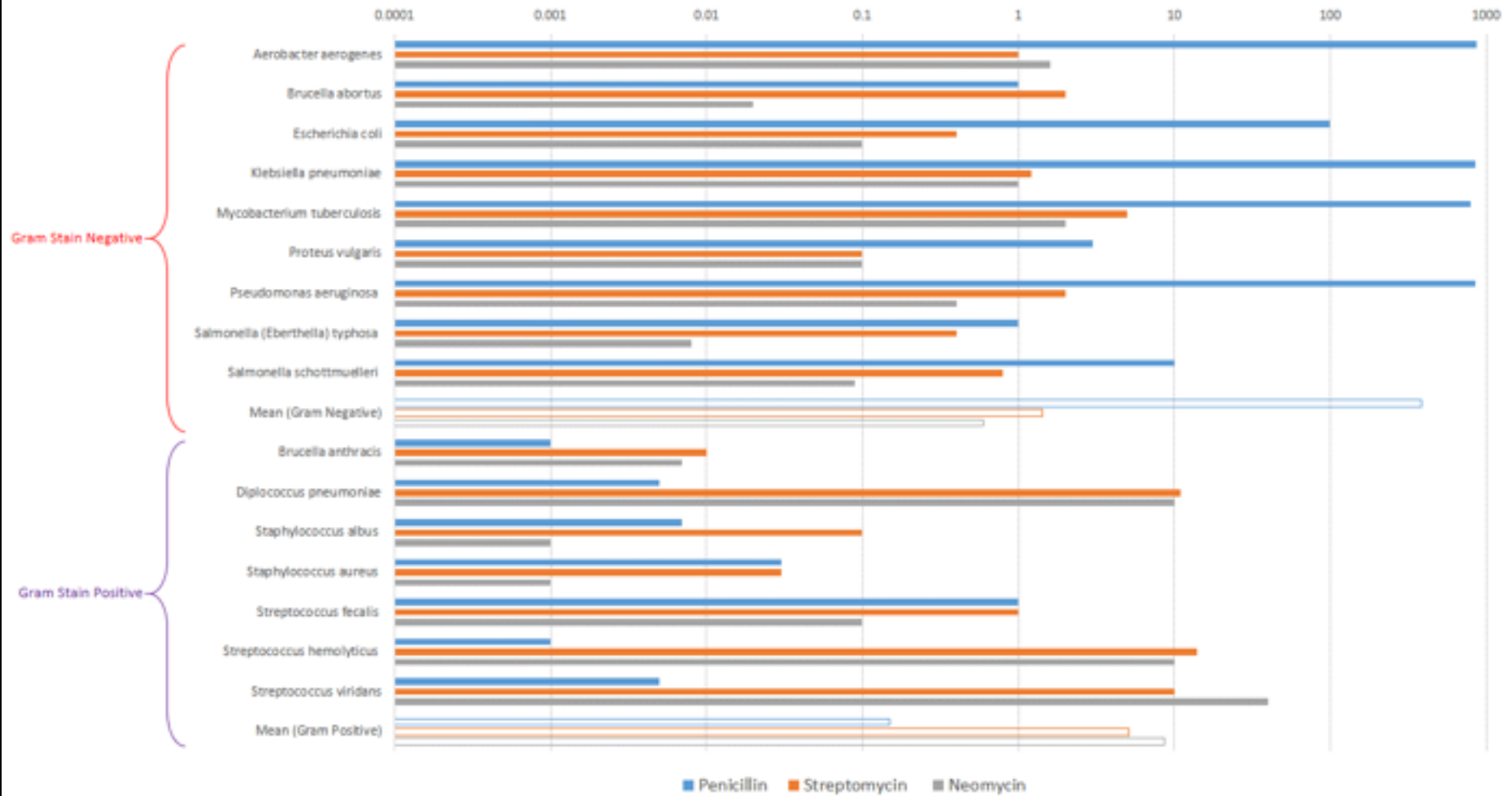
Gram Staining
Negative

Gram Staining
Positive

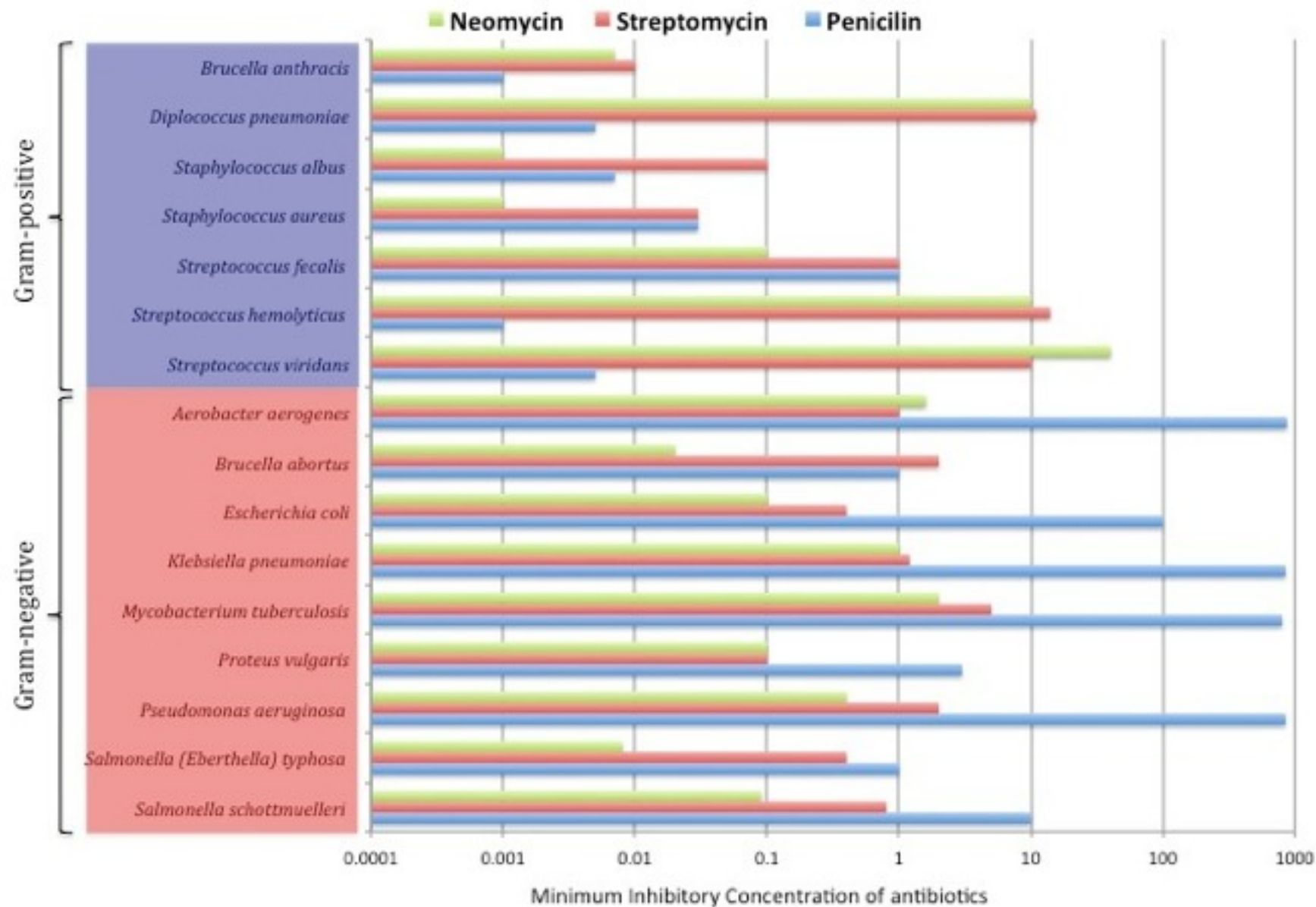




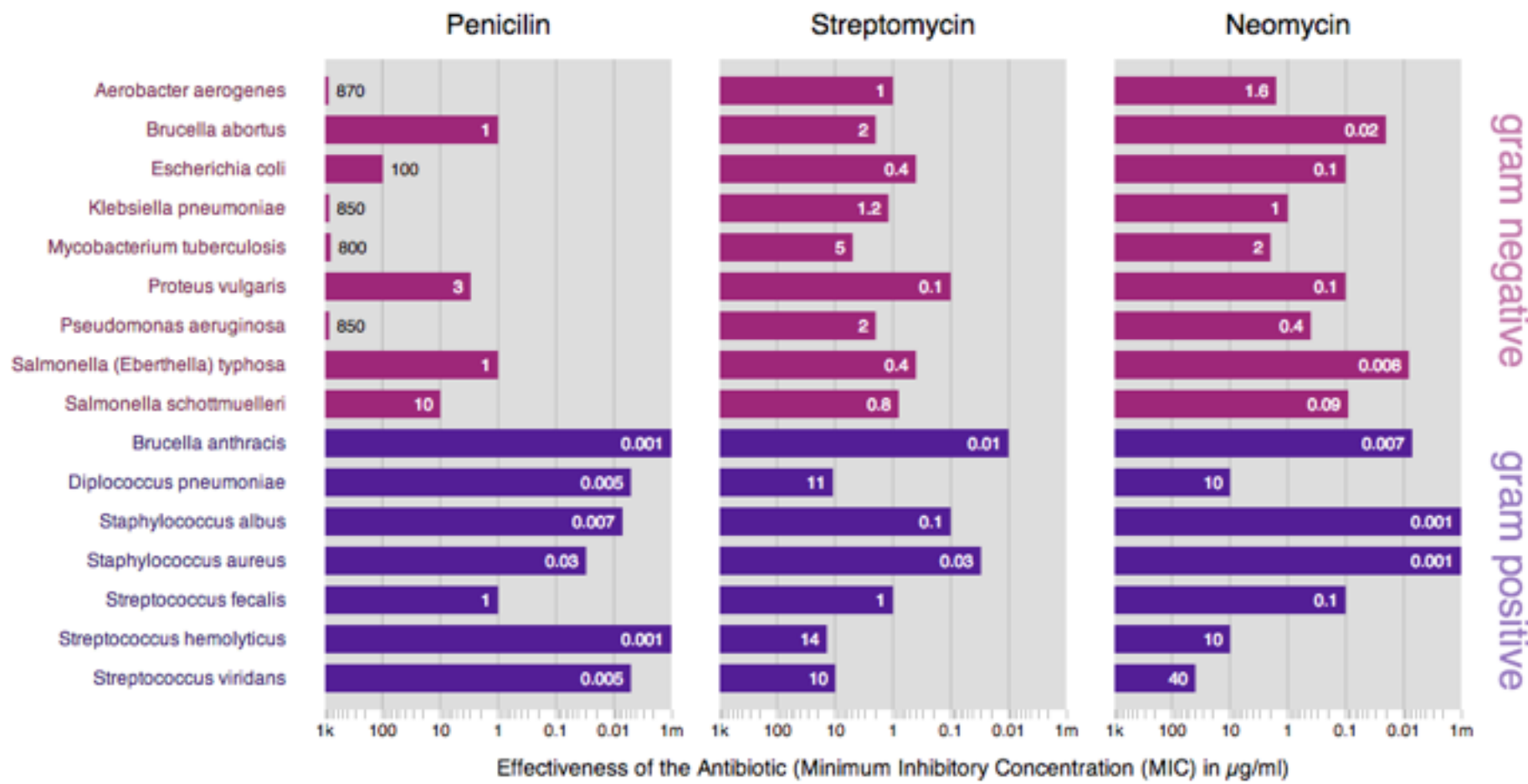
Minimum Inhibitory Concentration of Antibiotics (µg/ml)



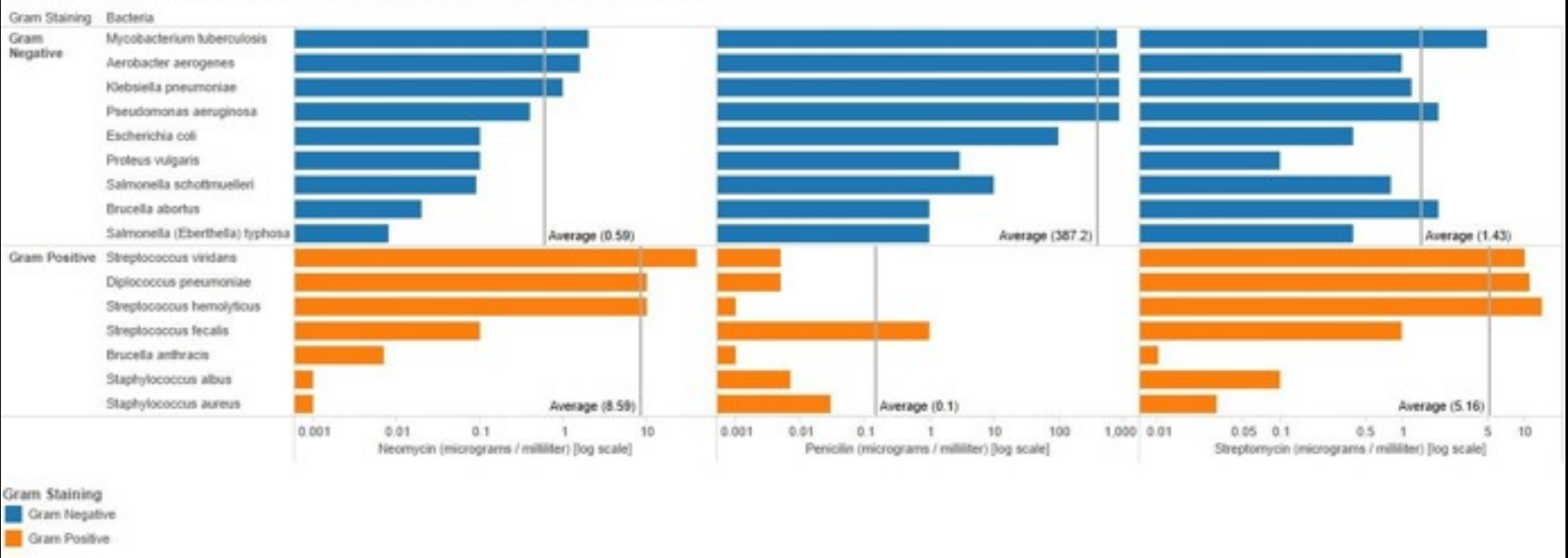
Effectiveness of three antibiotics against sixteen types of bacteria



Burtin's Antibiotics



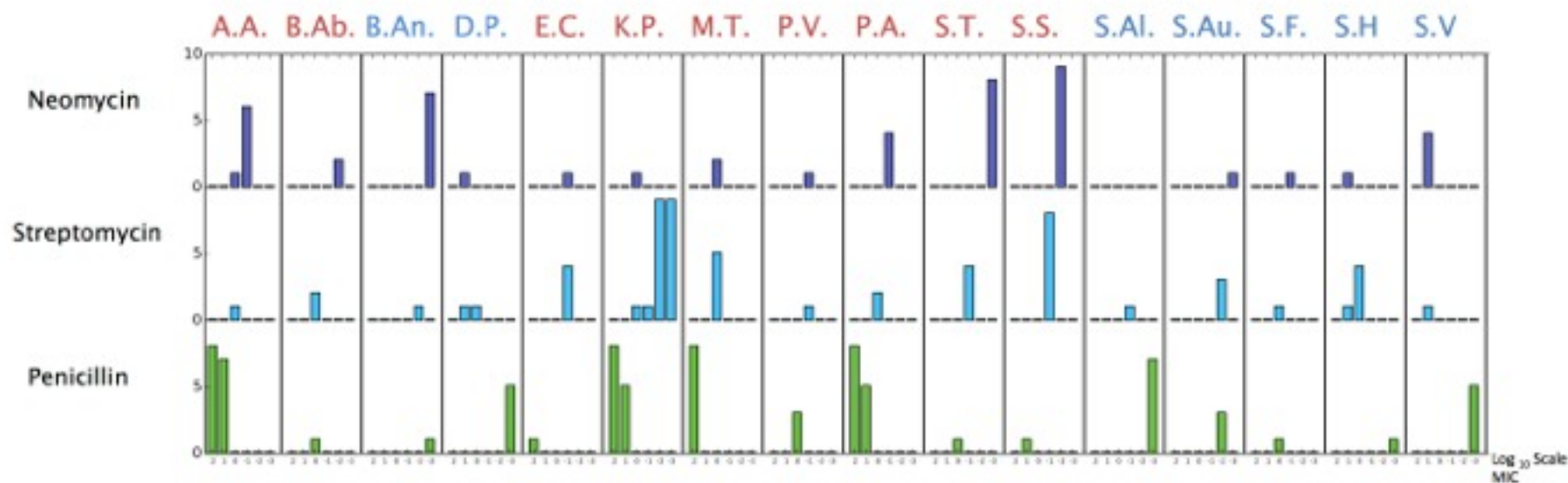
Effectiveness of 3 Popular Antibiotics Against Gram Negative/Postive Bacteria



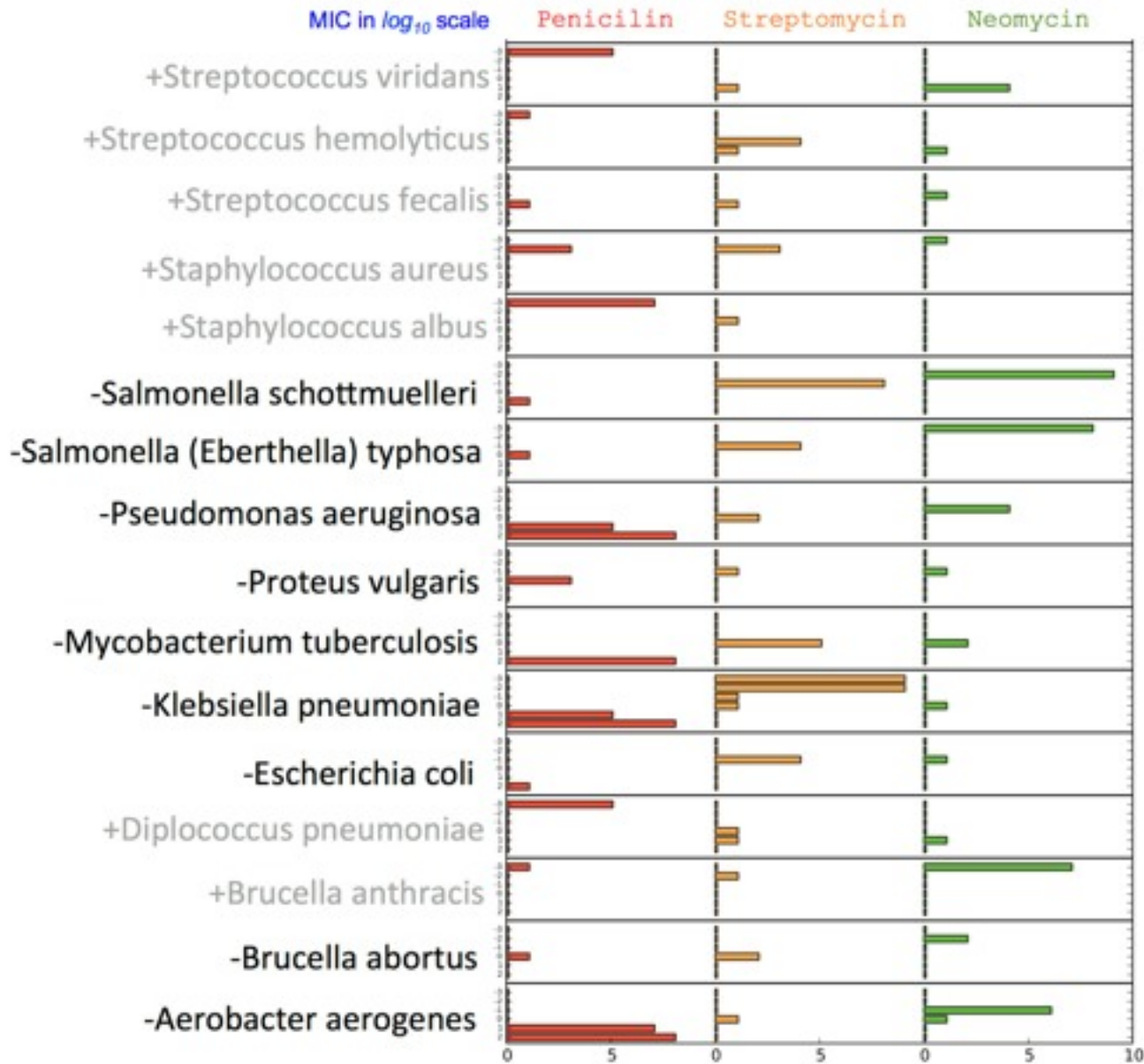
The Effectiveness of Antibiotics on Various Bacteria

Abbreviation	Bacteria
A.A.	Aerobacter aerogenes
B.Ab.	Brucella abortus
B.An.	Brucella anthracis
D.P.	Diplococcus pneumoniae
E.C.	Escherichia coli
K.P.	Klebsiella pneumoniae
M.T.	Mycobacterium tuberculosis
P.V.	Proteus vulgaris
P.A.	Pseudomonas aeruginosa
S.T.	Salmonella (Eberthella) typhosa
S.S.	Salmonella schottmuelleri
S.Al.	Staphylococcus albus
S.Au.	Staphylococcus aureus
S.F.	Streptococcus fecalis
S.H.	Streptococcus hemolyticus
S.V.	Streptococcus viridans

■ Gram Staining Negative
■ Gram Staining Positive

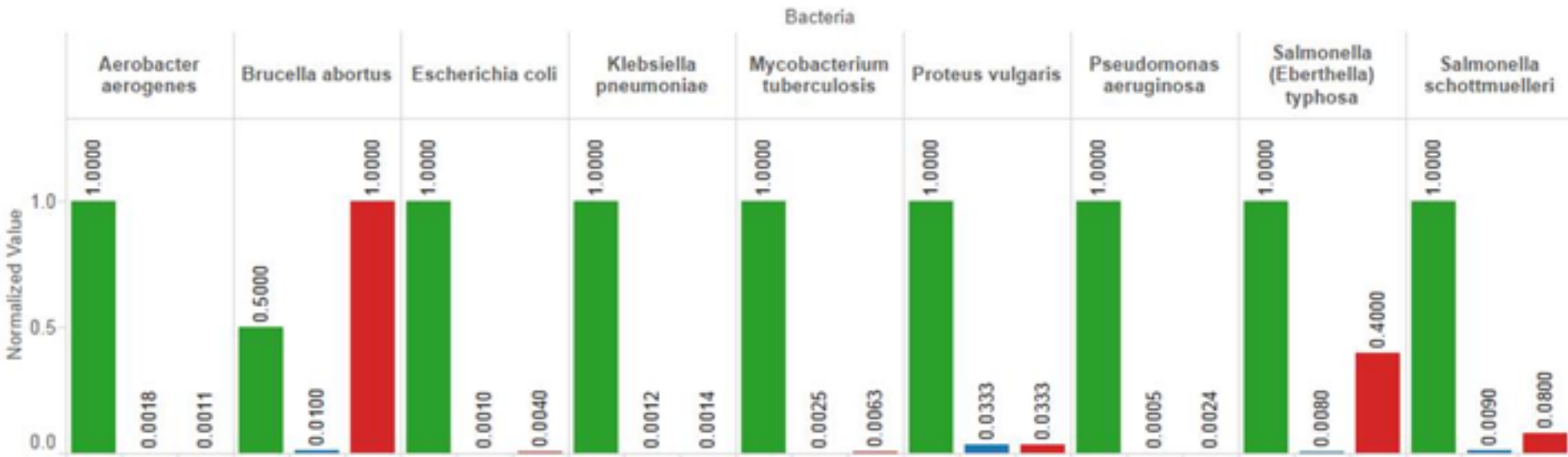


Effectiveness of 3 Antibiotics to 16 Bacterias

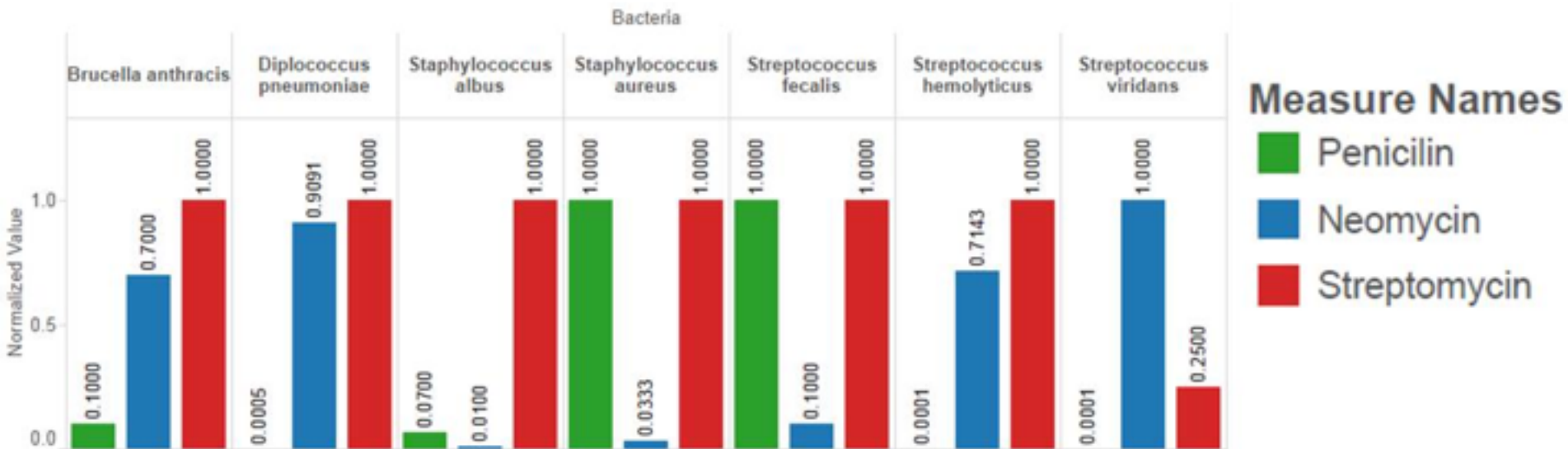


Gram Staining
 + positive
 - negative

Gram Staining / negative



Gram Staining / positive



Measure Names

- Penicilin
- Neomycin
- Streptomycin

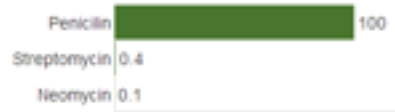
Minimum inhibitory concentration (MIC) normalized by the largest value of that bacteria

MIC of antibiotics on various bacteria

Aerobacter aerogenes



Escherichia coli



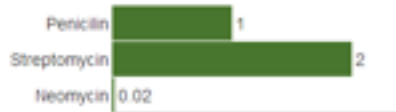
Pseudomonas aeruginosa



Staphylococcus aureus



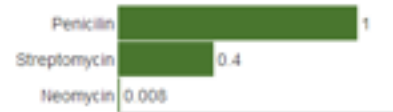
Brucella abortus



Klebsiella pneumoniae



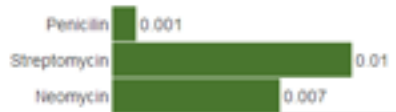
Salmonella (Eberthella) typhosa



Streptococcus fecalis



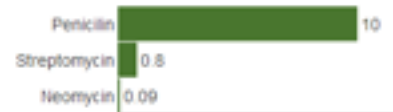
Brucella anthracis



Mycobacterium tuberculosis



Salmonella schottmuelleri



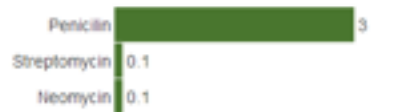
Streptococcus hemolyticus



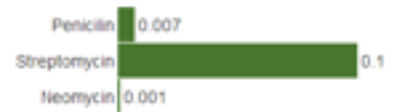
Diplococcus pneumoniae



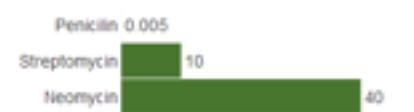
Proteus vulgaris



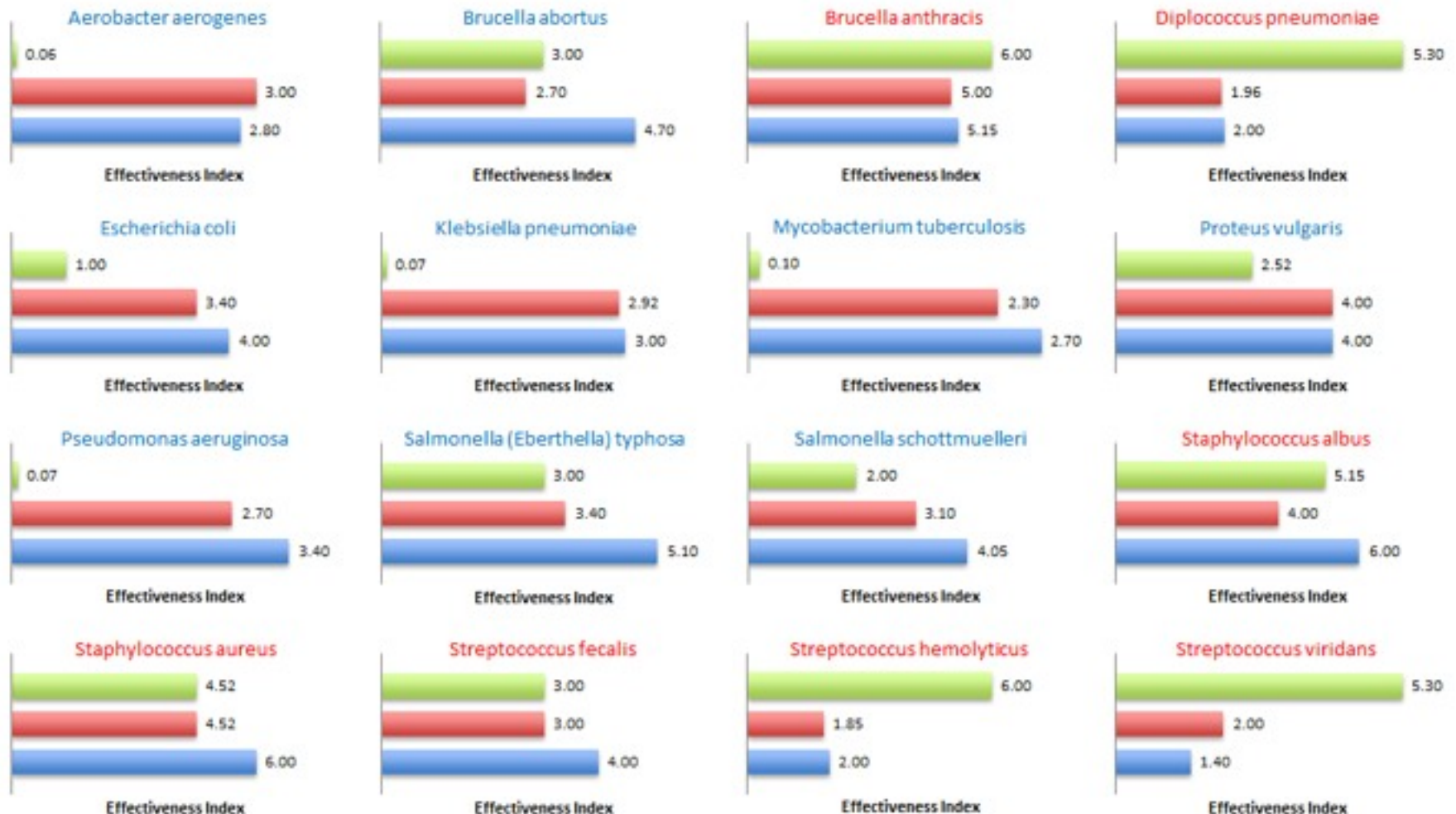
Staphylococcus albus



Streptococcus viridans



Antibiotic Effectiveness by Bacteria



Red bacteria names are Gram staining positive and blue bacteria names are Gram staining negative

Effectiveness Index = $\text{Log}(1000/\text{MIC})$

Legend:

Penicillin

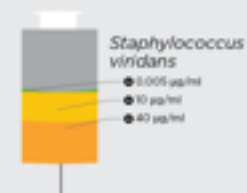
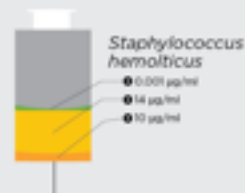
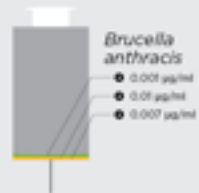
Streptomycin

Neomycin

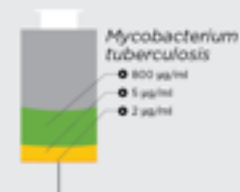
Antibiotics wwl



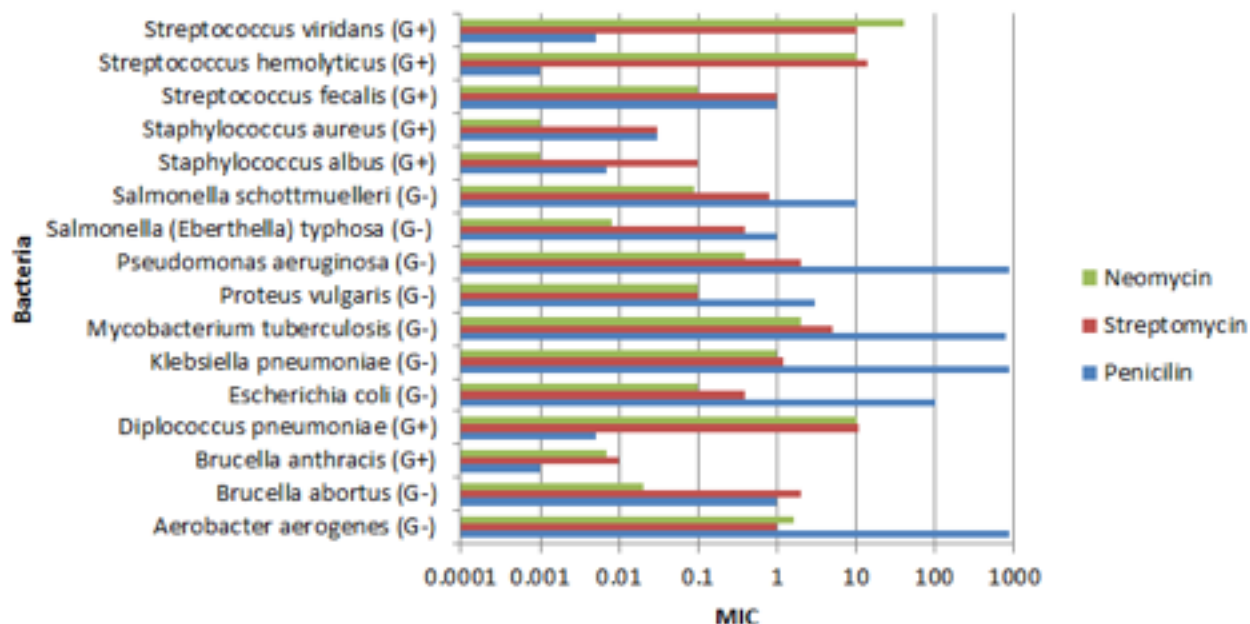
+



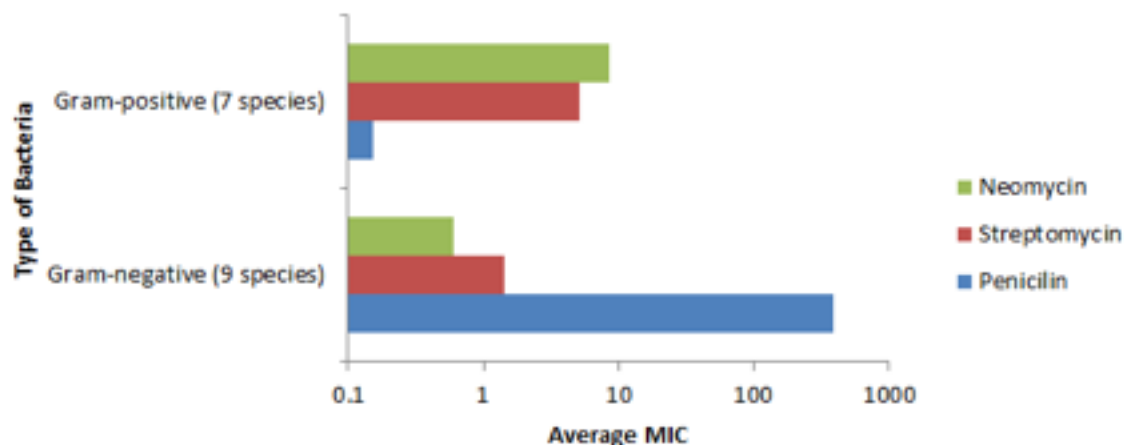
-



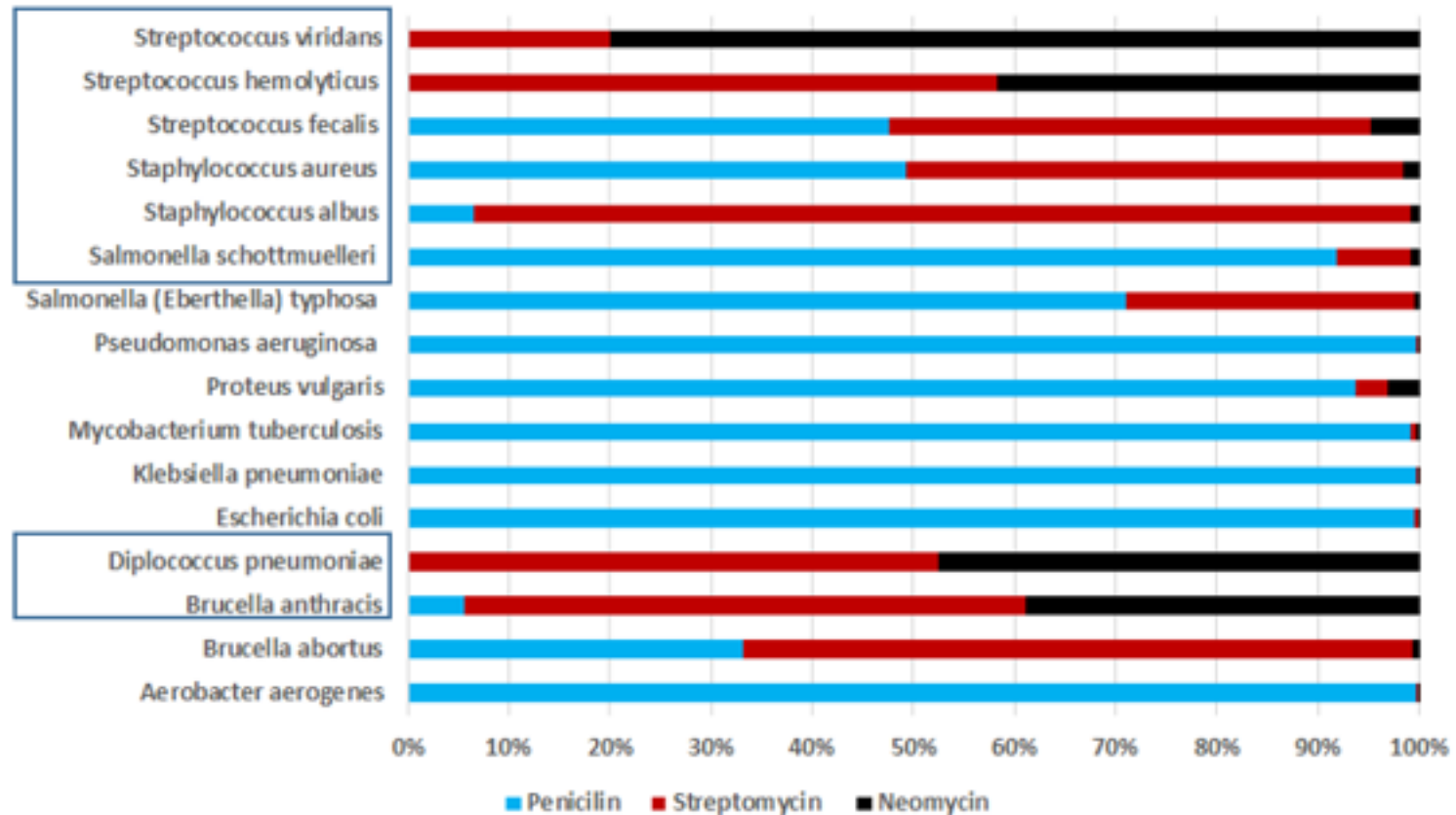
Minimum Inhibitory Concentration (MIC) of 3 Antibiotics on 16 Bacteria



Average Minimum Inhibitory Concentration of Above Antibiotics by Type of Above Bacteria

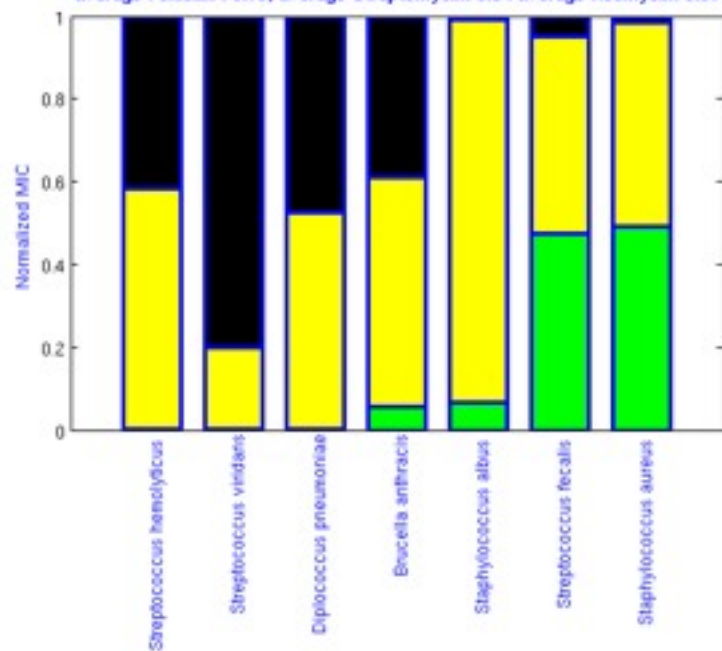


Effectiveness of Antibiotics on Bacteria

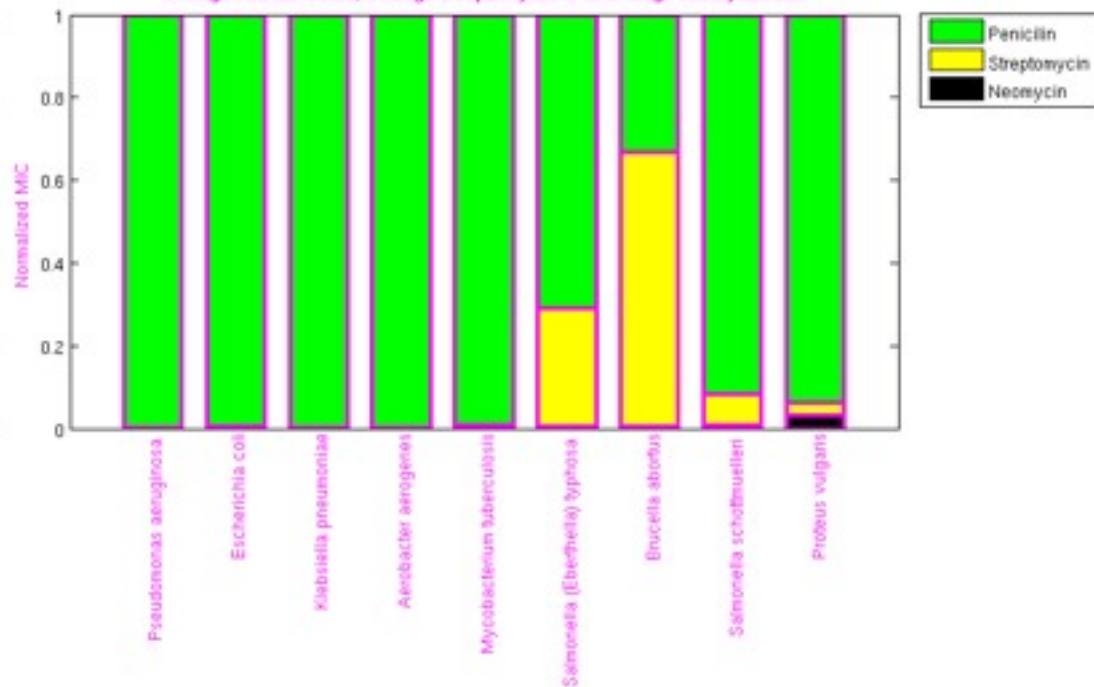


Positive Gram Staining

Normalized bar chart for Bacteria with positive Gram Staining
average Penicillin : 0.16, average Streptomycin: 0.54 average Neomycin: 0.31



Normalized bar chart for Bacteria with negative Gram Staining
average Penicillin : 0.80, average Streptomycin: 0.12 average Neomycin: 0.01



Effectiveness Comparison between Penicillin and Streptomycin / Neomycin

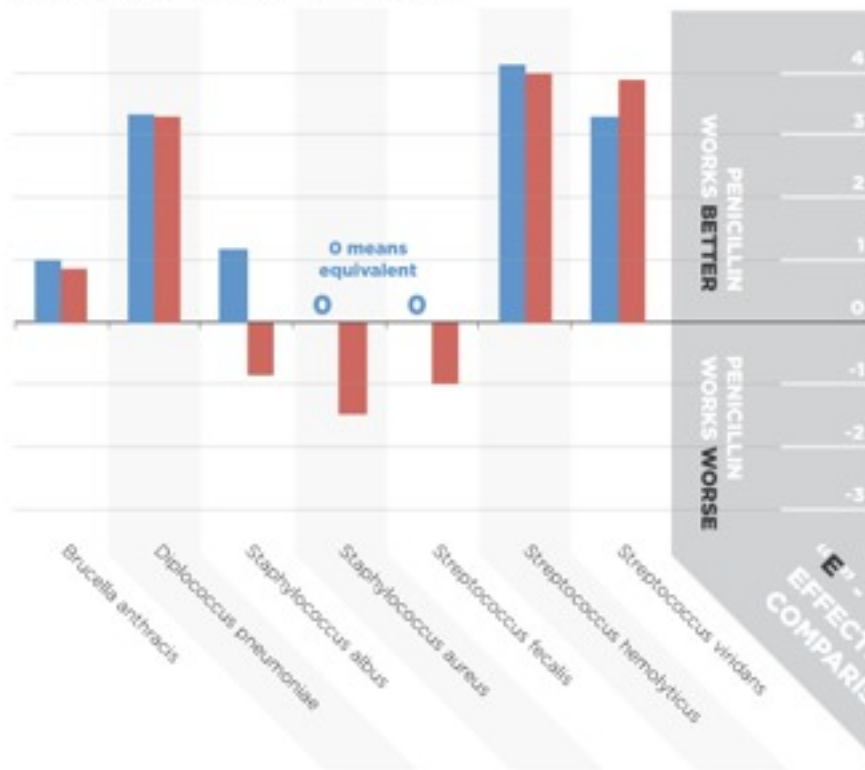
CSE 512 A1 | CHEN-HUNG WU

EFFECTIVENESS COMPARISON

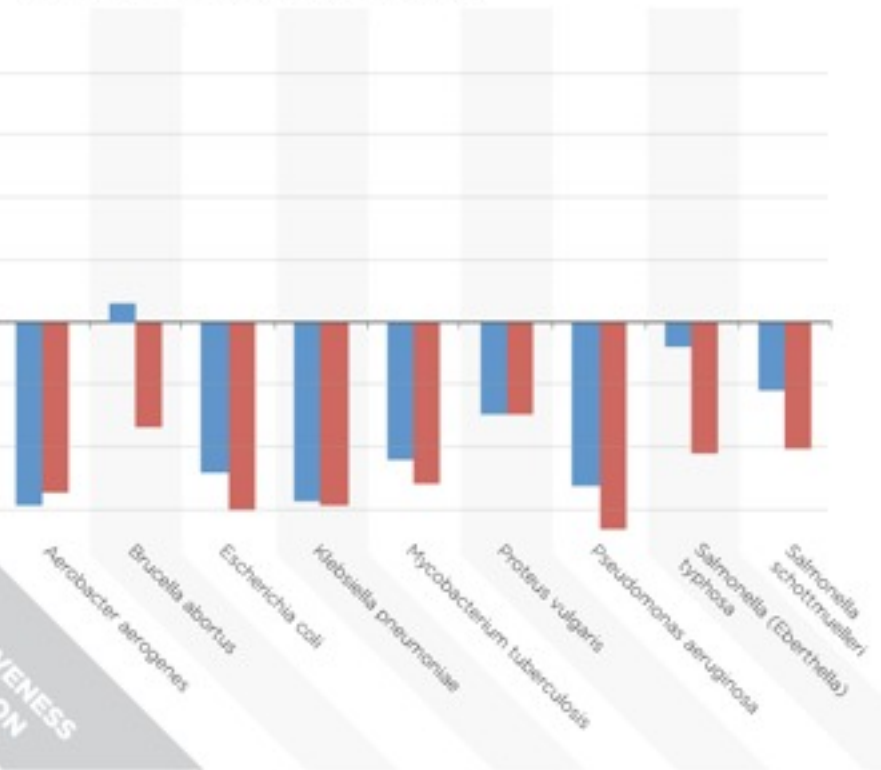
$$E = \log\left(\frac{\text{MIC}_x}{\text{MIC}_{\text{pen}}}\right)$$

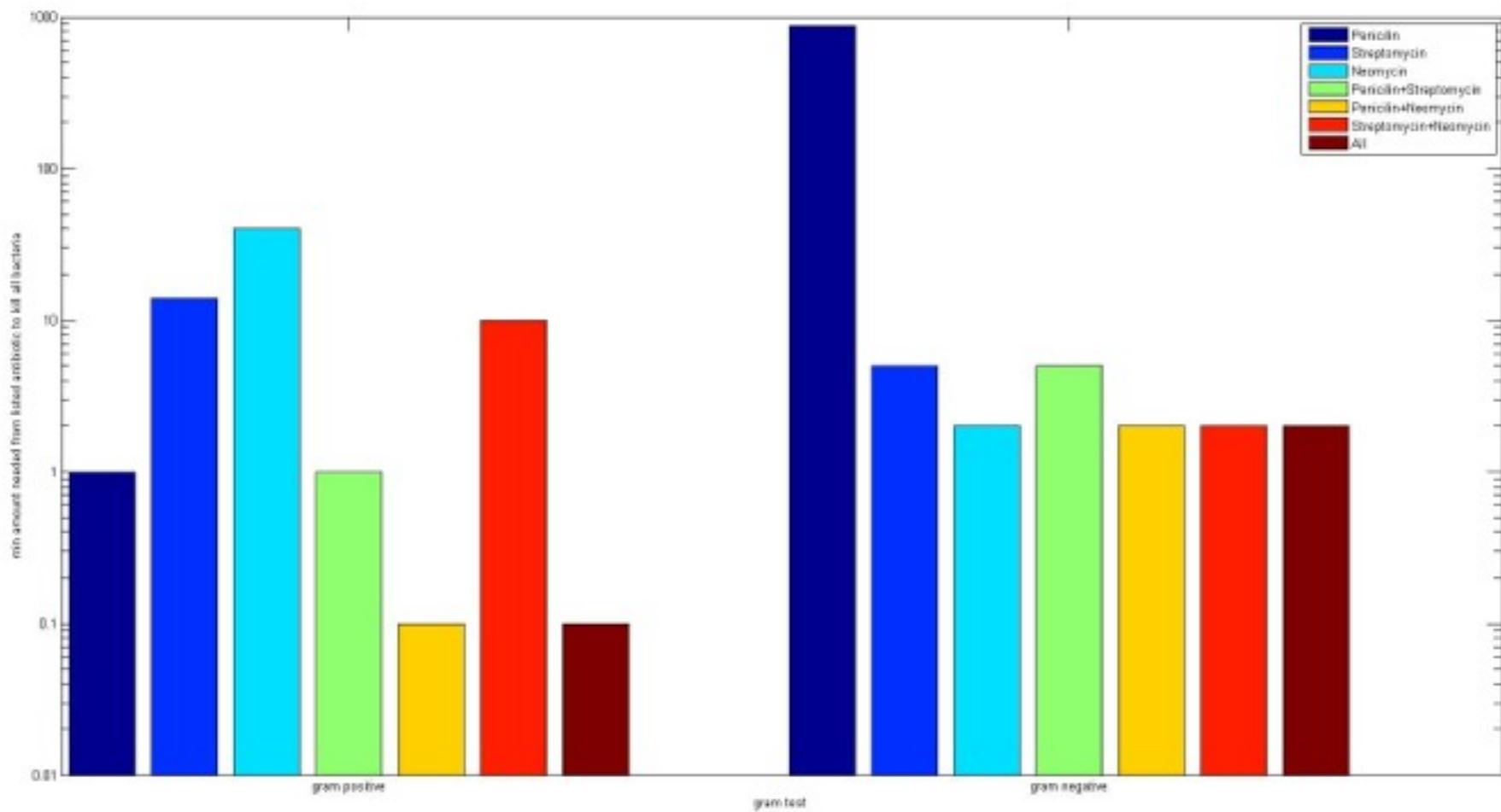
Minimum Inhibitory Concentration of x
Minimum Inhibitory Concentration of Penicillin

Gram Staining POSITIVE Bacteria

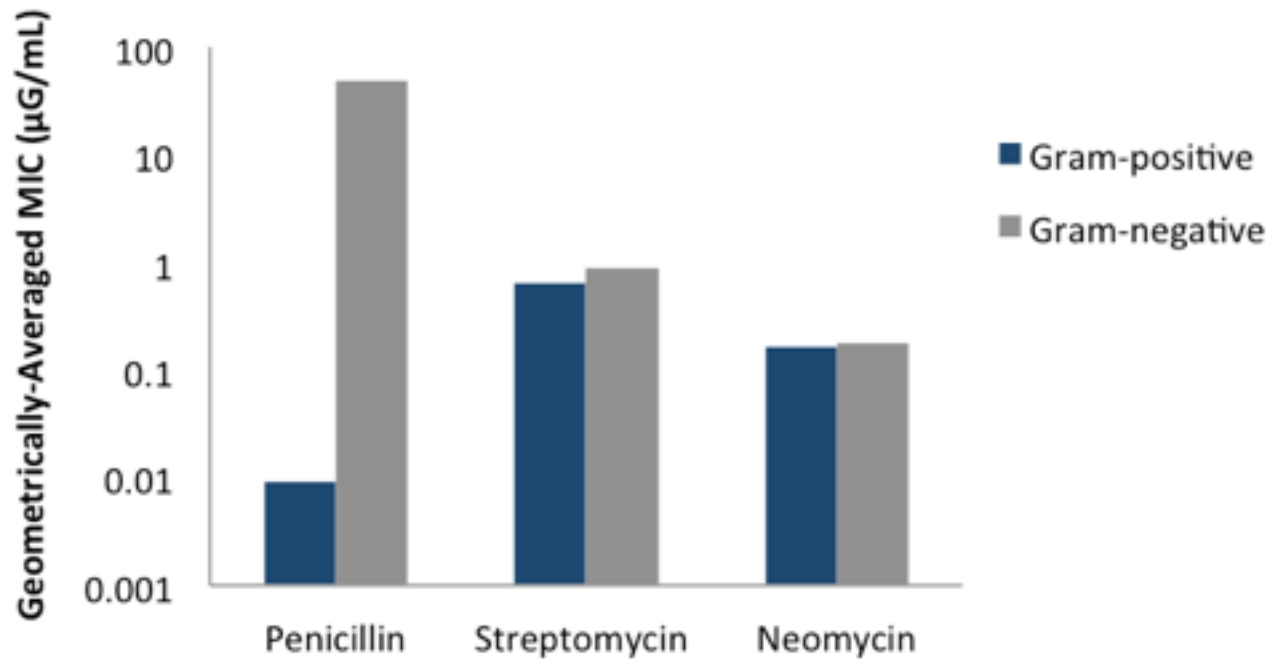


Gram Staining NEGATIVE Bacteria

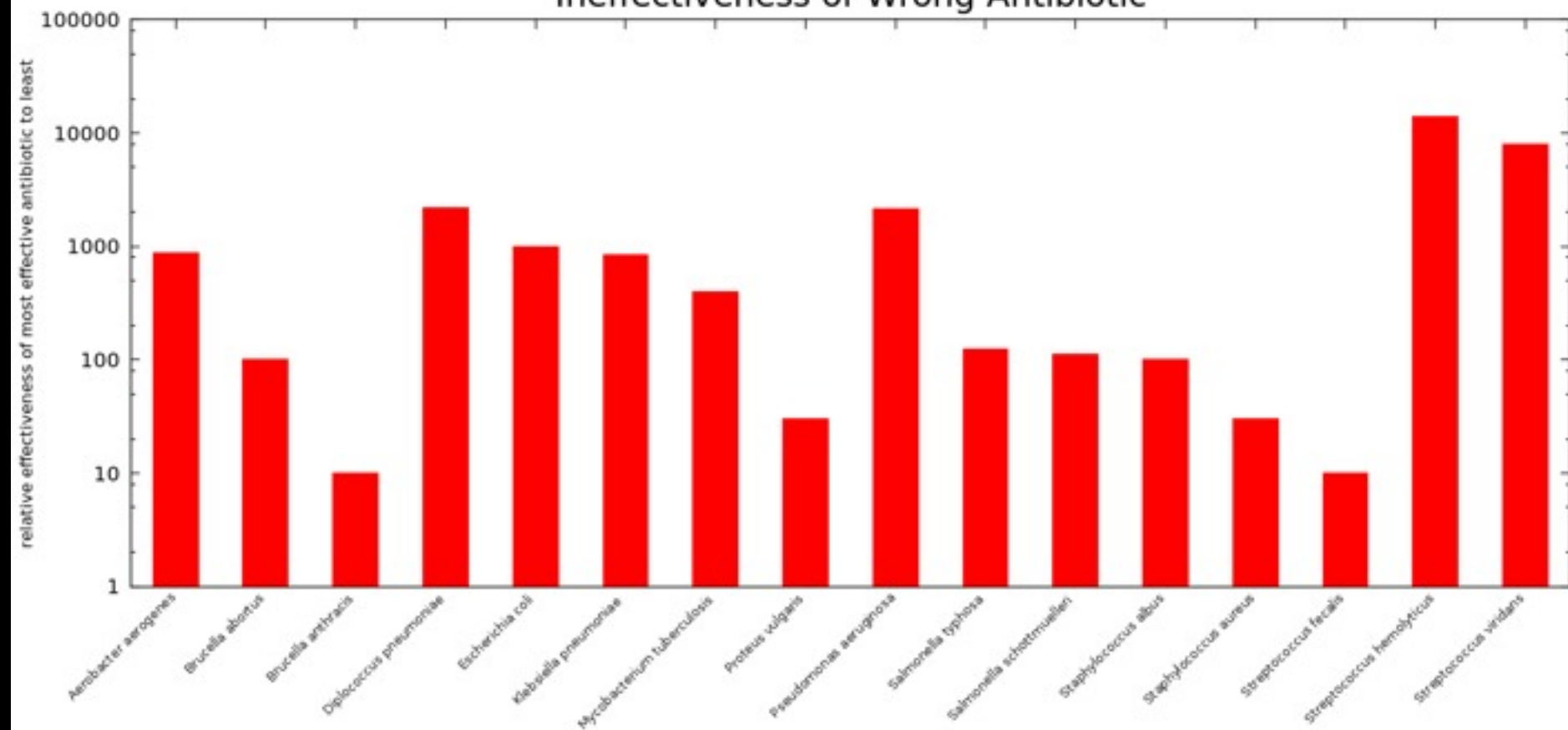




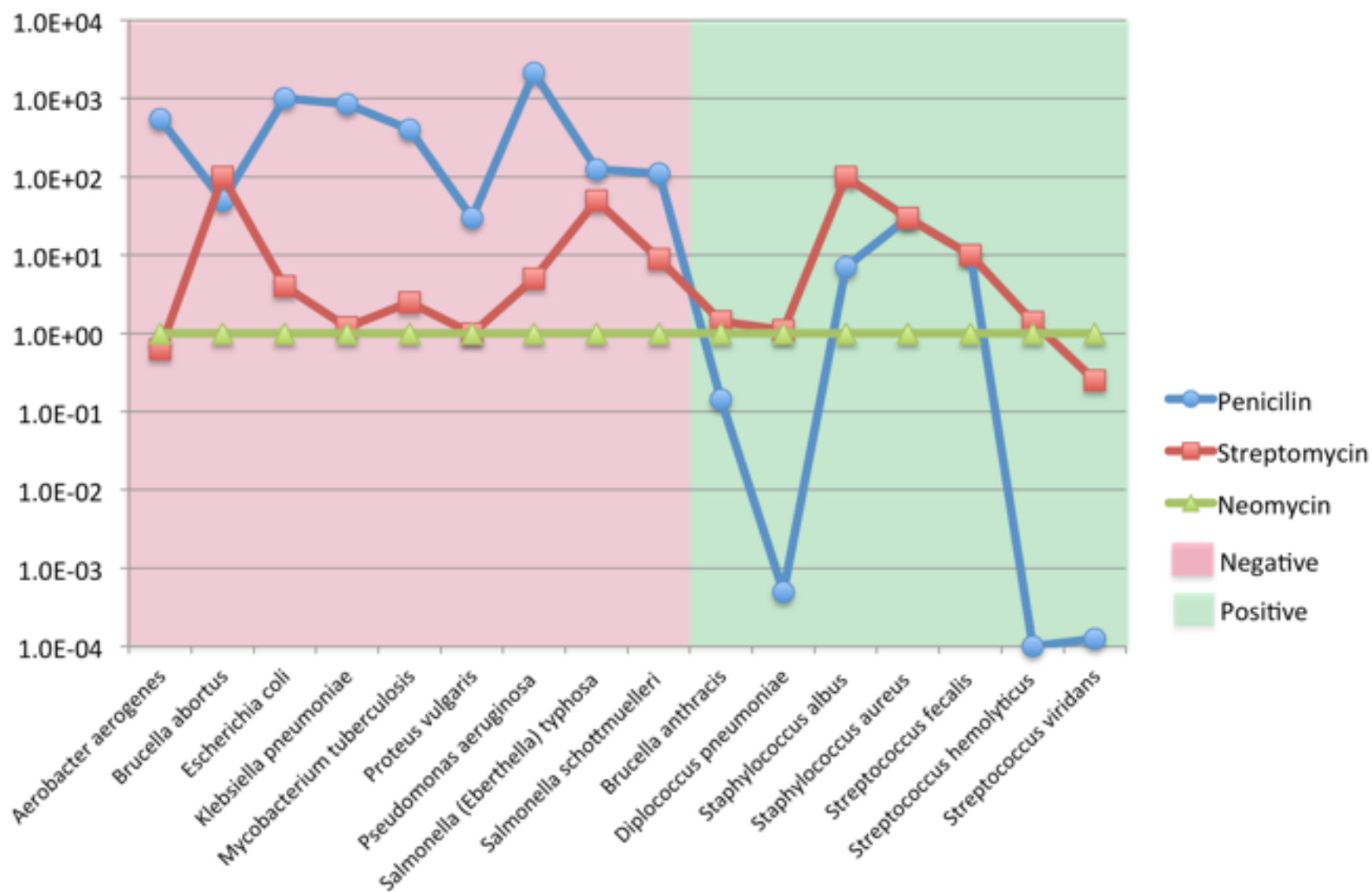
Average MIC of antibiotic on Gram-negative
versus Gram-positive bacteria



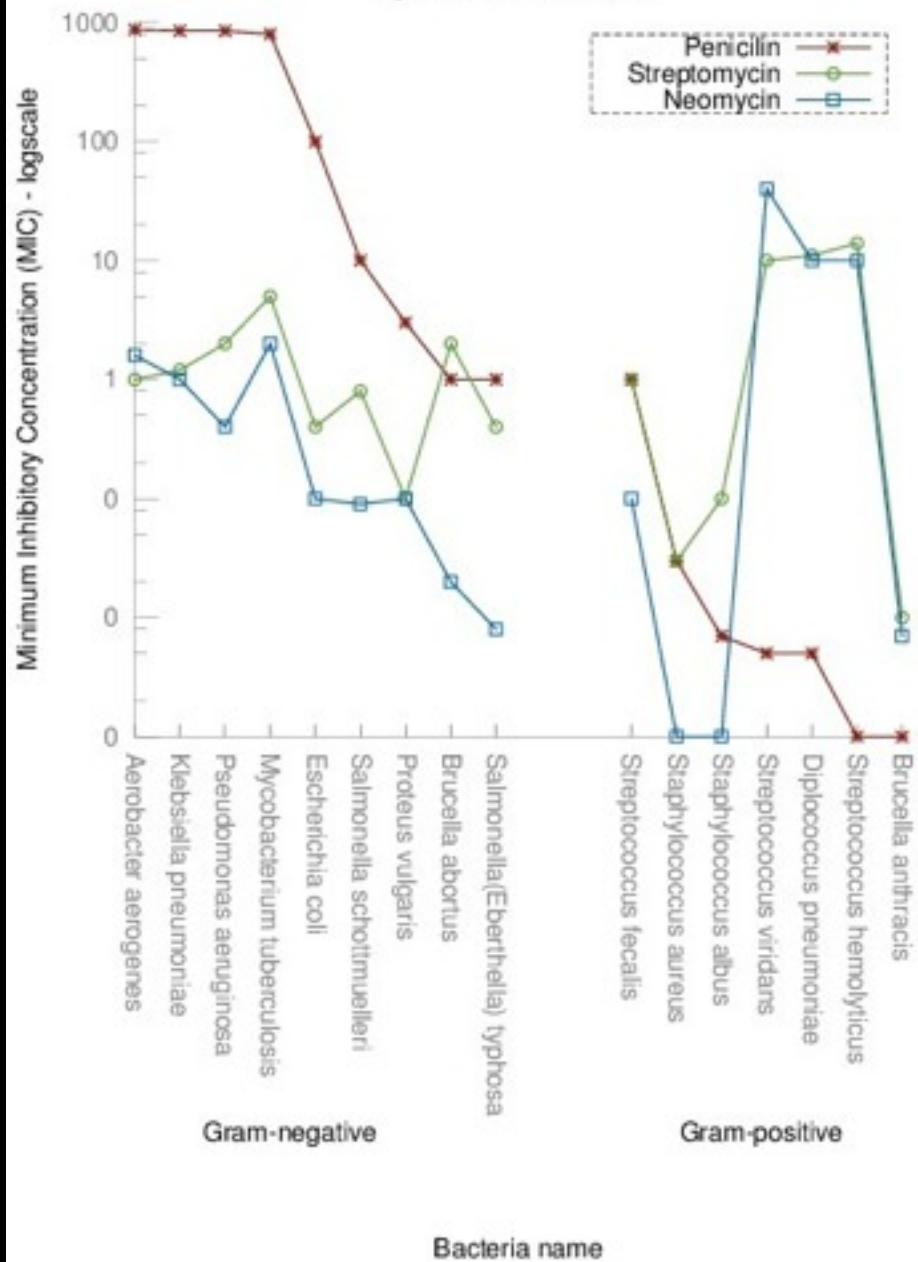
Ineffectiveness of Wrong Antibiotic



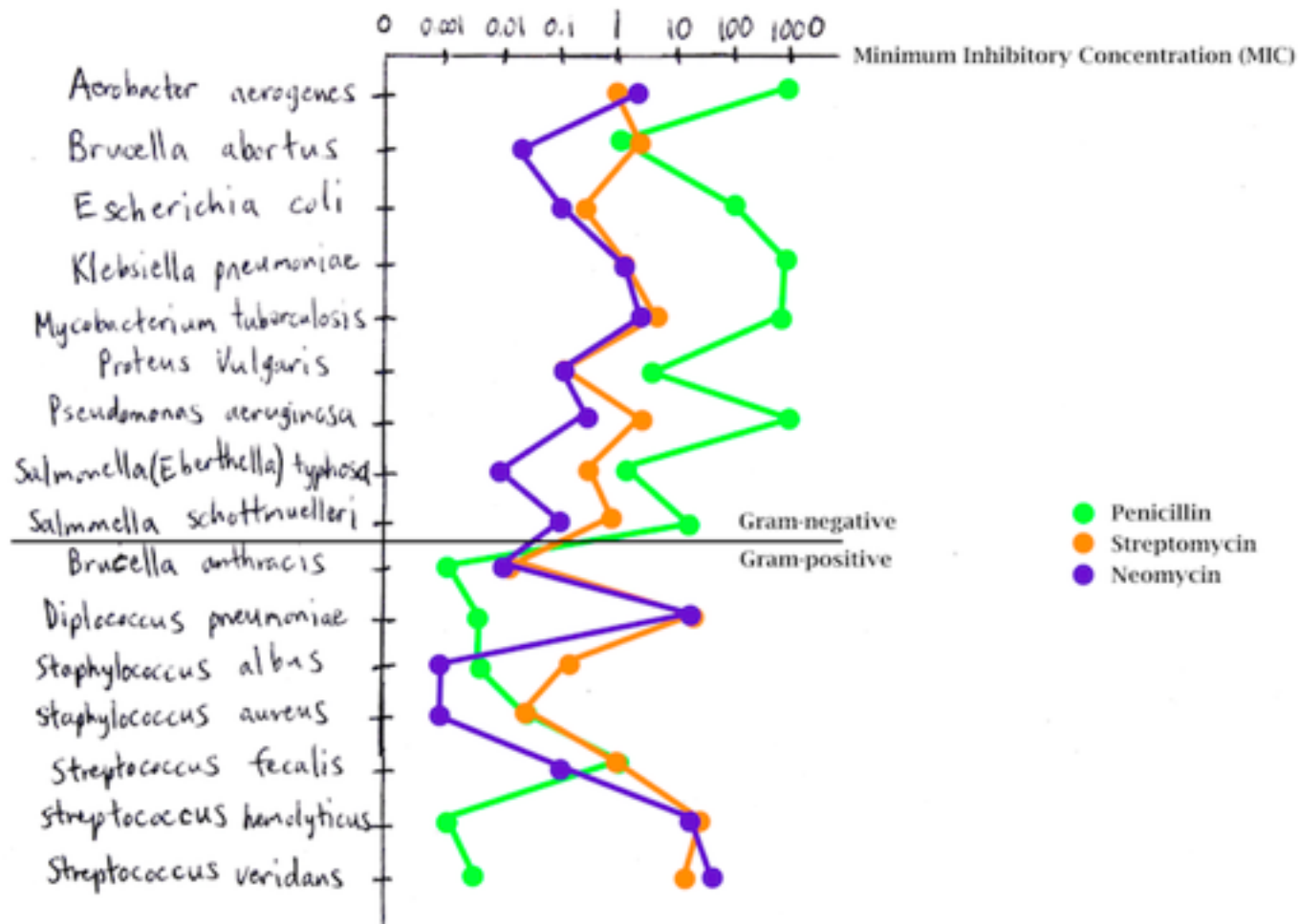
Line Charts



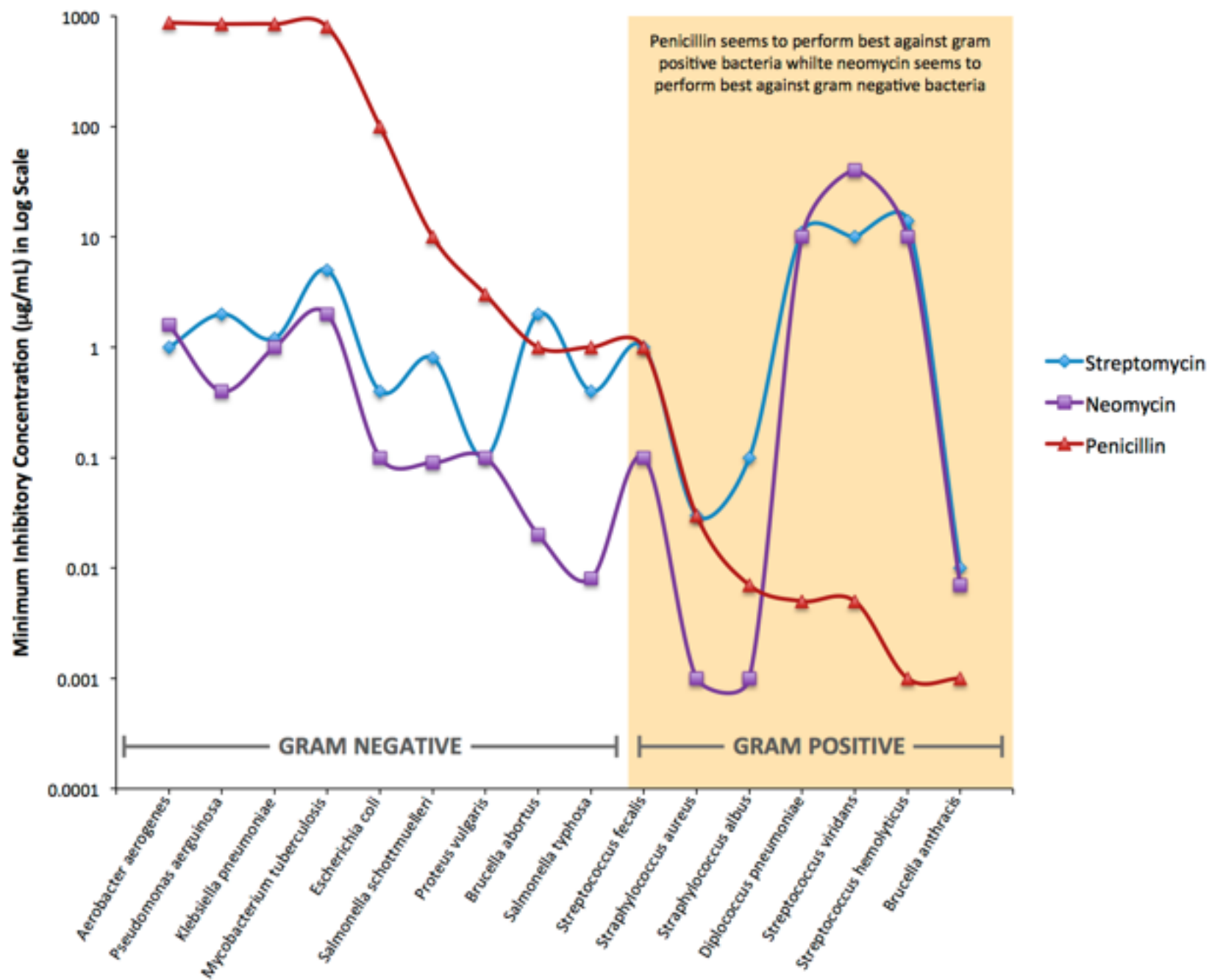
Effectiveness of Penicillin, Streptomycin and Neomycin against various bacteria.



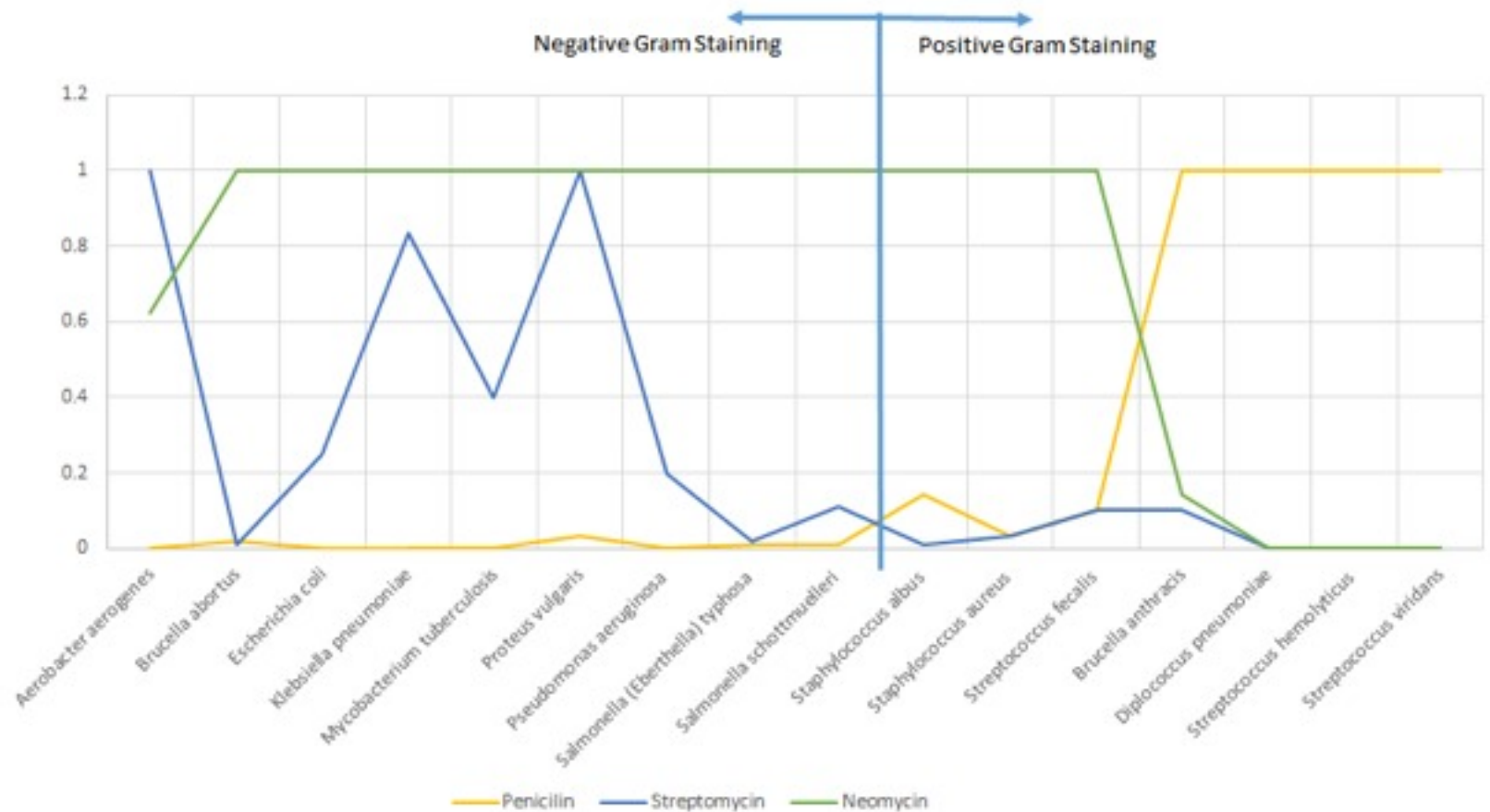
Measuring the effectiveness of popular antibiotics on various bacteria



Effectiveness of Streptomycin, Neomycin, and Penicillin Against 16 Bacteria

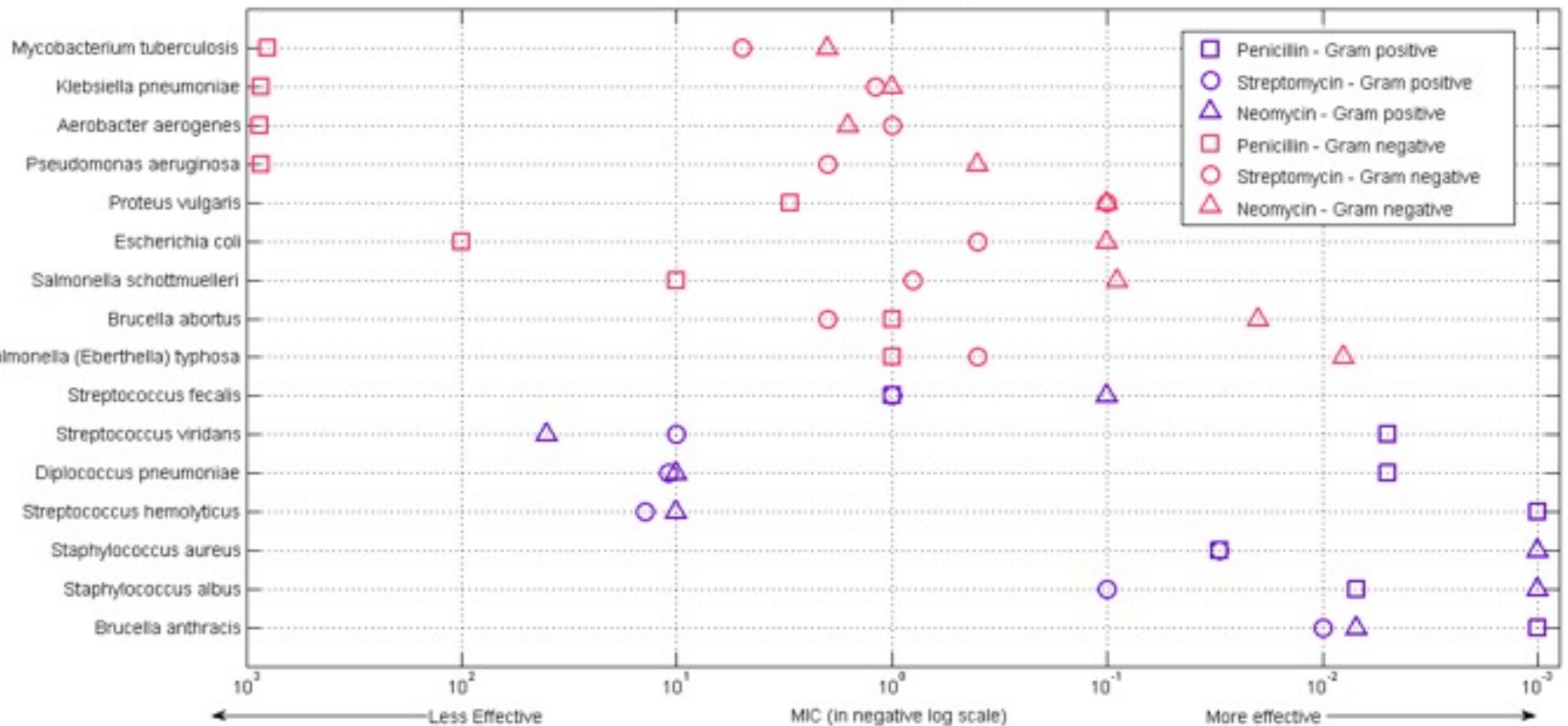


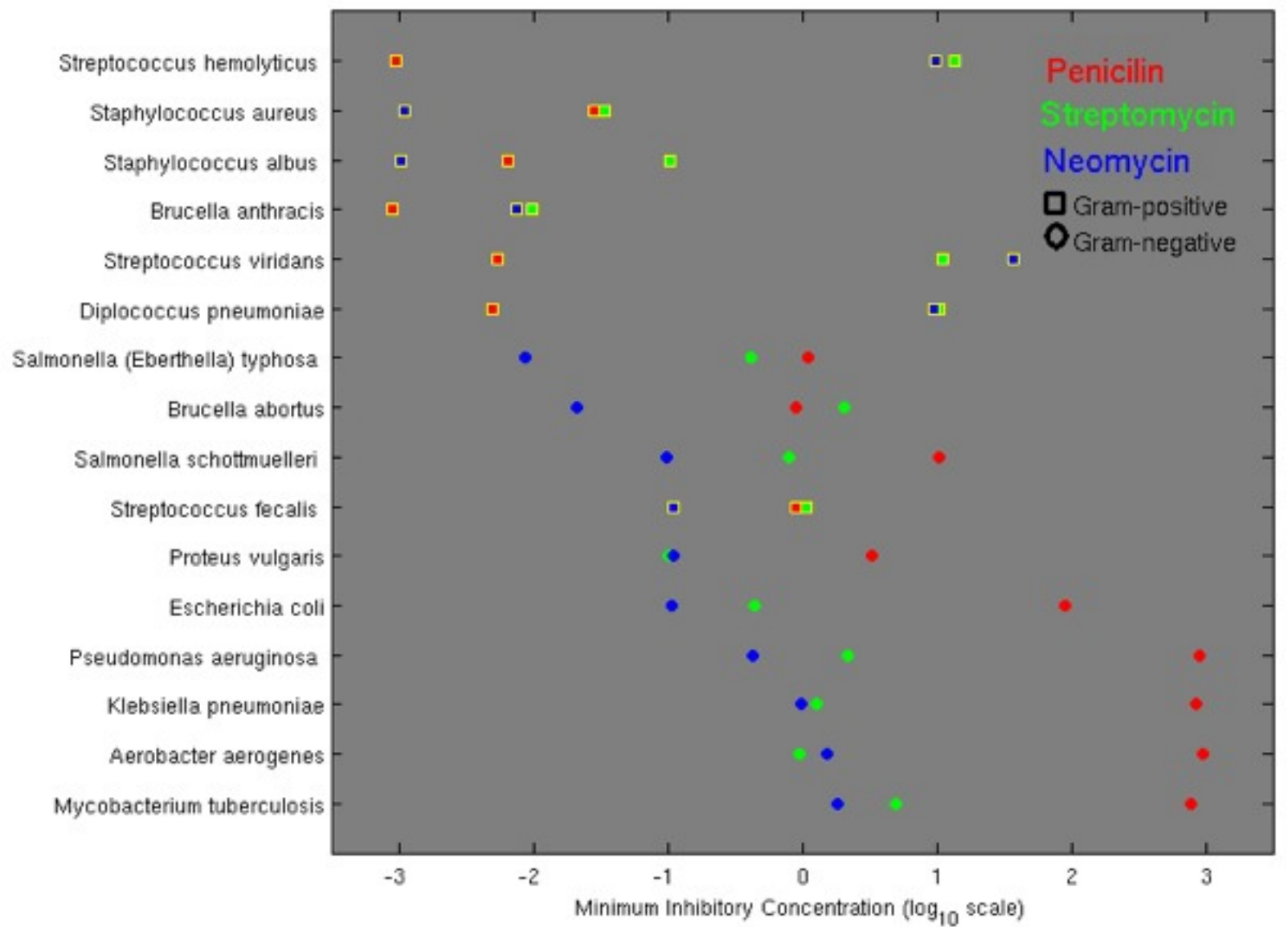
Normalized Effectiveness of 3 antibiotics on 16 Bacterium

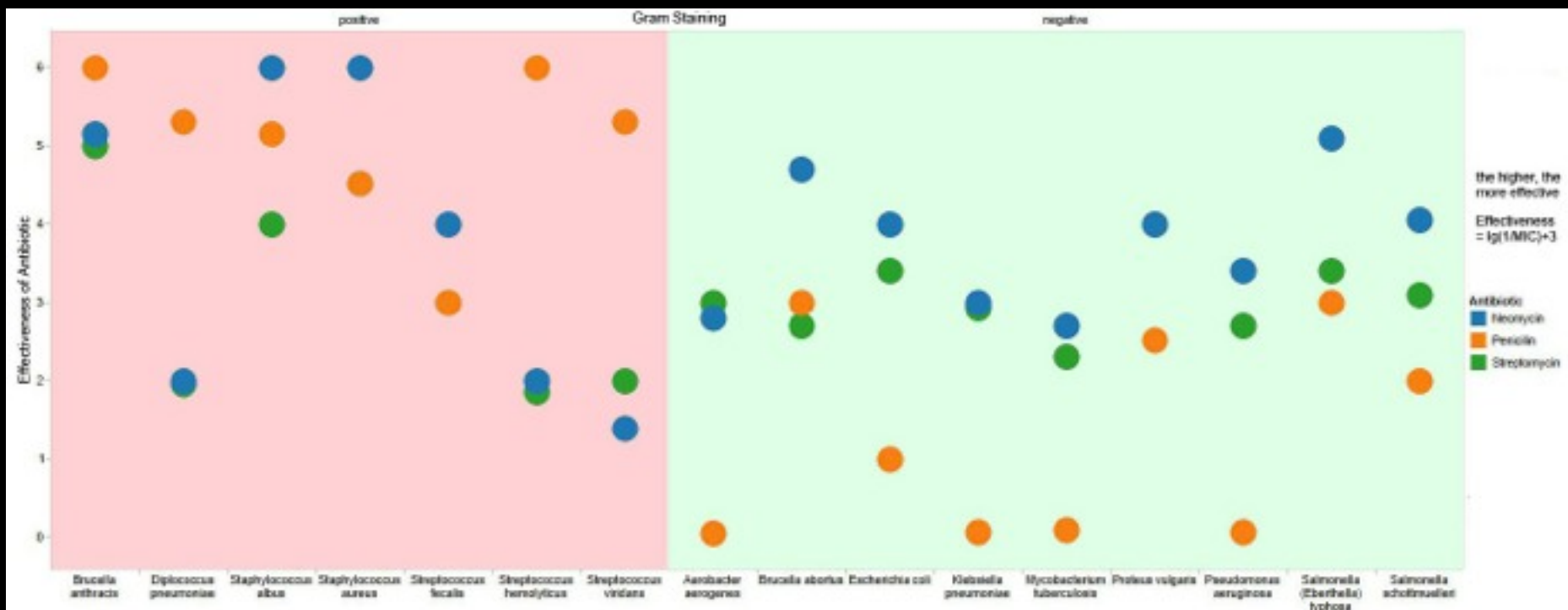


Dot Plots

Effectiveness of Antibiotics (Burtin's Data)

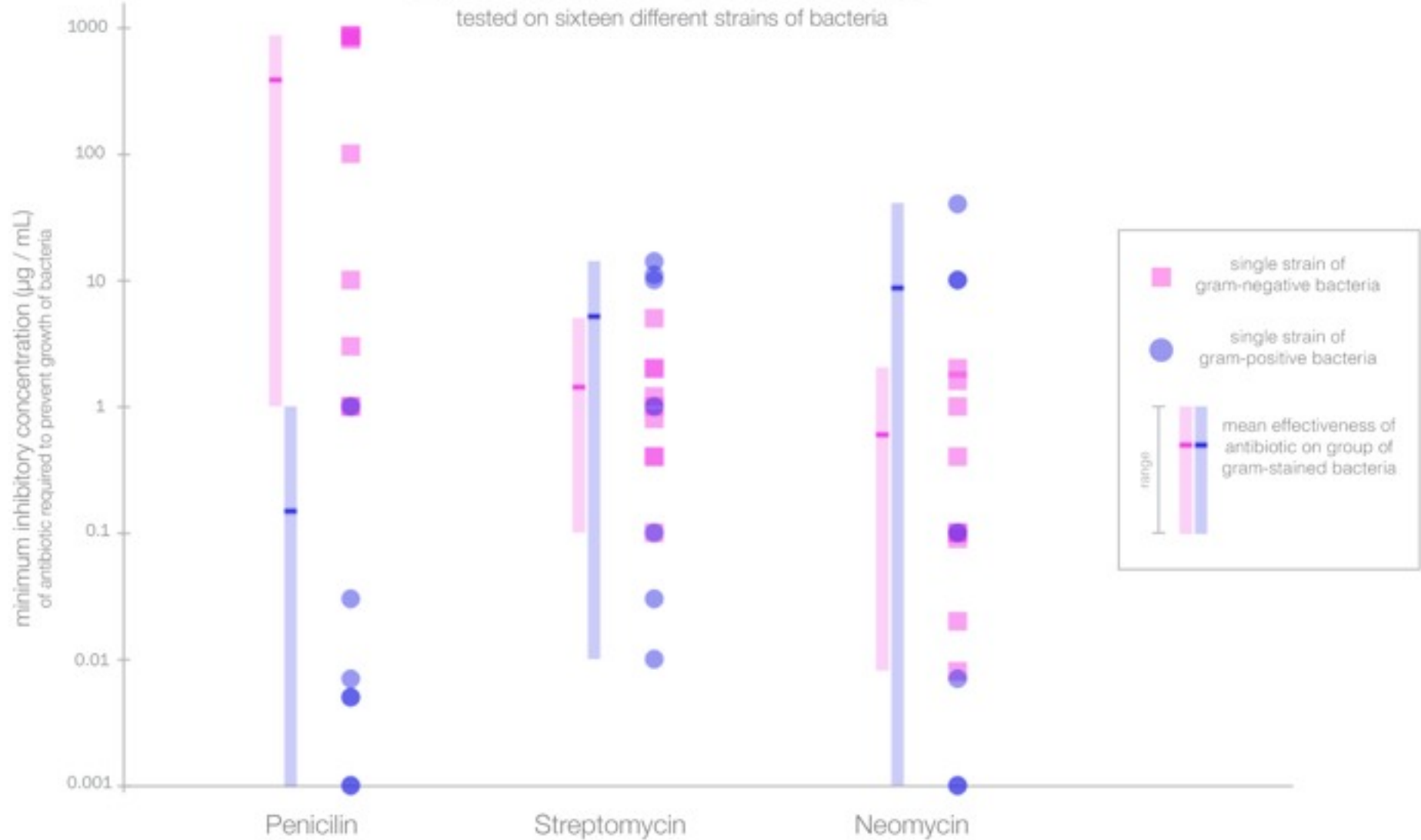




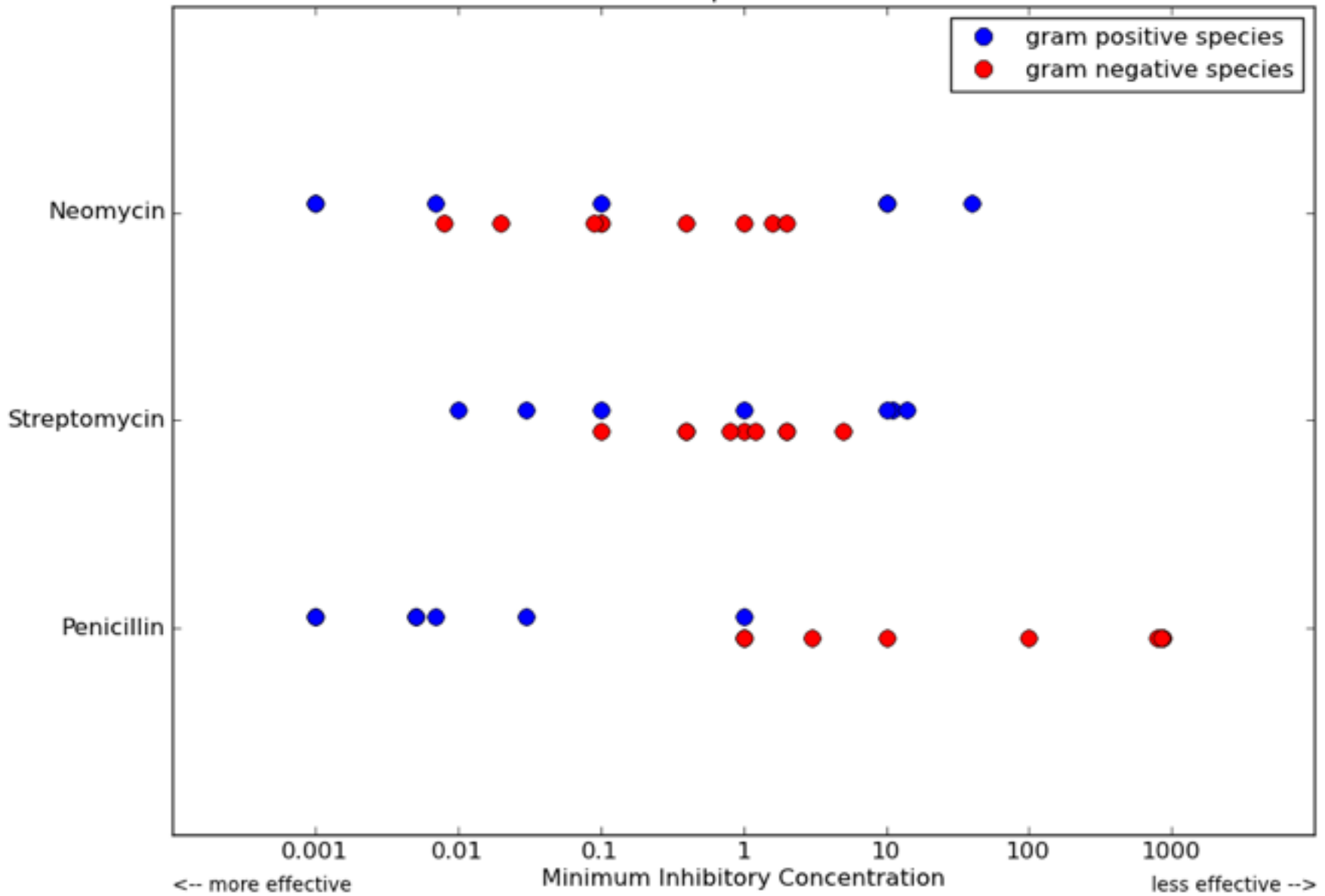


Effectiveness of Antibiotics

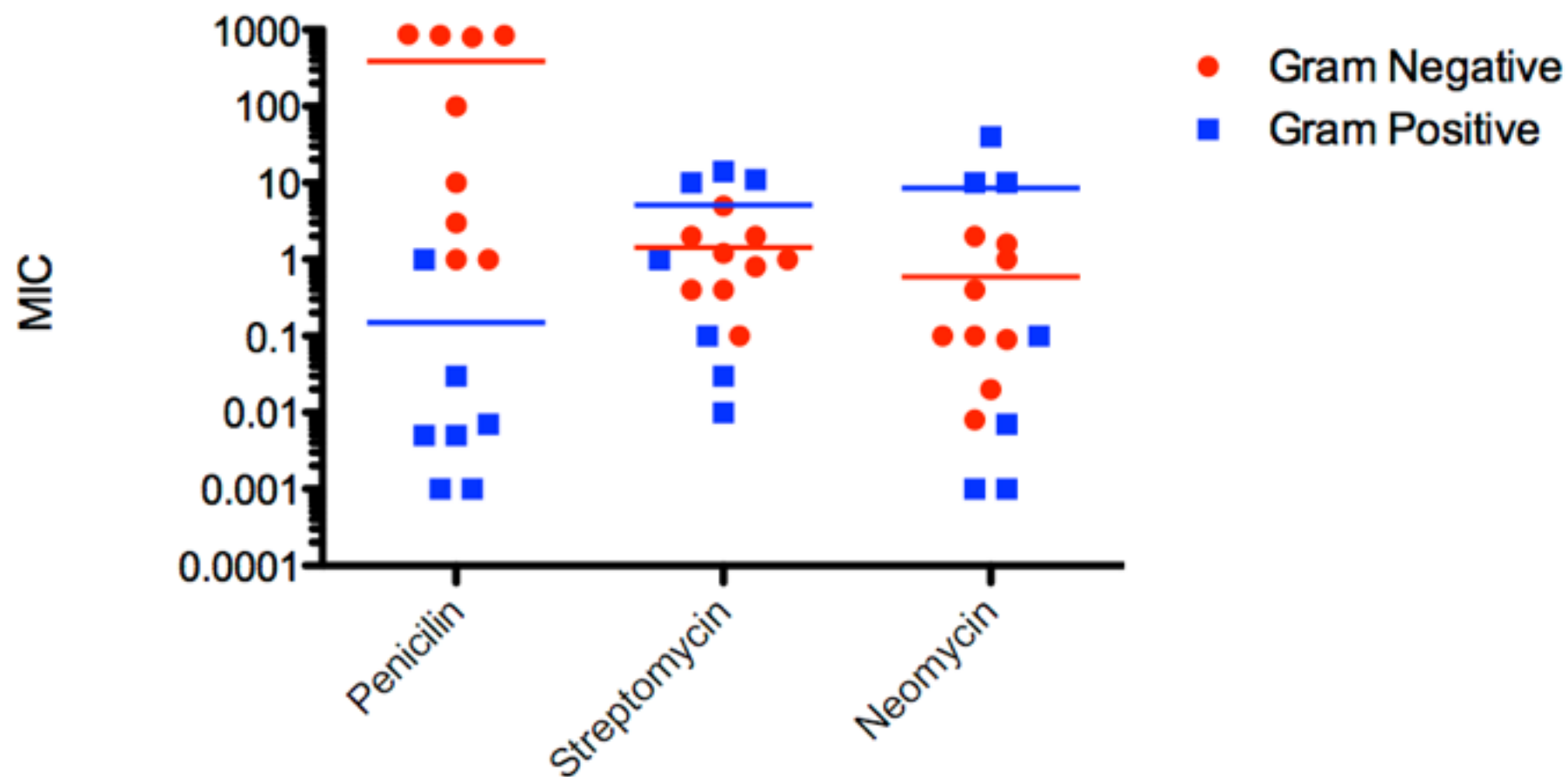
tested on sixteen different strains of bacteria



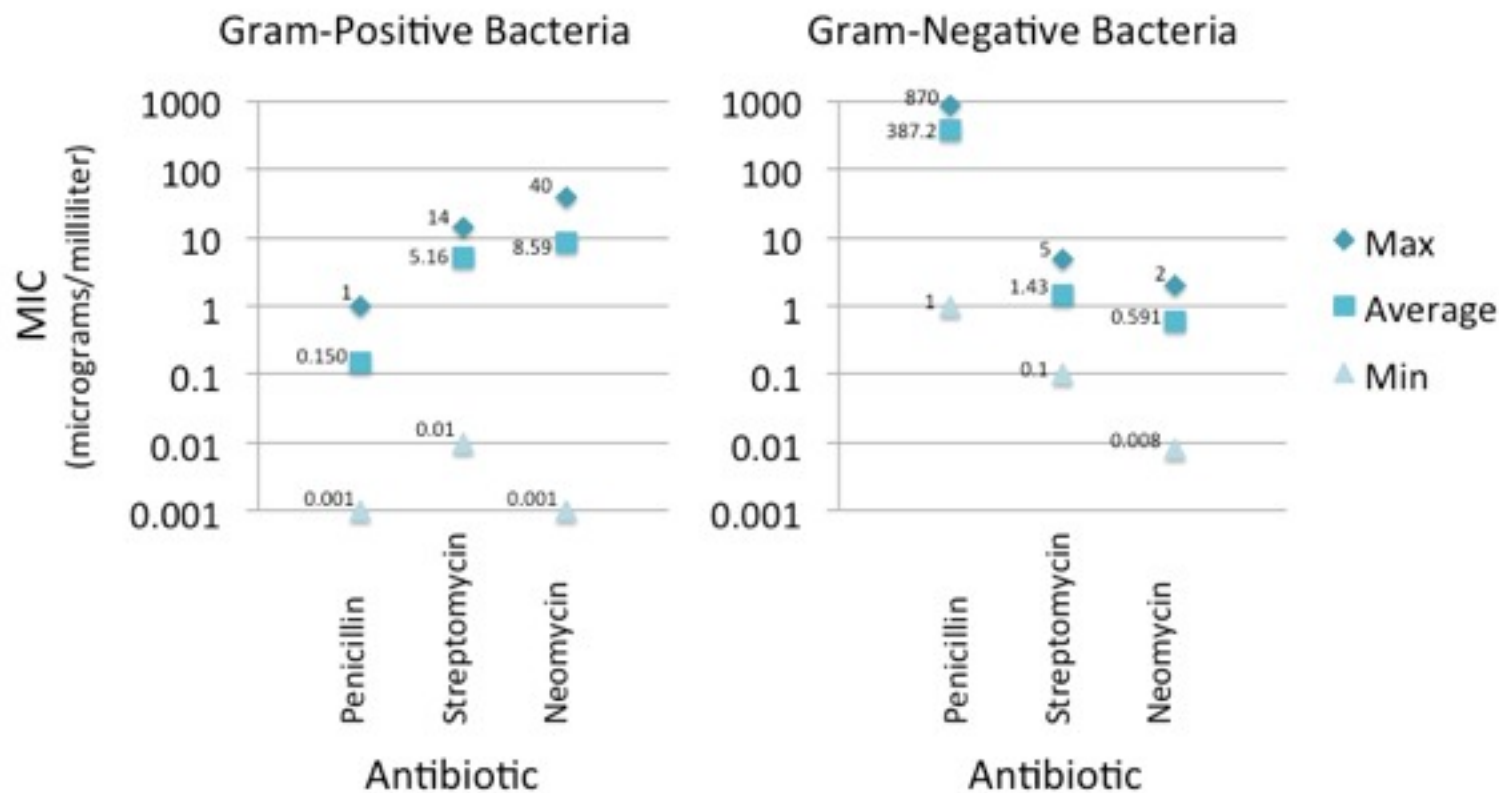
Antibiotic concentrations required to inhibit selected bacteria

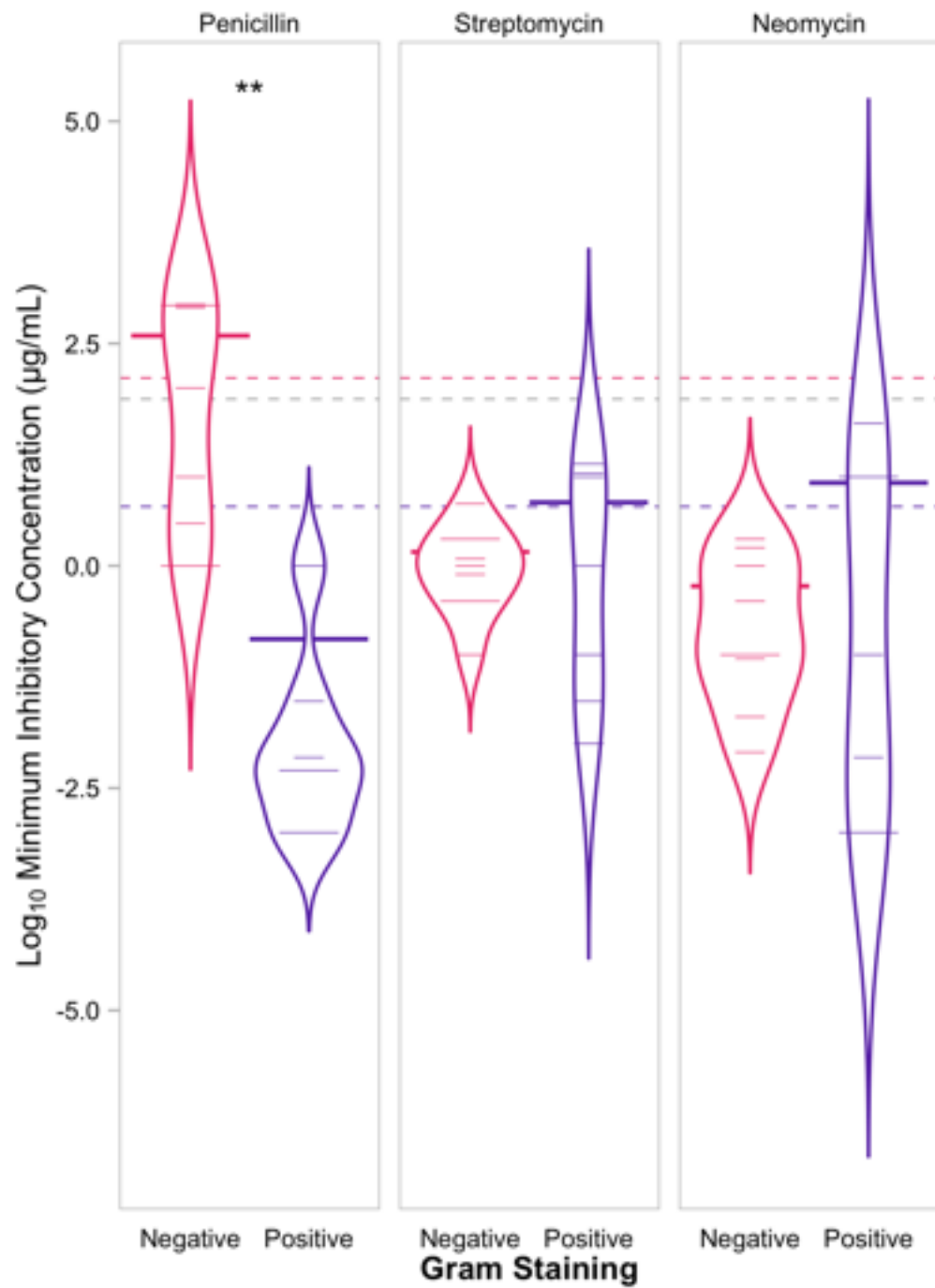


Correlationship between Gram staining and antibiotics



Minimum Inhibitory Concentration (MIC) for Antibiotics to Inhibit Bacteria Growth





Tables / Heat Maps

Which Antibiotic Is Most Effective For Which Bacteria?

Bacteria		Penicillin	Antibiotic Streptomycin	Neomycin
	Aerobacter aerogenes	870	1	1.6
	Brucella abortus	1	2	0.02
	Brucella anthracis	0.001	0.01	0.007
	Diplococcus pneumoniae	0.005	11	10
	Escherichia coli	100	0.4	0.1
	Klebsiella pneumoniae	850	1.2	1
	Myobacterium tuberculosis	800	5	2
	Proteus vulgaris	3	0.1	0.1
	Pseudomonas aerouginosa	850	2	0.4
	Salmonella (Eberthella) typhosa	1	0.4	0.008
	Salmonella schottmuelleri	10	0.8	0.09
	Staphylococcus albus	0.007	0.1	0.001
	Staphylococcus aureus	0.03	0.03	0.001
	Streptococcus fecalis	1	1	0.1
	Streptococcus hemolyticus	0.001	14	10
Streptococcus viridans	0.005	10	40	

Values in the table indicate the minimum inhibitory concentration measured for an antibiotic. The smallest values for each bacteria are highlighted.

Minimum Inhibitory Concentration (MIC) of antibiotics required to prevent growth of bacteria

		MIC for Antibiotic		
		Penicillin	Streptomycin	Neomycin
Bacteria	Streptococcus hemolyticus	0.001	14	10
	Streptococcus viridans	0.005	10	40
	Diplococcus pneumoniae	0.005	11	10
	Staphylococcus albus	0.007	0.1	0.001
	Staphylococcus aureus	0.03	0.03	0.001
	Brucella anthracis	0.001	0.01	0.007
	Streptococcus fecalis	1	1	0.1
	Proteus vulgaris	3	0.1	0.1
	Brucella abortus	1	2	0.02
	Salmonella schottmuelleri	10	0.8	0.09
Gram-Negative	Salmonella (Eberthella) typhosa	1	0.4	0.008
	Mycobacterium tuberculosis	800	5	2
	Klebsiella pneumoniae	850	1.2	1
	Aerobacter aerogenes	870	1	1.6
	Escherichia coli	100	0.4	0.1
	Pseudomonas aeruginosa	850	2	0.4

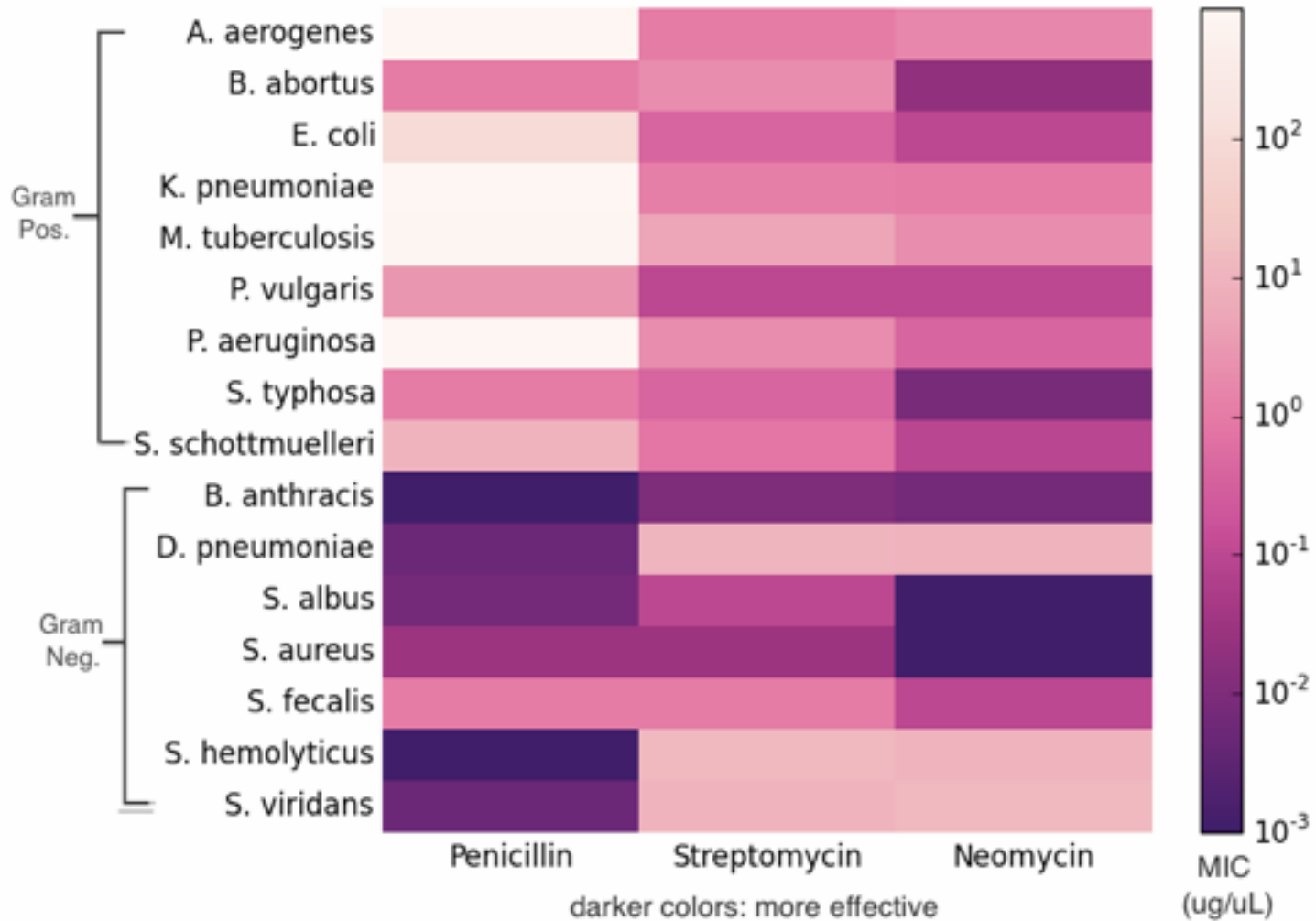
Bacterial Resistance to Antibiotics

Bacteria	Gram Staining	Antibiotics		
		Penicilin	Streptomycin	Neomycin
Aerobacter aerogenes	negative			
Brucella abortus	negative			
Brucella anthracis	positive			
Diplococcus pneumoniae	positive			
Escherichia coli	negative			
Klebsiella pneumoniae	negative			
Mycobacterium tuberculosis	negative			
Proteus vulgaris	negative			
Pseudomonas aeruginosa	negative			
Salmonella (Eberthella) typhosa	negative			
Salmonella schottmuelleri	negative			
Staphylococcus albus	positive			
Staphylococcus aureus	positive			
Streptococcus fecalis	positive			
Streptococcus hemolyticus	positive			
Streptococcus viridans	positive			

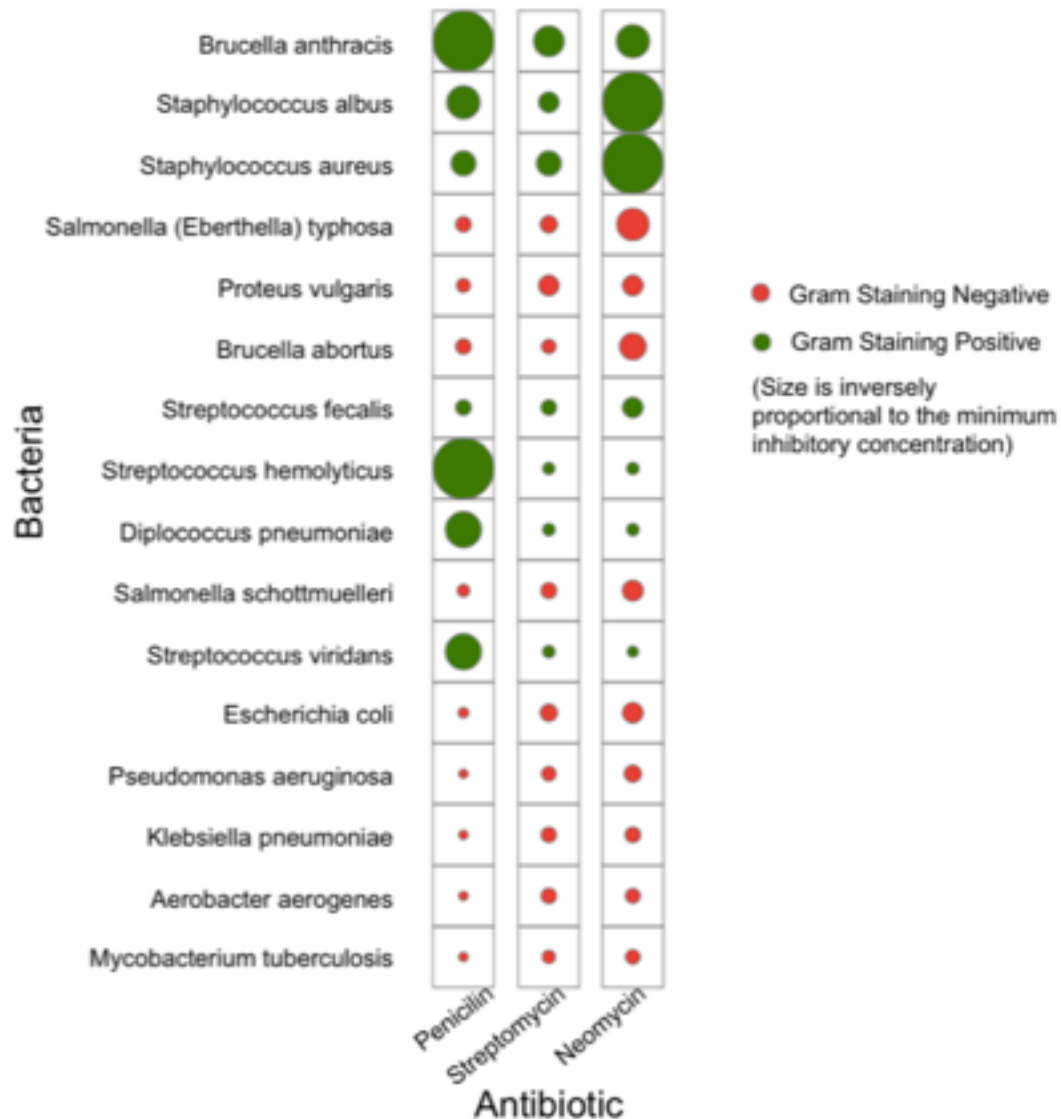
0.001 0.01 0.1 1 10 100 870

minimum inhibitory concentration (µg/L)

Effectiveness of Antibiotics



Effectiveness of the World War II "Wonder Drugs"





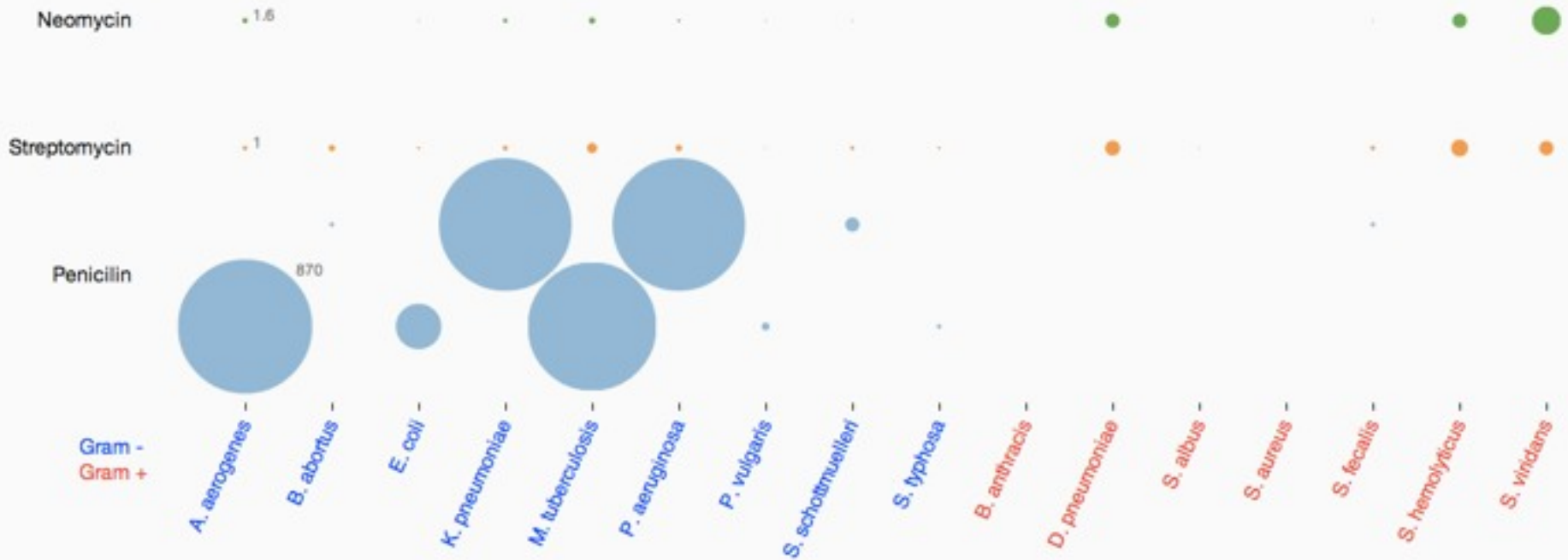
Gram Staining



Minimum Inhibitory Concentration (MIC)

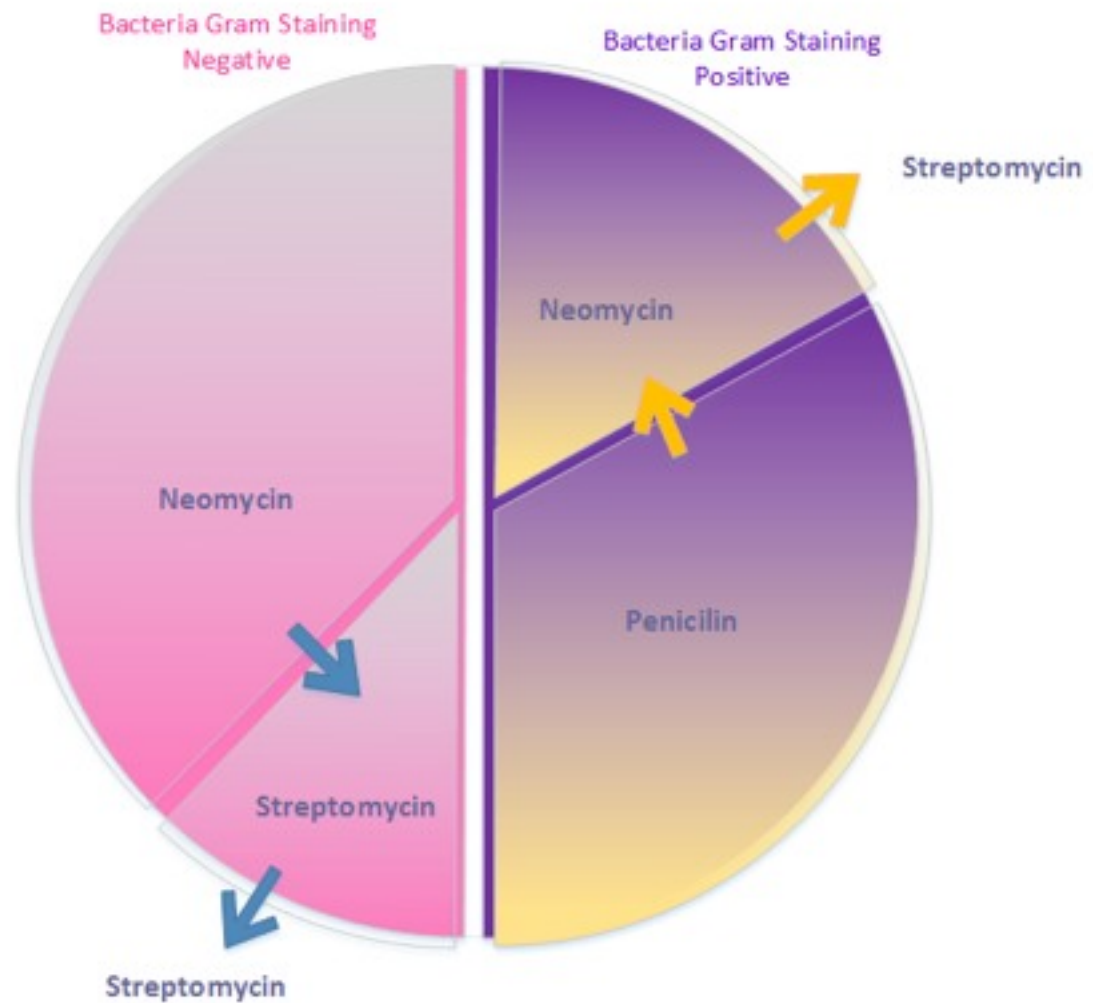


Burtin's MIC Values, Plotted as Area



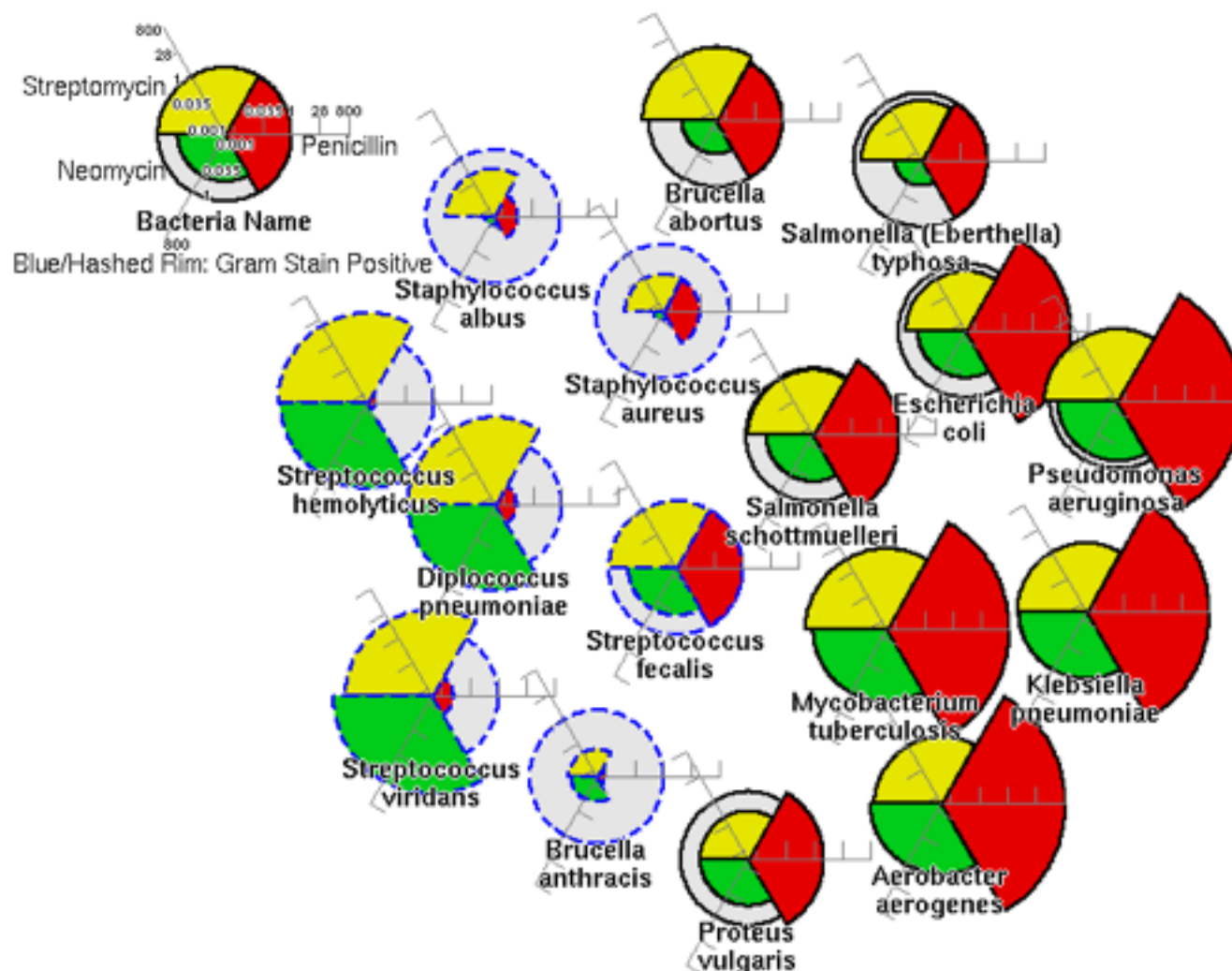
Other

What Antibiotics to Use

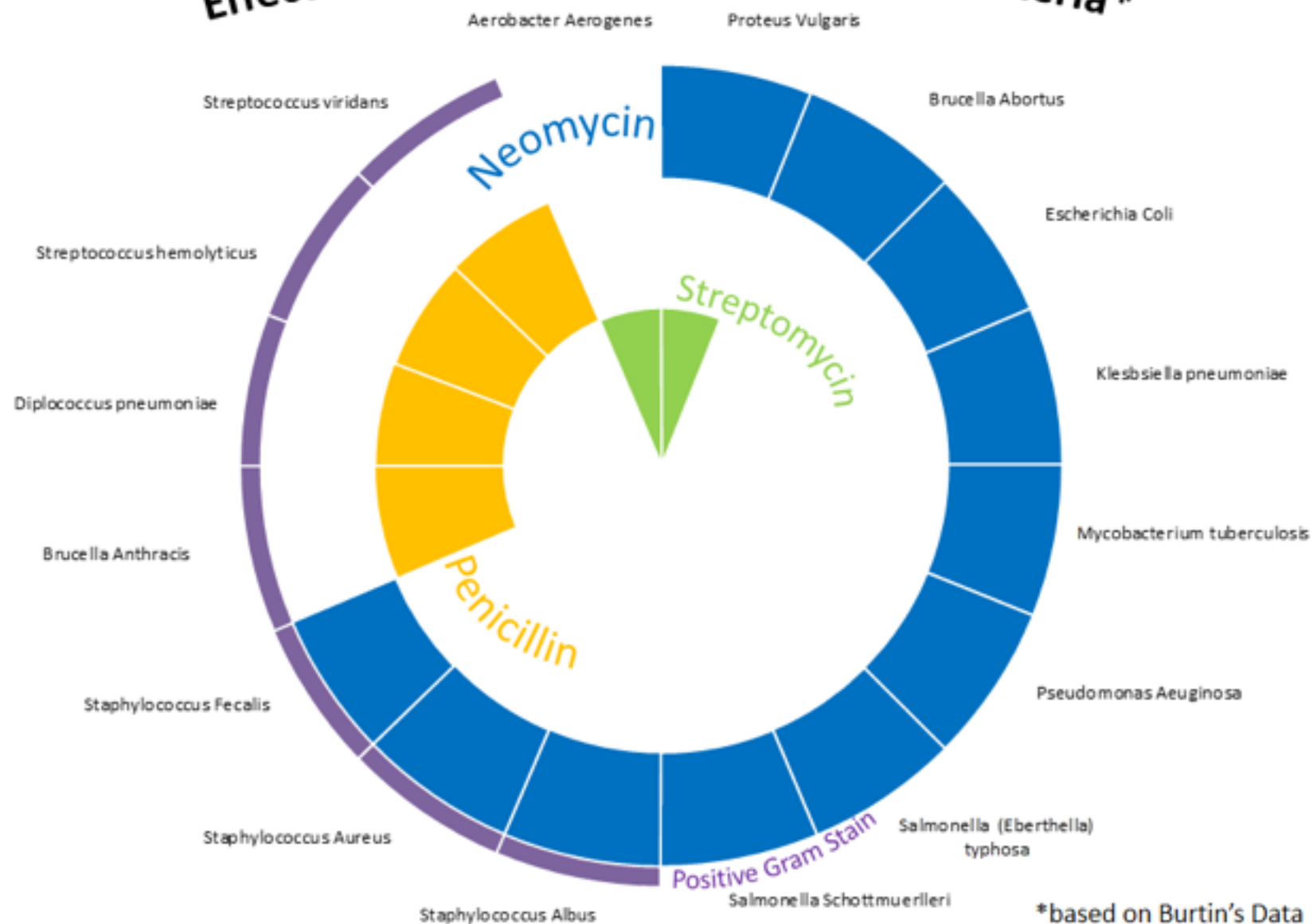


Dosage Needed to Combat Bacteria

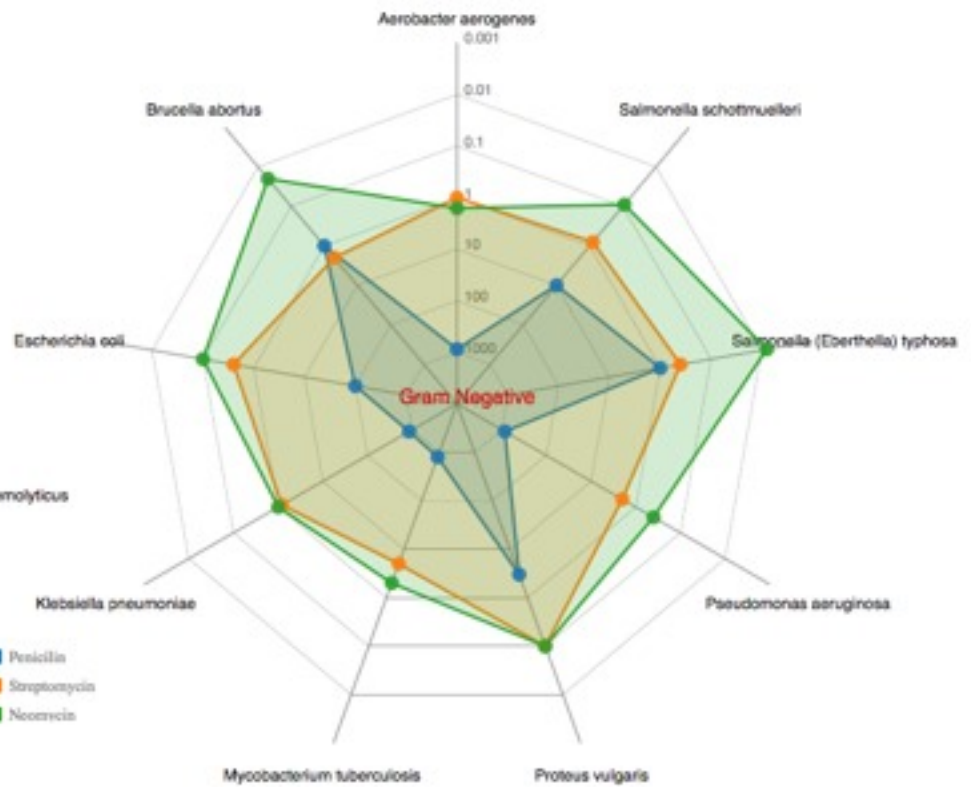
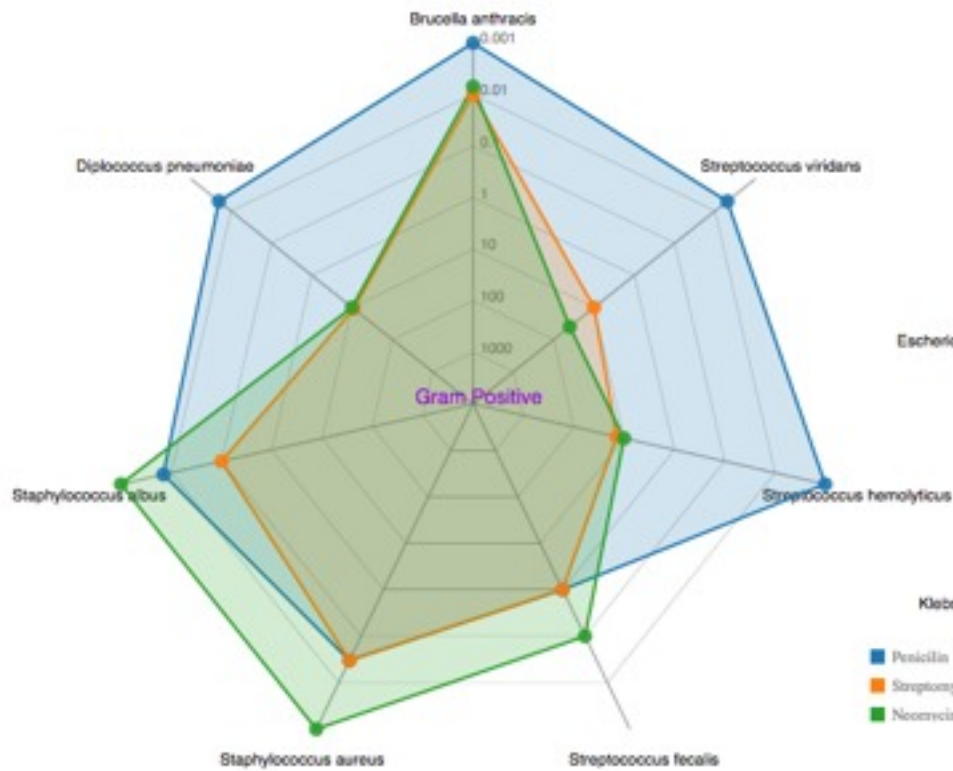
Size of wedge shows amount of an antibiotic needed to suppress bacterial growth
As Minimum Inhibitory Concentration (MIC), plotted on a log scale



Effective Antibiotics Against Various Bacteria*

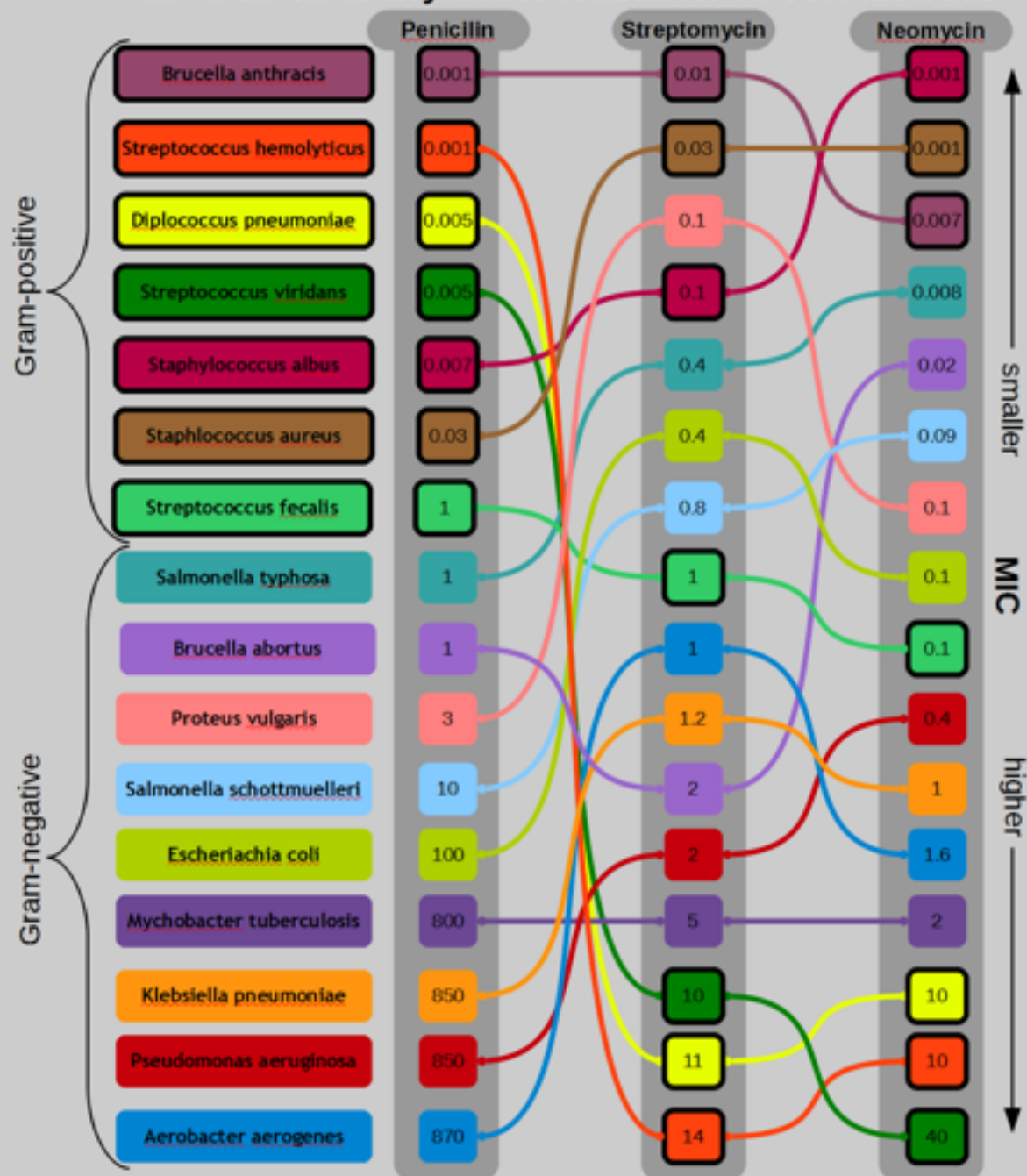


*based on Burtin's Data



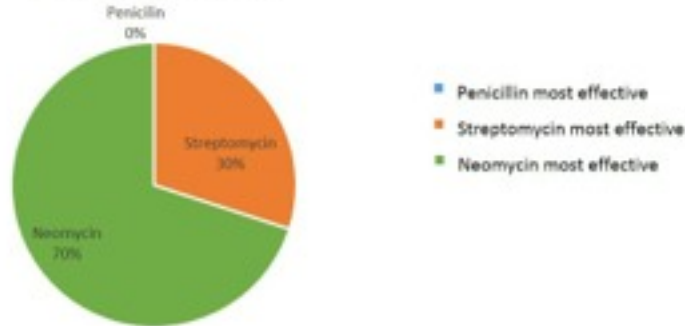
Minimum inhibitory concentration of Antibiotics

Minimum Inhibitory Concentration for Three Antibiotics

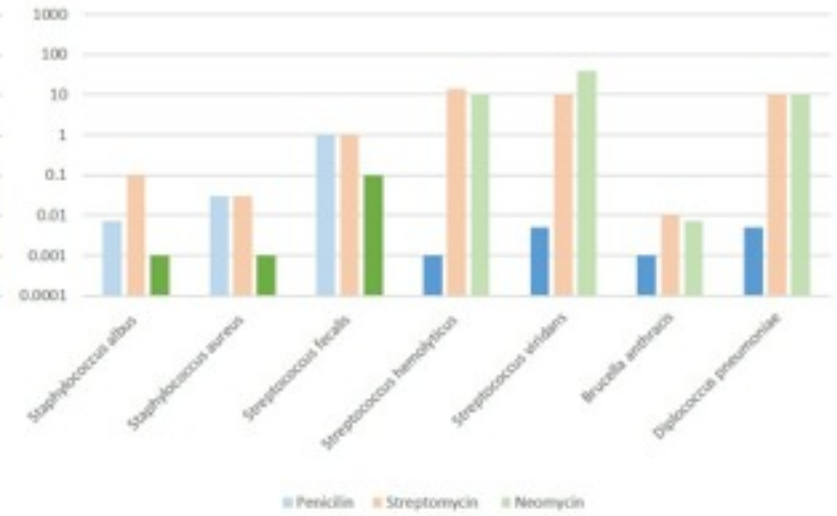
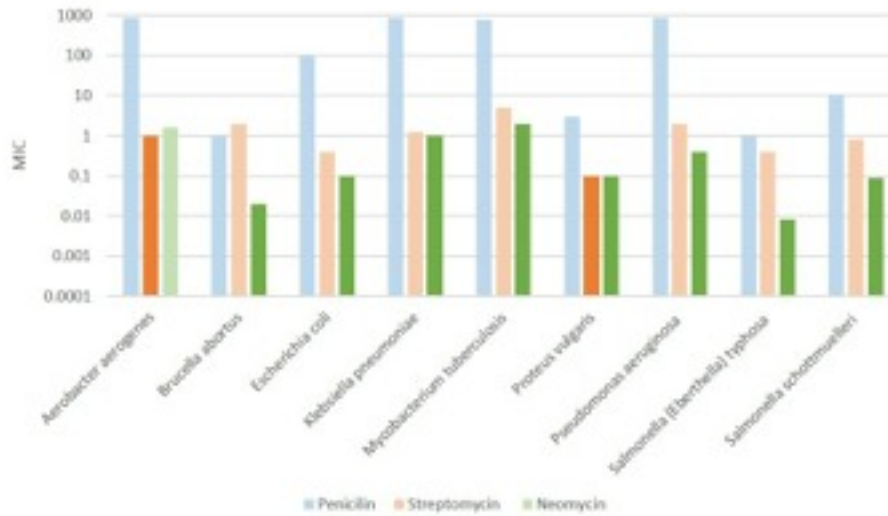
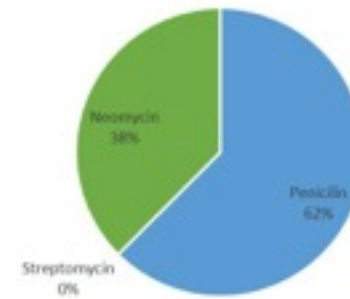


Effectiveness of Antibiotics against Bacteria

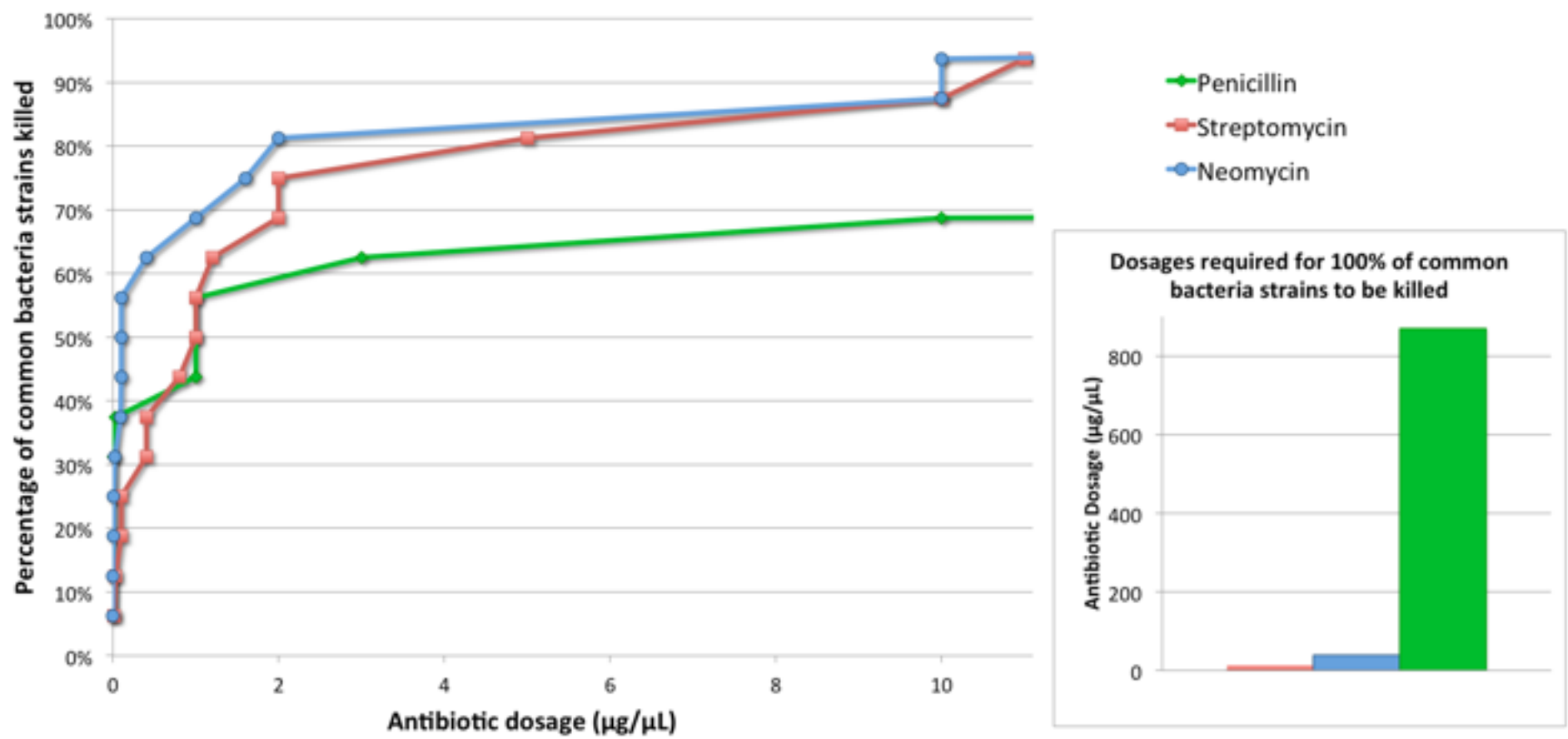
Gram Staining Negative



Gram Staining Positive

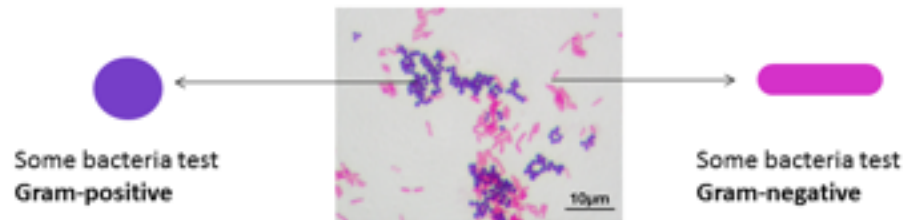


Effectiveness of antibiotics on 16 common bacteria strains



Which antibiotic should I take?

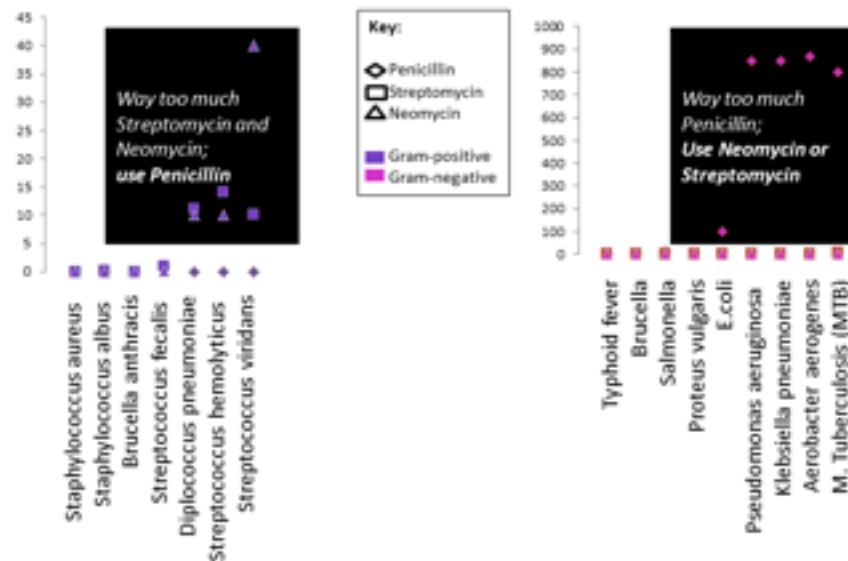
Even if you know you have a bacterial infection, it's hard to tell what kind. You need to know that to know which antibiotic is most effective. Your doctor can use the "Gram" stain test and look at the bacteria:



Some antibiotics work well for some bacteria and some don't.

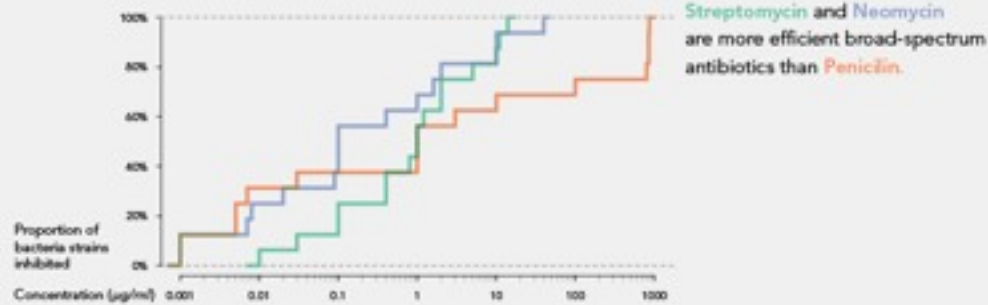
For example, to kill **Tuberculosis** bacteria, you would need to use a *lot* (400 times!) more Penicillin than Neomycin. Antibiotics are blind: they kill many organisms in your body, not just the bad bacteria. So, it would be much better for your body to use Neomycin, and it would still work. Some of the **Gram-negative** bacteria need a lot more Penicillin, so **Streptomycin** and **Neomycin** is better.

How much antibiotics need to be applied to kill **different kinds of bacteria?**

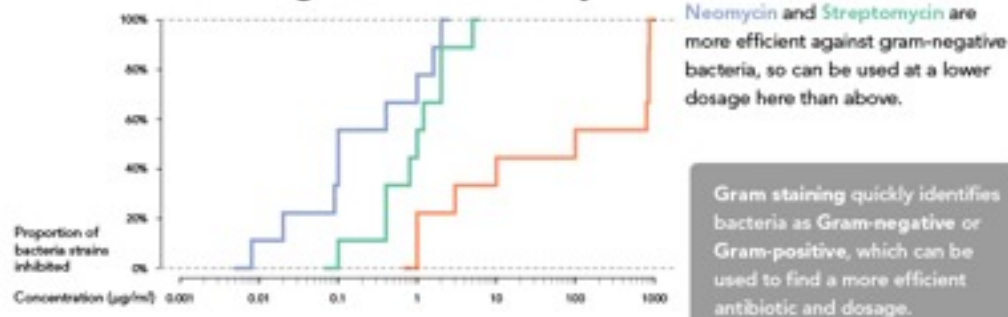


Dose efficiency of three antibiotics against —

All bacteria

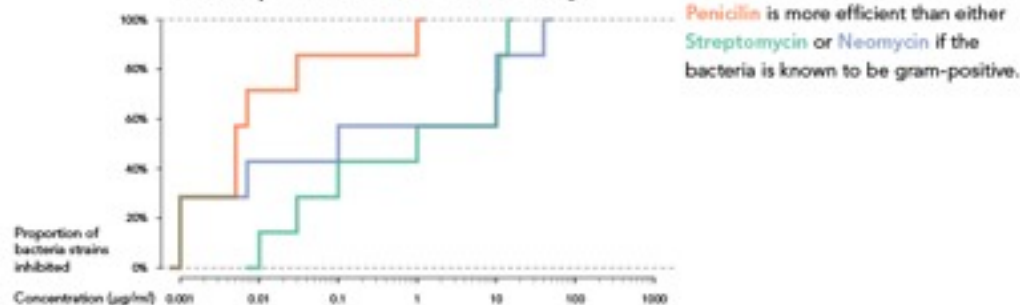


Gram-negative bacteria only



Gram staining quickly identifies bacteria as Gram-negative or Gram-positive, which can be used to find a more efficient antibiotic and dosage.

Gram-positive bacteria only



Visualization Design

In-Class Exercise

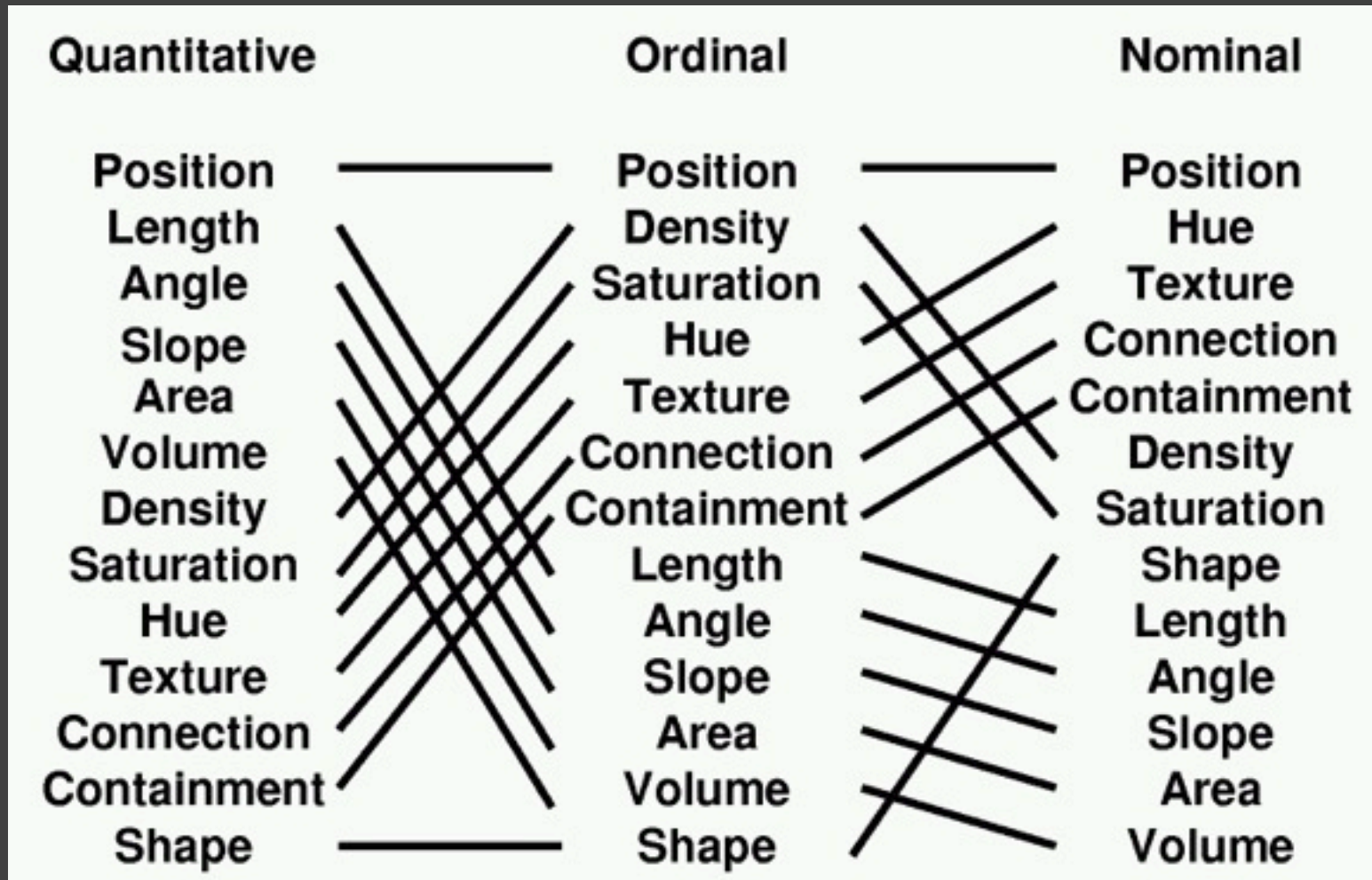
In-Class Design Exercise

Task: Analyze and Re-design visualization

- Identify data variables (n, o, q) and encodings
- Critique the design: what works, what doesn't
- Sketch a re-design to improve communication
- Be ready to share your thoughts with the class

Break into groups with those sitting near you
(~4 people per group)

Mackinlay's Ranking



Conjectured *effectiveness* of the encoding

Teacher Salaries: Is It Really That Bad?

National and State averages for K-12 Public-School Teachers



UNITED STATES

Avg. salary: \$47,874

Avg. vacation days: 63

HOURLY

Hours per week on-site: 36.5
Public-School Teacher: \$34.06
Private-School Teacher: \$21.08
Average Worker: \$25.08
Police: \$22.64
Fire: \$17.90



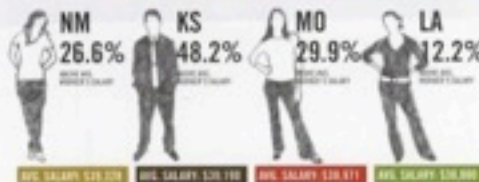
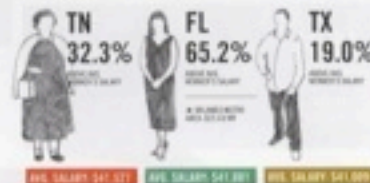
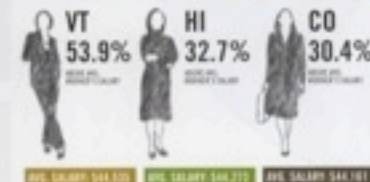
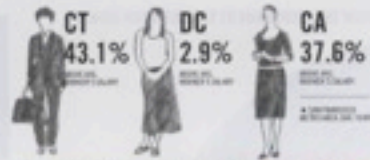
CANADA

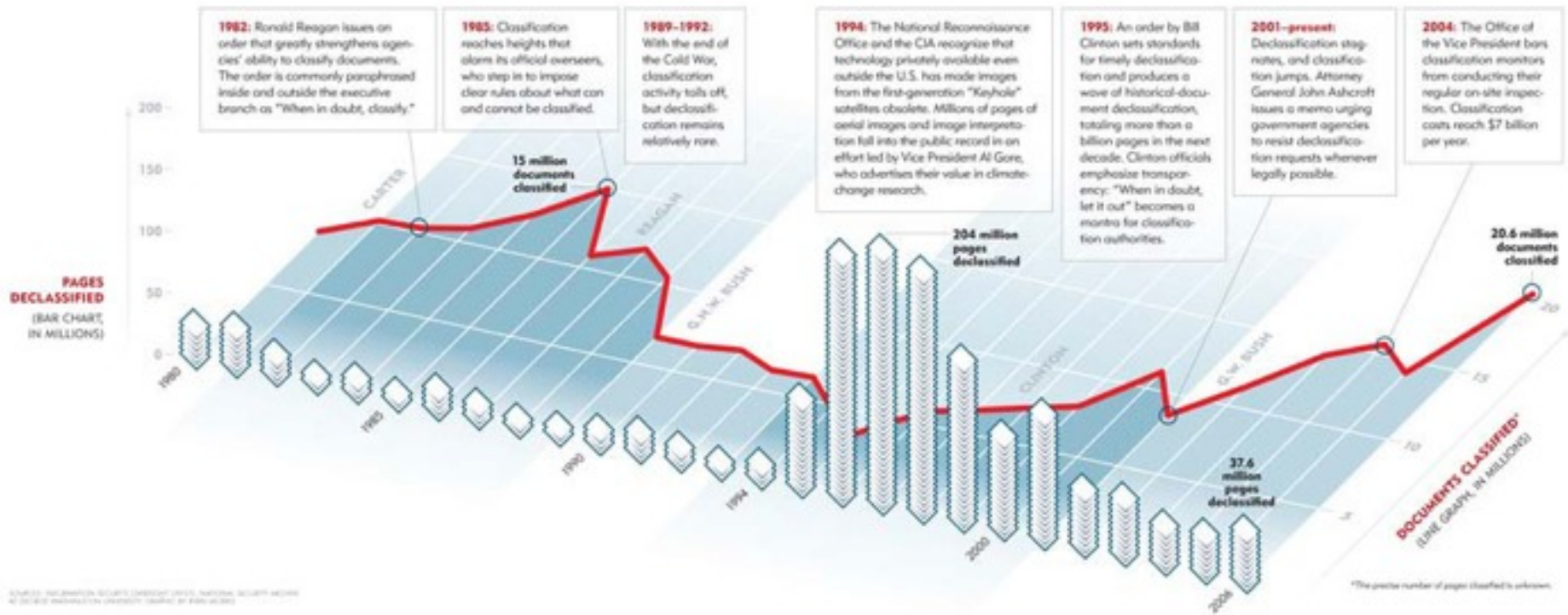
Avg. salary: \$47,800

Avg. vacation days: 50

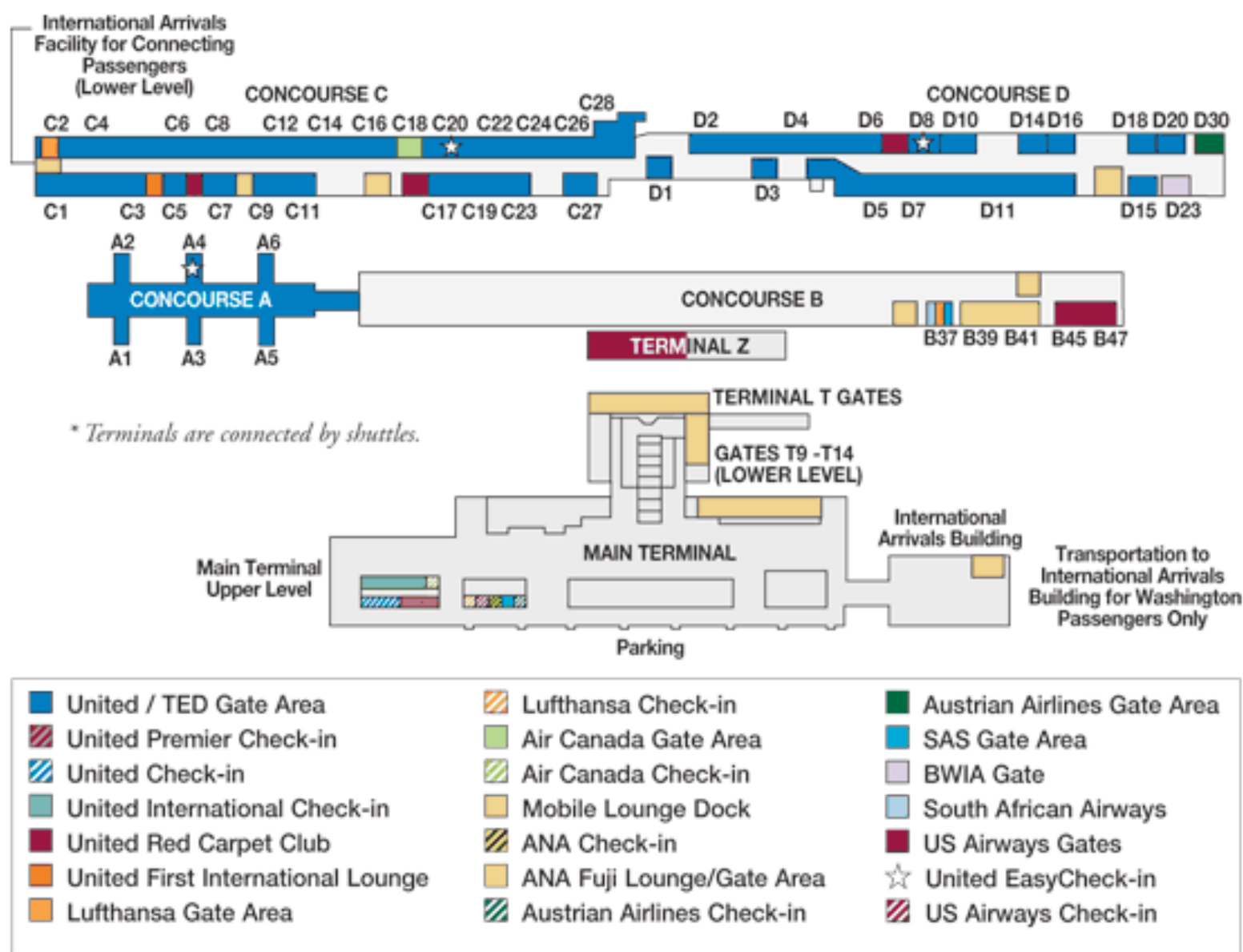
HOURLY

Hours per week on-site: 35.6





Source: *The Atlantic* 300 no. 2 (September 2007)
Number of Classified U.S. Documents

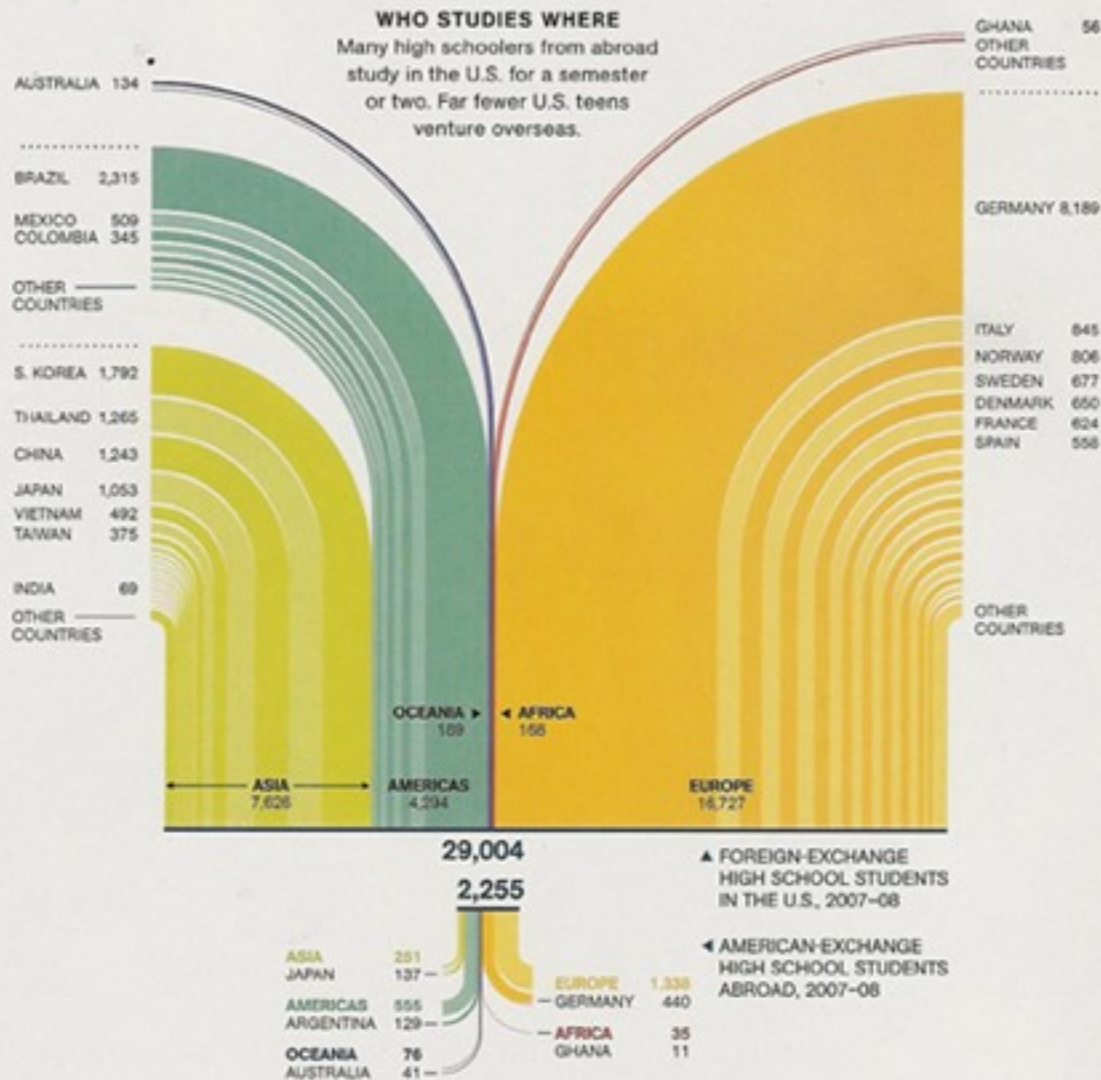


EasyCheck-in is available at this airport.



11 2006

Washington Dulles Airport Map
Source: United Airlines Hemispheres



Source: *National Geographic*, September, 2008, p. 22.
 Silver, Mark. "High School Give-and-Take."

IT WAS A VERY GOOD YEAR?

Robert Parker's ratings for vintages of Napa Valley cabernet sauvignon



RATINGS

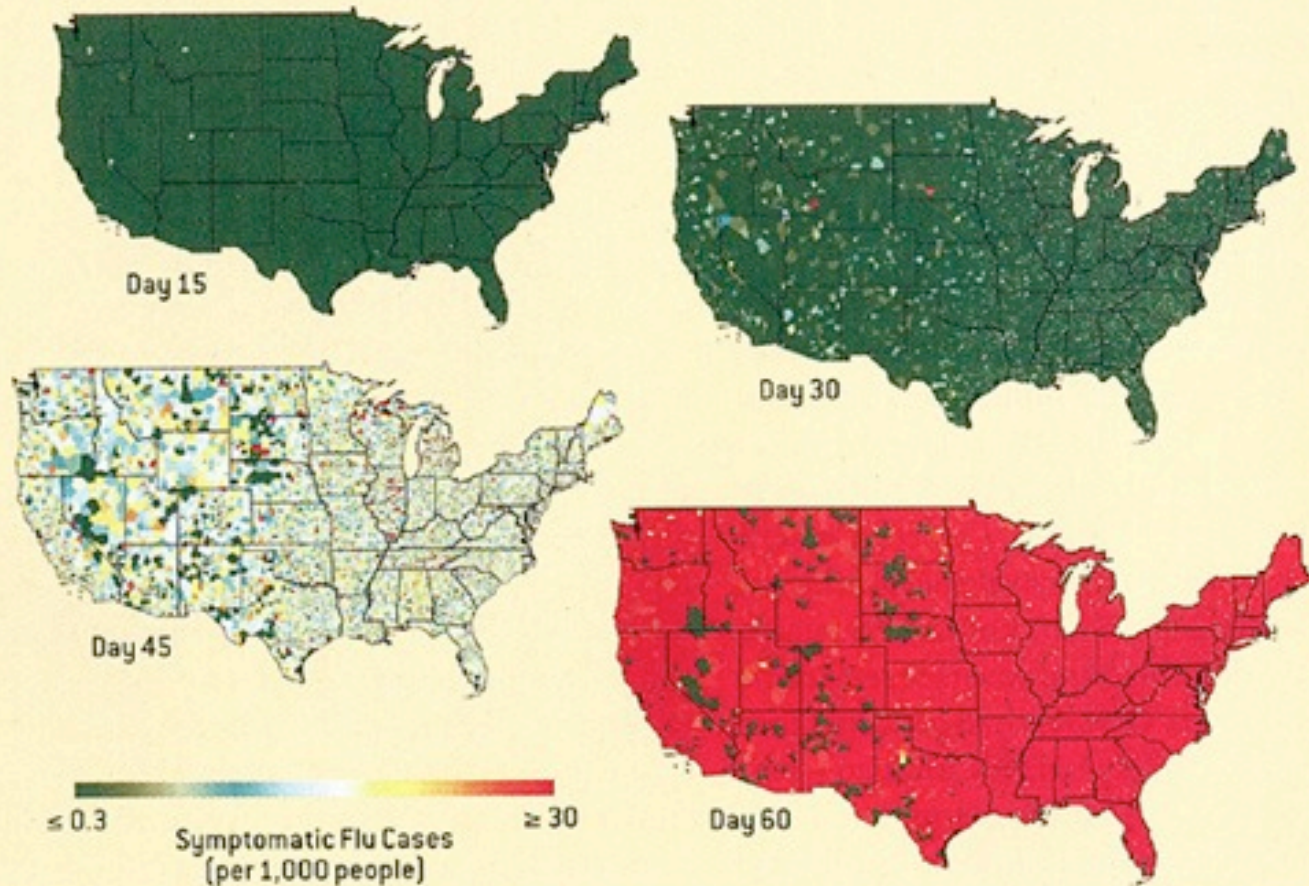
96-100	Extraordinary
90-95	Outstanding
80-89	Above average
70-79	Average
60-69	Below average
<59	Unacceptable

T= Still tannic, youthful, or slow to mature
R= Ready to drink
E= Early maturing and accessible
I= Irregular
C= Caution, may be too old

2001
96
points. It was a relatively modest year in terms of yield from the vineyards, and that worked to the vintner's advantage. The results: some of Napa's most concentrated, structured, long-lived wines. Built for aging, they are rich, densely colored,

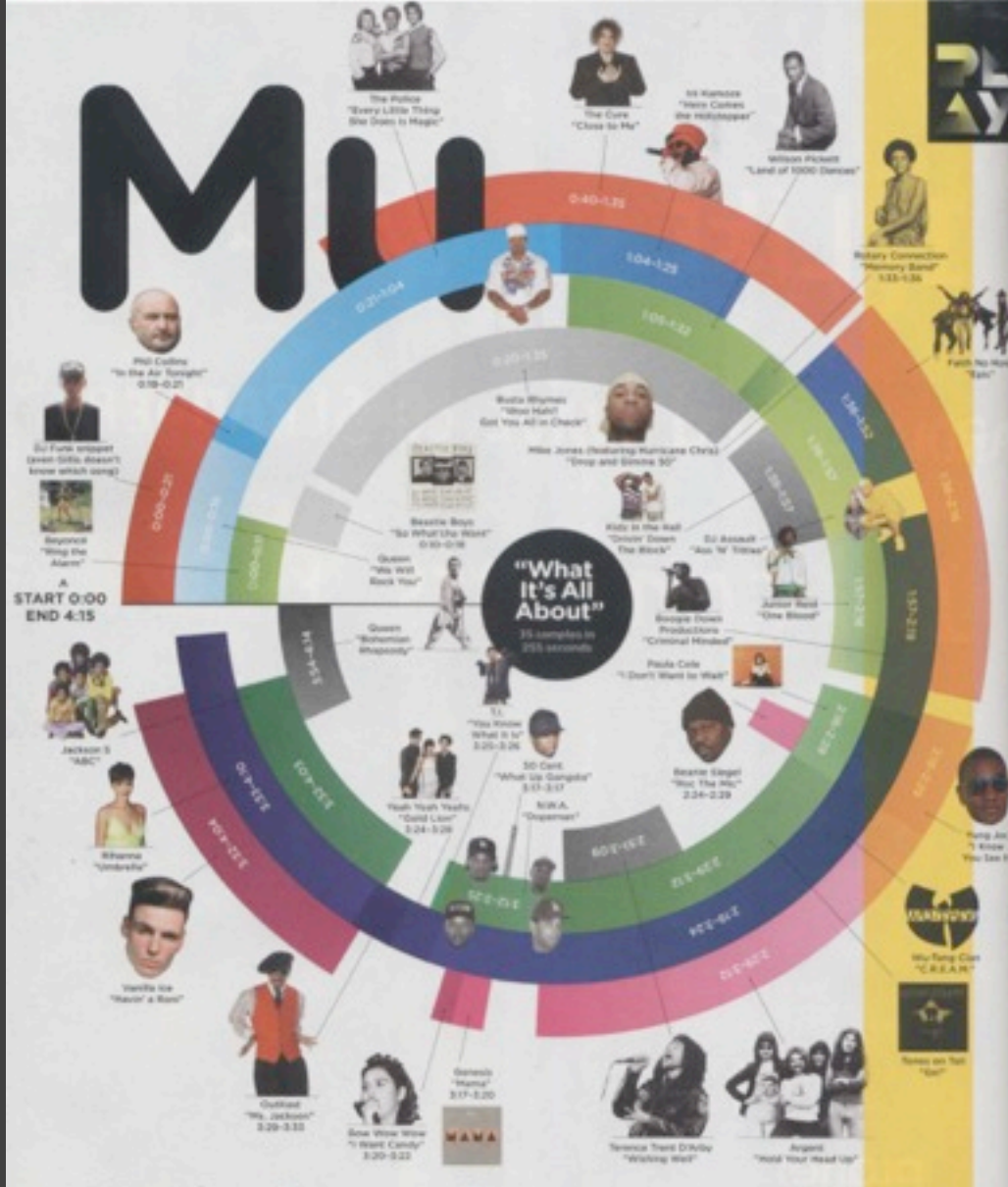
Pandemic Flu Hits the U.S.

A simulation created by researchers from Los Alamos National Laboratory and Emory University shows the first wave of a pandemic spreading rapidly with no vaccine or antiviral drugs employed to slow it down. Colors represent the number of symptomatic flu cases per 1,000 people (see scale). Starting with 40 infected people on the first day, nationwide cases peak around day 60, and the wave subsides after four months with 33 percent of the population having become sick. The scientists are also modeling potential interventions with drugs and vaccines to learn if travel restrictions, quarantines and other disruptive disease-control strategies could be avoided.



Preparing for a Pandemic

Source: *Scientific American*, 293(5). November, 2005, p. 50



Source: *Wired Magazine*, September 2008 Edition
 Music: Super Cuts (page 92)