

Visualization of Public Health Data

Anamaria Crisan

PhD Student at UBC in Computer Science

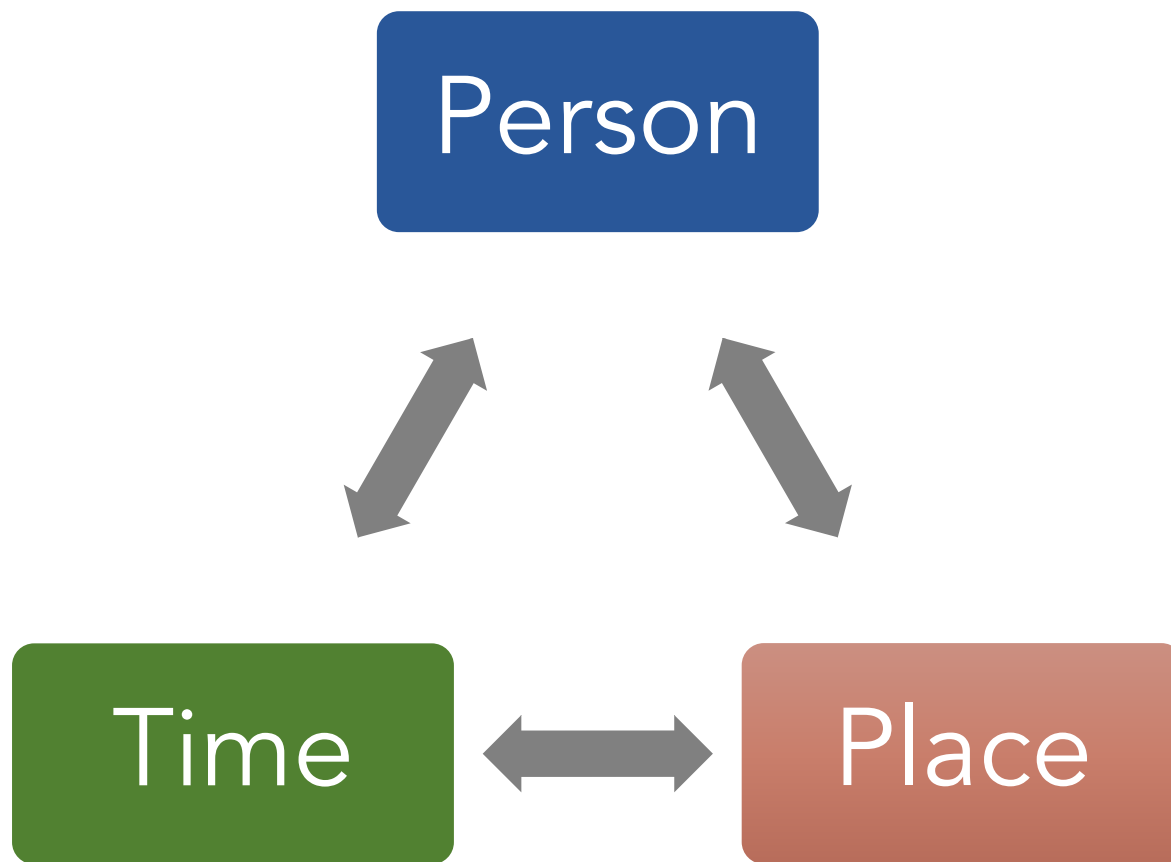
Supervisors:

Jennifer Gardy , Population and Public Health

Tamara Munzner, Computer Science

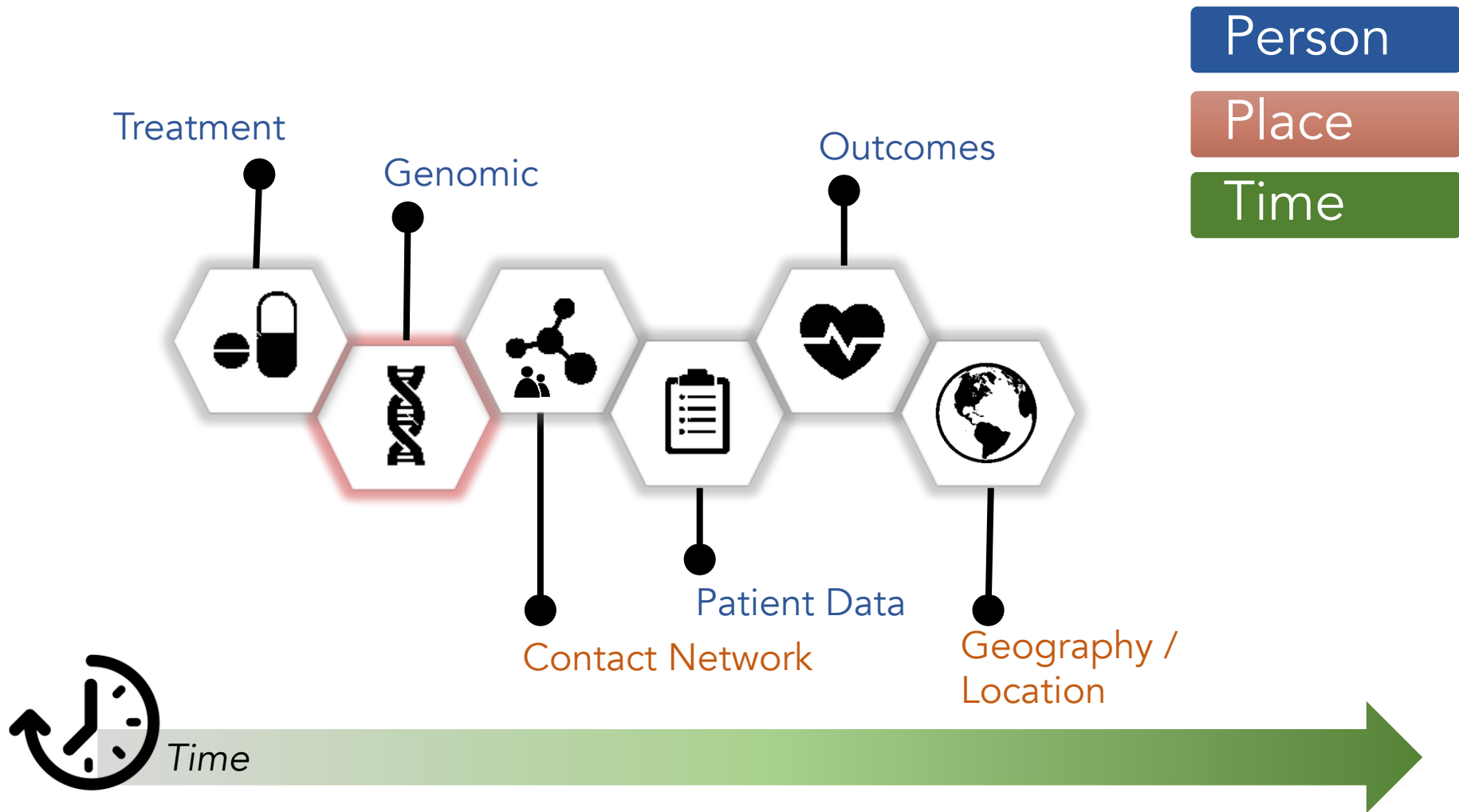
WHAT ARE PUBLIC HEALTH DATA?

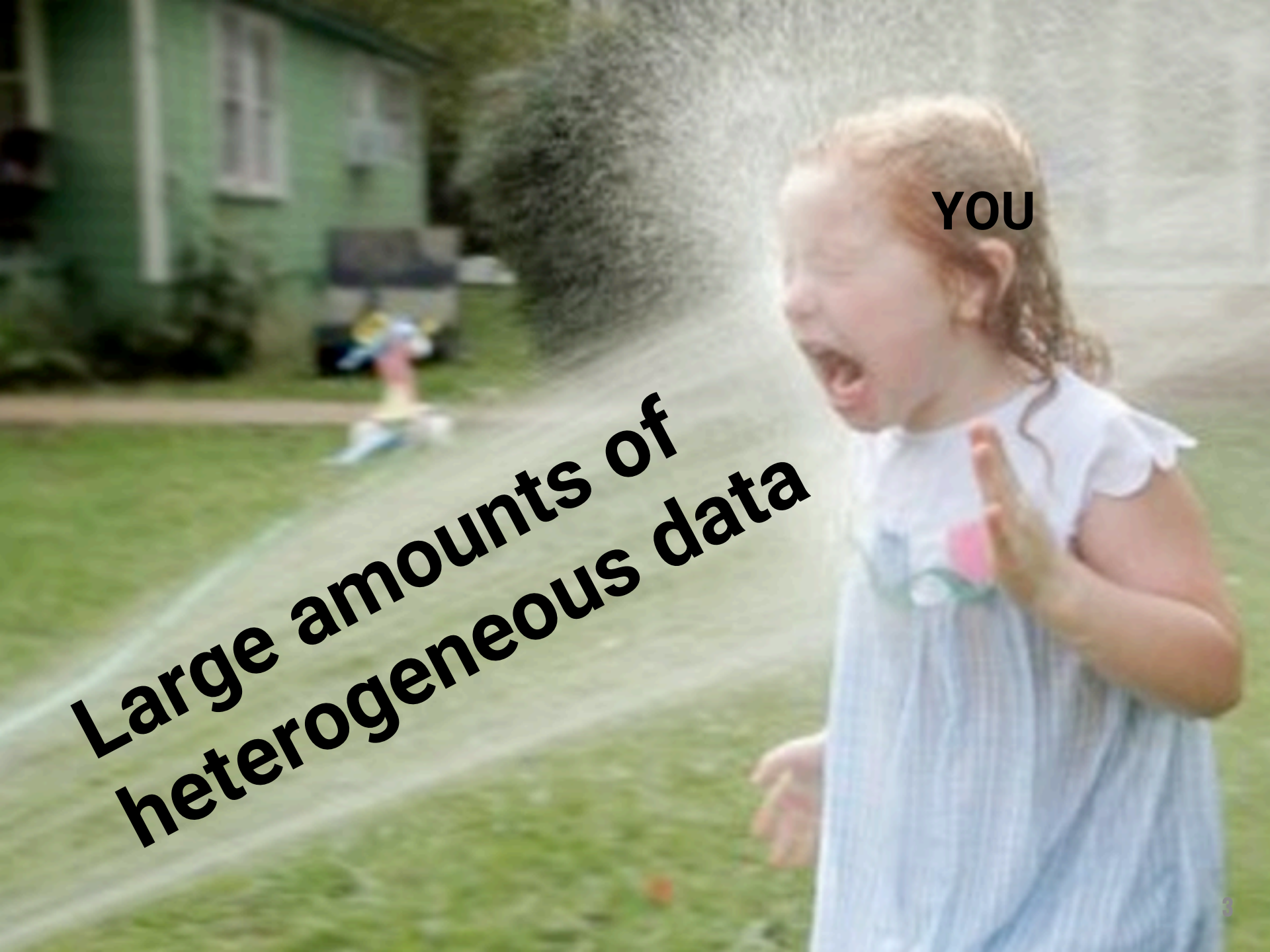
(FOR INFECTIOUS DISEASE MANAGEMENT)



WHAT ARE PUBLIC HEALTH DATA?

(FOR INFECTIOUS DISEASE MANAGEMENT)





YOU

**Large amounts of
heterogeneous data**

SUPPORT FOR DATA DRIVEN DECISIONS

- Public health has multidisciplinary decision making teams
 - More data & diverse data types = more informed decision making
 - BUT - not all stakeholders can interpret / understand data
- Support needed for decision making with heterogeneous data

Medical Health Officers



Clinicians



Nurses



Researchers



Community Leaders



PROPOSAL

Visualization of public health data **can improve knowledge sharing and decision making** in infectious disease prevention and control

WHY VISUALIZATION?

Least Understandable

Most Understandable

Probability

<

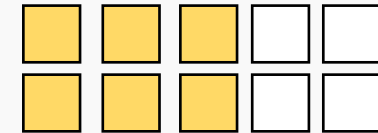
Frequency

<

Visualization

60%

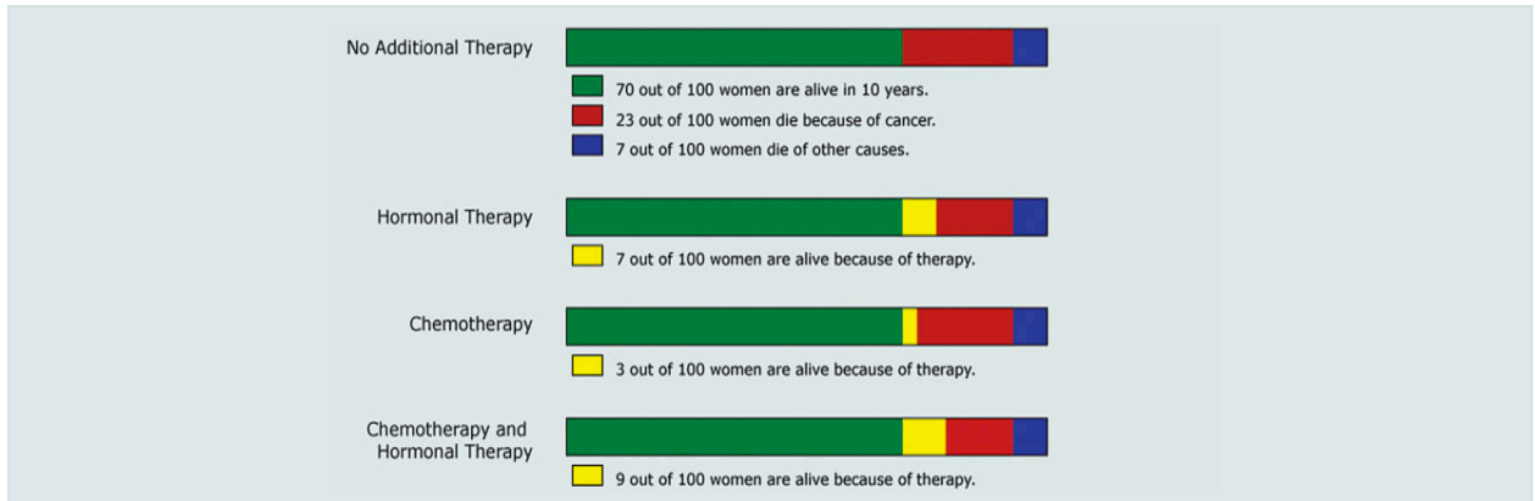
6 in 10



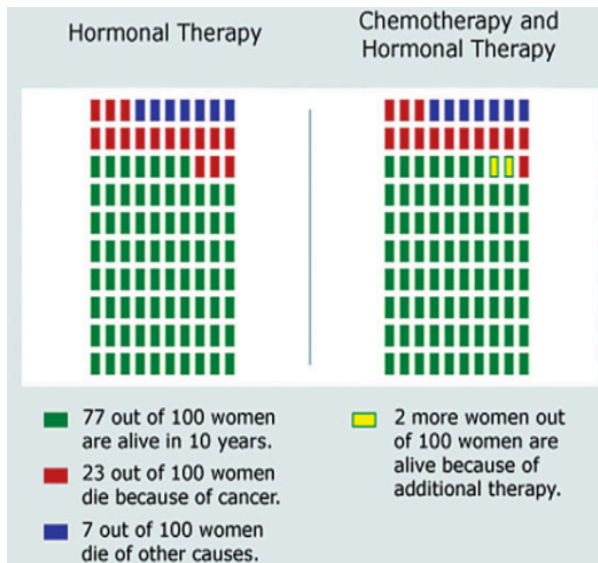
- Numeracy : the ability to reason with numbers
 - Individuals with low numeracy have a difficulty interpreting numbers and probabilities
 - Also true amongst educated professionals
- **Visualization can make data more accessible to diverse stakeholders on decision making teams**

BUT! VISUAL DESIGN ALSO MATTERS

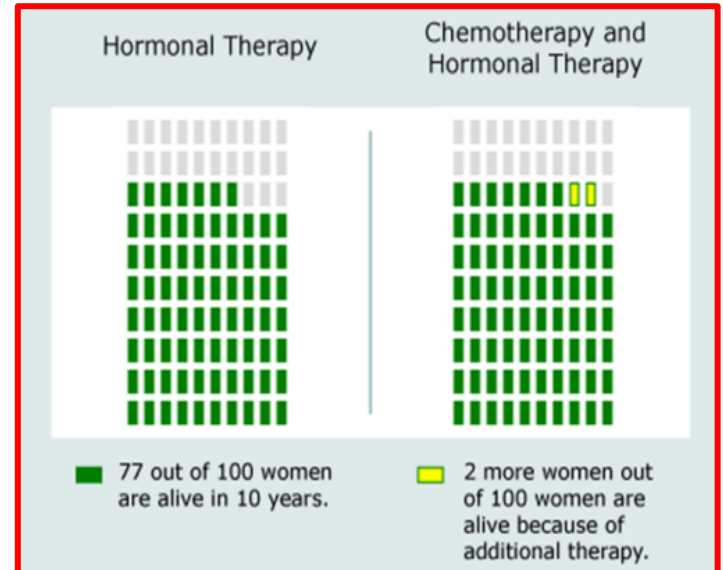
Baseline Visualization



Alternative 1



Alternative 2



EXAMPLE OF GUIDANCE : WWW.VIZHEALTH.ORG

VISUALIZING HEALTH

A Scientifically Vetted Style Guide for Communicating Health Data



About

The Wizard

Browse the Gallery

Using Visualizing Health

Disclaimer

Download Report

My goal

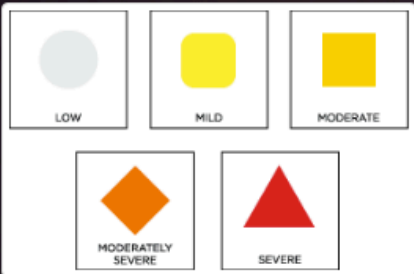
- Classifying risks
- Raise or lower concern
- Awareness of risk
- Differences in likelihood
- Risk tradeoffs

Details or gist?

- Verbatim recall
- Gist understanding

Data I have

- Benefit estimate
- Risk over time
- Case counts

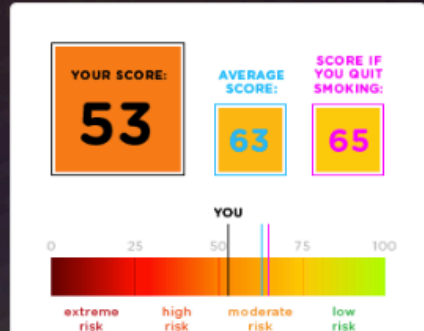


(84) Icons to show severity of side effects

...MORE LIKE THIS

Classifying risks

ALL TAGS



(43) Visualizing health scores

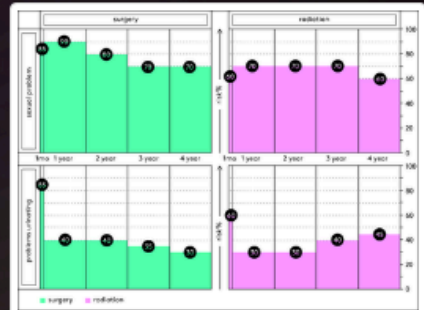
...MORE LIKE THIS

Classifying risks

Raise or lower concern

Awareness of risk

ALL TAGS



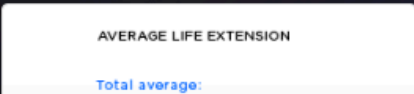
(66) Showing how side effects change over time

...MORE LIKE THIS

Risk tradeoffs

Differences in likelihood

ALL TAGS



APPLICATION TO PUBLIC HEALTH

- Lots of interest in Visualization in Public Health
- But - mainly developing *ad hoc* solutions
 - Visualization designers usually bioinformaticians (high numeracy, lack stakeholder context)
 - Stakeholders relying on Excel for visualizations
- **Need to make a case for better visualizations**
- **Need to treat data visualization as a research process**

VISUALIZATION DESIGN & ANALYSIS

Steps for visual design

1. Partner with a group of stakeholders that have a **problem**

VISUALIZATION DESIGN & ANALYSIS

Steps for visual design

1. Partner with a group of stakeholders that have a **problem**
2. Ask **what data** stakeholders use (is it available)?

VISUALIZATION DESIGN & ANALYSIS

Steps for visual design

1. Partner with a group of stakeholders that have a **problem**
2. Ask **what data** stakeholders use (is it available)?
3. Ask **what stakeholders do** with the data [**tasks**]

VISUALIZATION DESIGN & ANALYSIS

Steps for visual design

1. Partner with a group of stakeholders that have a **problem**
2. Ask **what data** stakeholders use (is it available)?
3. Ask **what stakeholders do** with the data [**tasks**]
4. Explore if other visualizations have addressed this **problem** and set of **tasks**

VISUALIZATION DESIGN & ANALYSIS

Steps for visual design

1. Partner with a group of stakeholders that have a **problem**
2. Ask **what data** stakeholders use (is it available)?
3. Ask **what stakeholders do** with the data [**tasks**]
4. Explore if other visualizations have addressed this **problem** and set of **tasks**
5. Test multiple alternatives (including new ones you develop) with stakeholders


VISUALIZATION DESIGN & ANALYSIS

Steps for visual design

1. Partner with a group of stakeholders that have a **problem**
2. Ask **what data** stakeholders use (is it available)?
3. Ask **what stakeholders do** with the data [**tasks**]
4. Explore if other visualizations have addressed this **problem** and set of **tasks**
5. Test multiple alternatives (including new ones you develop) with stakeholders
6. Gather qualitative & quantitative evaluation data

VISUALIZATION DESIGN & ANALYSIS

Steps for visual design AN ITERATIVE PROCESS

1. Partner with a group of stakeholders that have a **problem**
 2. Ask **what data** stakeholders use (is it available)?
 3. Ask **what stakeholders do** with the data [**tasks**]
 4. Explore if other visualizations have addressed this **problem** and set of **tasks**
 5. Test multiple alternatives (including new ones you develop) with stakeholders
 6. Gather qualitative & quantitative evaluation data
- 

EXAMPLE: TB GENOMIC CLINICAL REPORT

Current Report

Mycobacterium Whole Genome Sequencing Report from MGIT Positive Samples

Not for diagnostic use

11/12/2015

Sample Details

Sequencing Location	Oxford	Date received in Lab	
Local LIMS	12.0610882	Run date	01/01/20151008
Specimen ID			
Guid	b7aa98e0-3612-4c0b-a47b-471e0e78c72d		

Organism Identification

Predicted/closest match	
TBCOMP/microti	100%
TBCOMP	100%
TBCOMP/TB	96.77%
TBCOMP/tuberculosis-canetti	35.71%
MACCOMP	21.21%

Sample/Sequencing Quality

Total reads (~millions)	Mapped %	No reads mapped (~millions)	Coverage %
4.73	99.47	4.7	91.99

Resistance Summary

INH	RIF	EMB	PZA	QUI	SM	AG
U	S	S	S	S	S	S

Resistotype

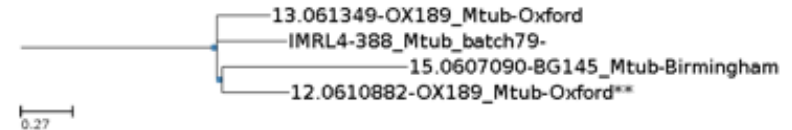
Drug	Mutation	Nucleotides	Support (ACGT)	Source - (R/Total)	Prediction
INH	katG_A727T	GCC->ACC	(160/0/1/0) (0/164/0/0) (0/167/0/0)	Unclassified	UNK

Relatedness

NB: This data may be added or updated at a later date
Nearest neighbour(s)

Sample -Plate Name	Date received in Lab	Centre	No. of SNPs apart
12.0610882-OX189_Mtub		Oxford	0
IMRL4-388_Mtub_batch79	1900-01-01		13
15.0607090-BG145_Mtub	2015-05-26	Birmingham	16
13.061349-OX189_Mtub		Oxford	8

The alignment width is 285. Multiply this number by the tree metrics.



Comments



Authorised

Signature:	Print name:
Position:	Date:

DESIGN PROCESS OVERVIEW

Question: Can we improve upon the existing report design

Note: Not a data vis project, but uses data vis methods and result will feed into other data vis projects

Phase 1: Expert consultations

Phase 2: Task Questionnaire

Design Sprint

Phase 3: Design choice Questionnaire

Phase 4: Evaluation of final report design

DESIGN PROCESS OVERVIEW

Question: Can we improve upon the existing report design

Note: Not a data vis project, but uses data vis methods and result will feed into other data vis projects

Phase 1: Expert consultations

Phase 2: Task Questionnaire

Design Sprint

Phase 3: Design choice Questionnaire

Phase 4: Evaluation of final report design

PHASE 1

EXPERT CONSULTATIONS

Participants: 7 = physicians (clinical & laboratory), public health researchers

Key Findings

- Different needs between physicians and researchers
- Physicians had greater time pressure
- Trust in lab and procedures
- Some data on report not necessary, other data confusing
- Constraints on delivery report due EHR

PHASE 2

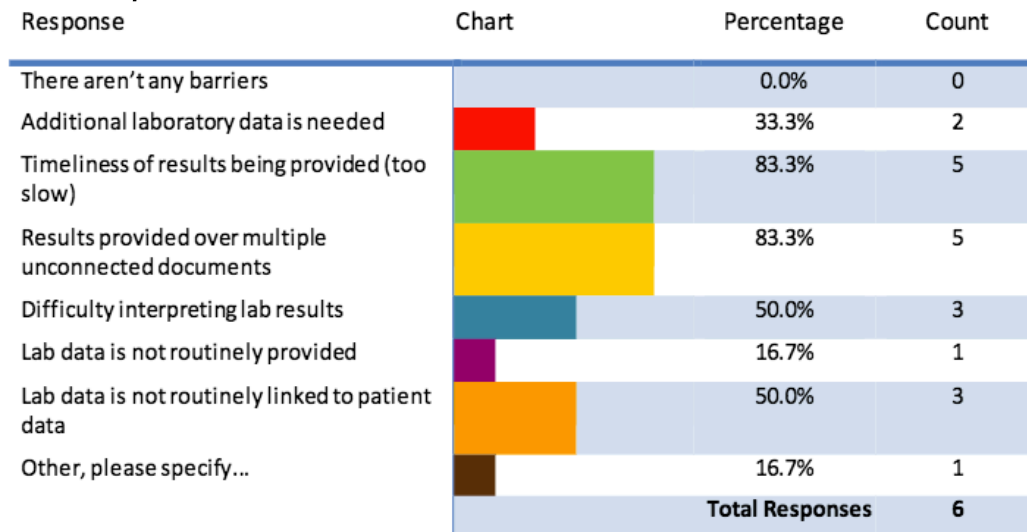
TASK QUESTIONNAIRE

Participants: 17 = physicians (clinical & laboratory), nurses, public health researchers, surveillance experts

Key findings

- Quantitative support for earlier qualitative findings
- Better granularity of data used, and confidence performing, different tasks

Q: *What could improve the efficiency of using molecular data?*



PHASE 2

DESIGN SPRINT



PHASE 2

DESIGN SPRINT

OVERVIEW

- NAME
- ID
- EXECUTIVE SUMMARY
- DATE SENT OUT
- DATE RECEIVED
- ORIG
- DEPT
- CLUST#

DIAGNOSIS

BLURB

ORGANISM

Species name (redia, redemptory, etc.)

T. ALT ORGANISM

ANY TB COVD PRO.

TREATMENT

BLURB

- DRUG TYPE (FULL NAME)
- SENSITIVE, RESISTANCE
- SUMMARY (MONO, N-DR, XDR)
- Q: HOW SHOULD BE SEEN

CLUSTER MEMBERSHIP

RESISTOTYPE → # OF SNPS? X

- Mutations → some well characterized → others poor → support maybe.

CLUSTER & GENOMIC INSIGHTS

BLURB

- lineage
- SNP threshold
- 0-5: part of cluster
- 6-12: peripheral to cluster but may be related.
- 12+: not related

RETAINING LIST WITH 20 MUTATIONS

- SUMMARISE CLUSTER MEMBERSHIP
- WHERE: } Nearest Neighbour
- HOW BIG

SAMPLE DATA

BLURB

- DATE OF REPORT RELEASE
- DATE OF SAMPLE COLLECTION
- LAB ID (Oxford)
- SAMPLE LIMS
- GWID
- SAMPLE & SEQUENCE QUALITY
- REFERENCE (GENOME PIPELINE ID)
- RESISTOTYPE RESULT

Title **date**

Sample Details

Sample Quality

Resistance

UNKNOWN	EMB
RESISTANT	INH
1	2
2	3
3	4

if there is any →

look by white text

Origin ID

Microb. Tub.

Resistance Summary

Sensitive	Resistance	Unknown
RIF SM	INH	EMB
BUL AG	XY	

* For more info, see ... in the next page

Relationships

Always Related	# Likely Related
10	
11 of 100 up to 0-5	6-30

small font

1/2

2/2

Comments

Amplified

Signature:

Drum:

Price class:

Same:

WISH IDENTIFICATION DETAILS

NAME

SAMPLE ID

DATE

LOCATION

RESIST

% IDENTITY

100%

91%

87%

SENSITIVITIES

RESIST

UNKNOWN

EMB

RESISTANT	INH
1	2
2	3
3	4

DETAILS

SUMMARY

RESISTANT

CLUSTERING DETAILS

SNP

SNP DISTANCE

1

5

SNPs DETAILS

DETAILS

DATE

July 21, 2016

PAGE X OF Y

PHASE 3 DESIGN CHOICE QUESTIONNAIRE

Participants: 42

Goal: Compare control (existing report) with options developed in the design sprint

Tuberculosis Sequencing Results Page 1 of 2

Patient Information

Patient Name	Bob Johnson	Sample Type	Sputum
Patient ID	123456789	Sample Site	-
Patient DOB	01-01-1950	Sample Date	01-01-2020
Location	Default	Specimen ID	123456789

Summary of Findings

Based upon an analysis of the specimen's genomic data, this patient has **mycobacterium tuberculosis** that is predicted to be resistant to 2 antibiotics: **Isoniazid**, **Rifampin**. This case belongs to a cluster of cases with similar genomic findings.

Diagnosis

Metabology genomic data from the specimen was compared to mycobacterium and non-mycobacterium reference genomes for species where available.

The specimen was classified as **mycobacterium tuberculosis**.

Treatment

Metabology drug sensitivities were predicted using the genomic sequence data in accordance to the method reported in published literature.

The specimen was considered to be **multi-drug resistant (MDR) TB**.

Summary of sensitive findings:

Drug	Prediction	Status	Comment
Isoniazid	Resistant	2	Gene: katG, Antic Acid Changes: S315T
Rifampin	Resistant	2	Gene: rpoB, Antic Acid Changes: S531L
Ethambutol	Sensitive	✓	-
Pyrazinamide	Sensitive	✓	-
SM	Sensitive	✓	-
AD	Sensitive	✓	-

Page 1 of 2

Tuberculosis Genome Sequencing Results Page 2 of 2

Epidemiologic Summary

Metabology Patients are automatically assigned to clusters based upon their genotype relative to phylogenetic differences. Clustering thresholds are defined according to [this reference page](#).

The specimen belongs to a **previously existing cluster**.

Similarity	SNP Difference	Cluster (over past 5 years)	Members (Known)
Highly	0 to 5		2
Peripheral	6 to 12		6

Quality Summary

The whole genome sequence analysis of the isolate was considered **HIGH QUALITY** as the number of reads was greater than 6.7 million with 99.43% mapped and a coverage of 81.99%.

Comments

Advanced for 1st Ref 2

Authorised by: Dr. John Smith, Laboratory Director, Signature Date: 01-01-2020

Page 2 of 2

Mycobacterium Whole Genome Sequencing Report

Report Date: 01-01-2020
Laboratory: Oxford
Reviewed by: Dr. John Smith

Patient Details

Patient Name	Bob Johnson	Requester	Dr. Paul Giddens, Dr. Birmingham, UK
Patient ID	123456789	Copy to	-
Patient DOB	01-01-1950		
Location	Default		

Sample Details

Sample Type	Sputum	Sample Date	01-01-2020
Sample Site	-	Specimen ID	123456789

Speciation

Reported Species: **Mycobacterium tuberculosis**

Drug Sensitivities

Drug	Prediction	Status	Comment
Isoniazid	Resistant	2	-
Rifampin	Resistant	2	-
Ethambutol	Sensitive	✓	-
Pyrazinamide	Sensitive	✓	-
SM	Sensitive	✓	-
AD	Sensitive	✓	-

Relatedness

Number of isolates	likely related (more than 5 SNP difference)	possibly related (6-50 SNP difference)
2	2	6

For further information on related isolates and existing clusters, please contact the Public Health lab at 012-456-1999

Resistotype

Drug	Prediction	Gene	Mutation
Isoniazid	Resistant	katG	S315T
Rifampin	Resistant	rpoB	S531L

Sequence Quality

The whole genome sequence analysis of the isolate was considered **HIGH QUALITY** as the number of reads was greater than 6.7 million with 99.43% mapped and a coverage of 81.99%.

Reviewer Comments

No additional comments

Authorization

Signature	Print Name	Dr. John Smith
Bob Johnson	Pauline	Lab Director

Mycobacterial Genome Sequencing Results Page 1 of 2

Patient Information

Patient Name	Bob Johnson	Report ID	0123456789
Report Date	1 JAN 2020	Location	OXFORD
Sample Type	SPUTUM	Sample Size	1.00g (100)
Reporting Lab	OXFORD	Report Date	1 JAN 2020

Summary

The specimen from **Bob Johnson** is positive for **Mycobacterium tuberculosis**. It is predicted to be resistant to **isoniazid** and **rifampin**. It belongs to a cluster of genetically related cases.

Diagnosis

The specimen is positive for **mycobacterium tuberculosis**.

Treatment

Based on predicted drug sensitivities, the isolate has **multi-drug resistant (MDR) TB**.

First Line Drugs:

Isoniazid	Resistant (Gene: katG)
Rifampin	Resistant (Gene: rpoB)
Ethambutol	Sensitive
Pyrazinamide	Sensitive

Second Line Drugs:

Streptomycin	Sensitive
Cycloserine	Sensitive
Clarithromycin	Sensitive
Moxifloxacin	Sensitive
Amikacin	Sensitive
Kanamycin	Sensitive
Capreomycin	Sensitive

Epidemiology

This case belongs to a cluster of 6 genetically related cases, suggesting recent transmission.

Comments

This sample was sequenced using the initial sequencing run did not provide high quality data for further analysis.

Authorised by: Dr. John Smith, Laboratory Director, Signature Date: 1 JAN 2020

Page 1 of 2

MYCOBACTERIAL GENOME SEQUENCING REPORT

Report Issued by: OXFORD Report Date: 1 JAN 2020

1 PATIENT INFORMATION

Name: Bob Johnson Identifier: 123456789
Birth Date: 1 Jan 1950 Sample Date: 1 Jan 2020
Location: Birmingham Gender: M

2 SPECIES IDENTIFIED BY SEQUENCING

100% identical to **Mycobacterium tuberculosis**

3 PREDICTED ANTIBIOTIC RESISTANCE

Resistant to **isoniazid**, **rifampin**.

4 EPIDEMIOLOGICAL RELATIONSHIPS

Belongs to a cluster of 6 genetically related cases, suggesting recent transmission.

5 SEQUENCING QUALITY

Sequenced 2 Aug 2018 on an Illumina MiSeq, yielding 4.71M reads, 4.70M (99.43%) mapped to the H37Rv (NC_009622.2) reference genome.

6 COMMENTS

This sample was sequenced using the initial sequencing run did not provide high quality data for analysis.

DIAGNOSIS DETAILS

Species	% Identity
Mycobacterium tuberculosis	100%
Mycobacterium avium complex	87%
Mycobacterium goodii	83%

TREATMENT DETAILS

Drug	Gene	Mutation	Coverage	Support
Isoniazid	katG	S315T	41x	4047/16231
Rifampin	rpoB	S531L	30x	3038/16331

EPIDEMIOLOGY DETAILS

Isolate	Year	SNP Distance
01234_0	2019	0
01234_1	2019	4
01234_2	2019	8
01234_3	2019	7
01234_4	2019	10
01234_5	2019	9
01234_6	2019	16
01234_7	2019	9

GENOME SEQUENCING DETAILS

LOCUS LINK ID	123456789	STATUS	STRANDED WGS
RNA DATE	1 JAN 2020	PLATE/STRIP	ELUMINATED
HTG METHOD	4.7M	INTERPRETED BY	4.7M (99.43%)
REFERENCE GENOME	NC_009622.2		

Page 1 of 2

MYCOBACTERIAL GENOME SEQUENCING REPORT

Report Issued by: OXFORD Report Date: 1 JAN 2020

7 Technical Details

This section of the report provides the technical details for the specimens presented on the first page.

Resistotype

The resistotype describes the mutations that are predicted to confer drug resistance.

Drug	Gene	Mutation	Locality	Coverage	Support
Isoniazid	katG	S315T	Middle of	41x	4047/16231
Rifampin	rpoB	S531L	Middle of	30x	3038/16331

Related Isolates

The following graph and table describe isolates that have been identified as being genetically similar to this patient's isolate.

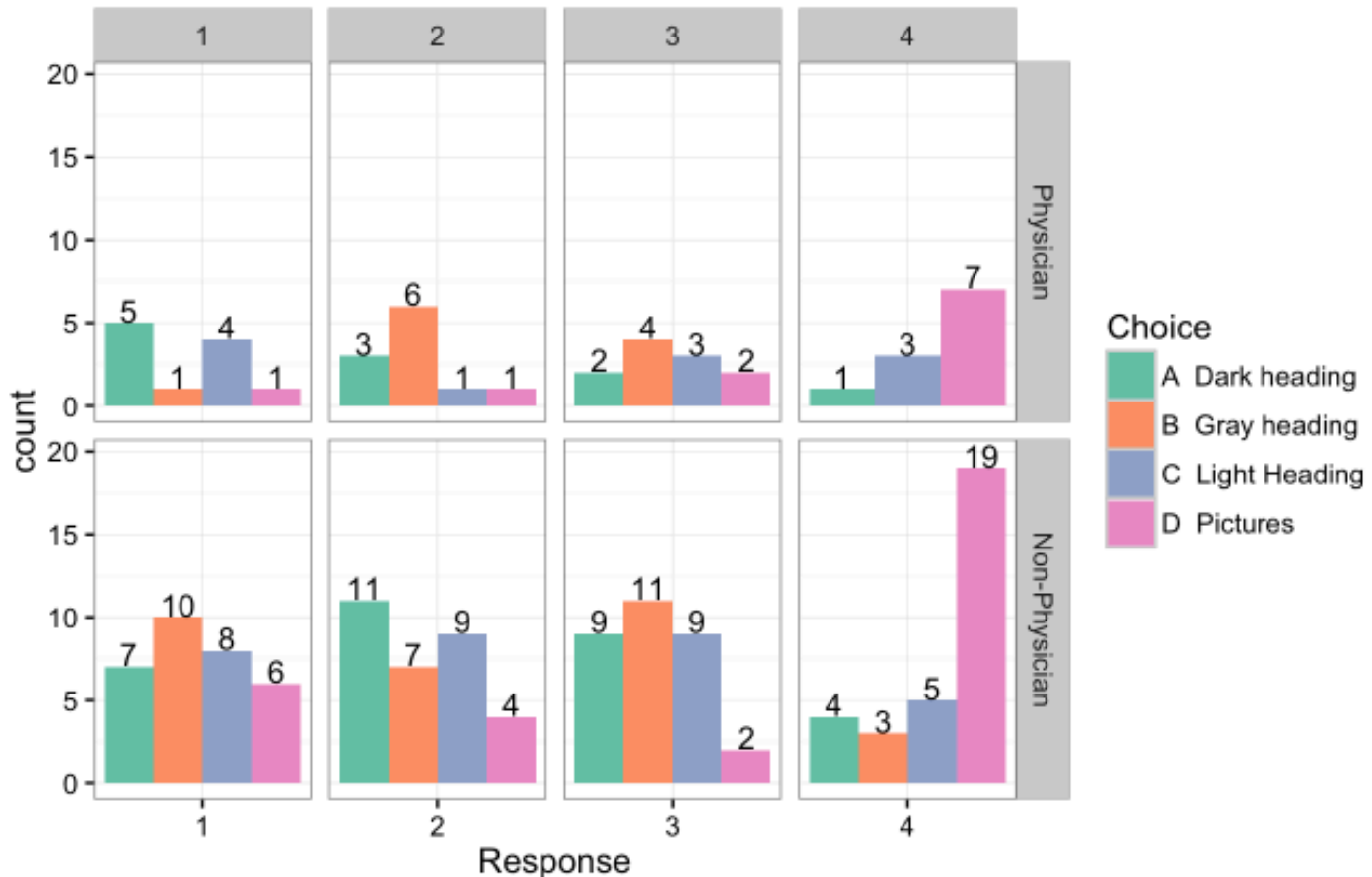
Isolate	Year	SNP Distance
01234_0	2019	0
01234_1	2019	4
01234_2	2019	8
01234_3	2019	7
01234_4	2019	10
01234_5	2019	9
01234_6	2019	16
01234_7	2019	9

PHASE 3

DESIGN CHOICE QUESTIONNAIRE

Key finding #1: Comparing whole reports not very useful

The previous 4 report prototypes demonstrate different ways of presenting lab data from whole genome sequencing of a tuberculosis isolate Which of the reports to you prefer

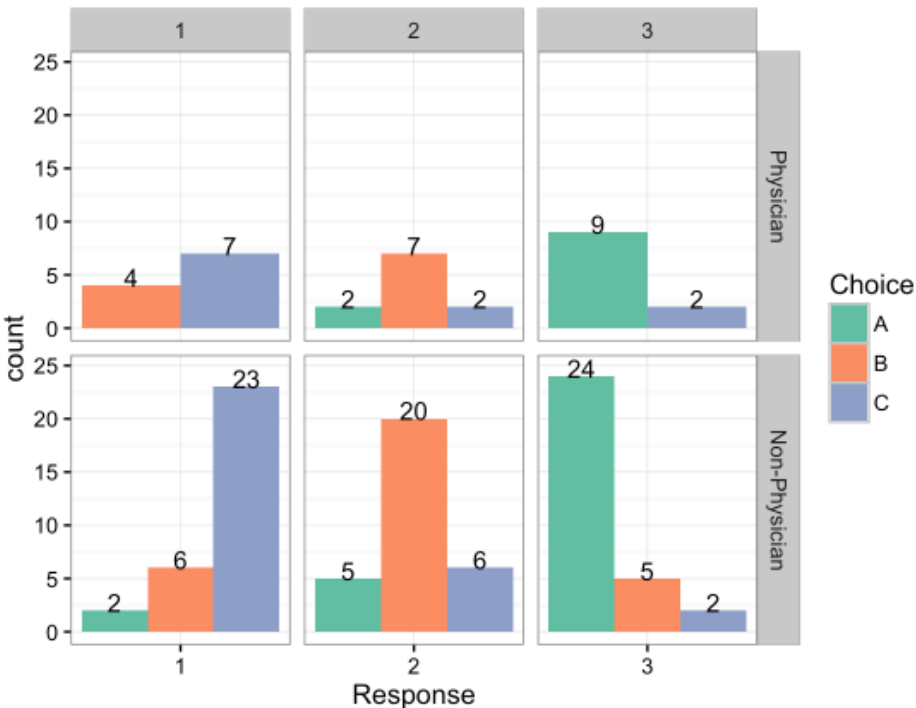


PHASE 3

DESIGN CHOICE QUESTIONNAIRE

Key finding #2: Generally strong preference patterns, consistent between clinicians and non-clinicians

Depending on the resistance mutations observed an isolate might be identified as having multidrug resistant TB MDR TB There are many ways this could be noted on the report



A

Drug Susceptibility

Drug	Prediction
Isoniazid	Resistant
Rifampin	Resistant
Ethambutol	Sensitive
Pyrazinimide	Sensitive

B

Drug Susceptibility

Based on predicted antibiotic sensitivities, this individual has multidrug-resistant (MDR) TB.

Drug	Prediction
Isoniazid	Resistant
Rifampin	Resistant
Ethambutol	Sensitive
Pyrazinimide	Sensitive

C

Drug Susceptibility

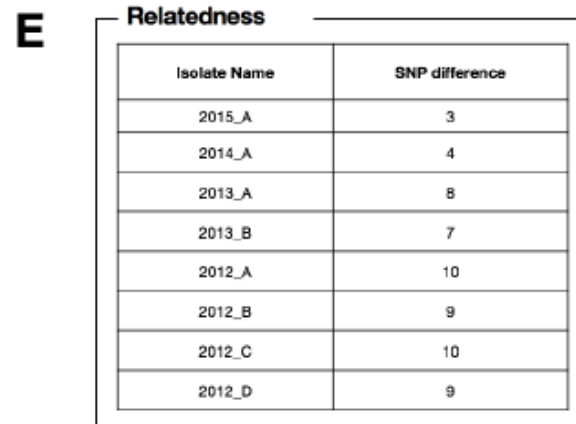
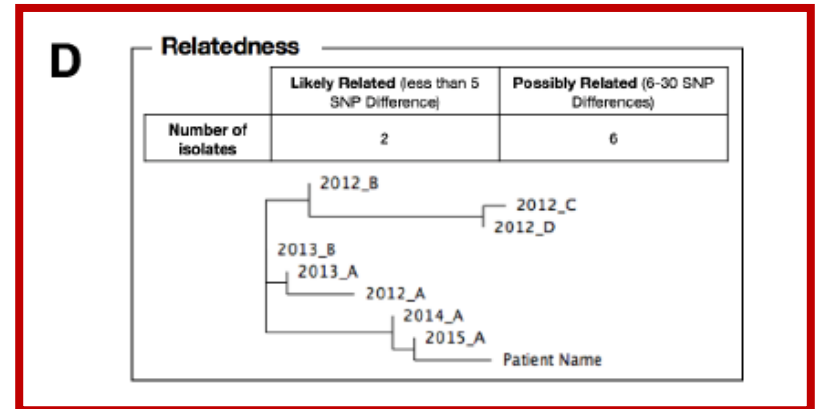
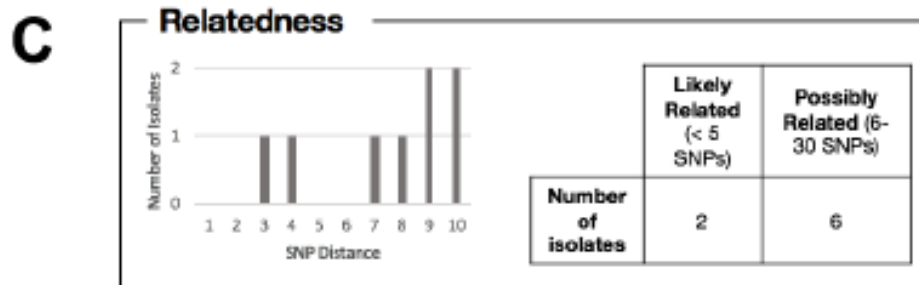
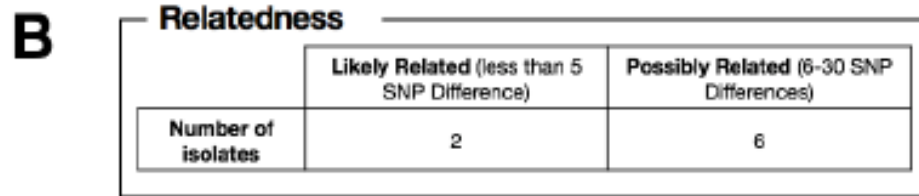
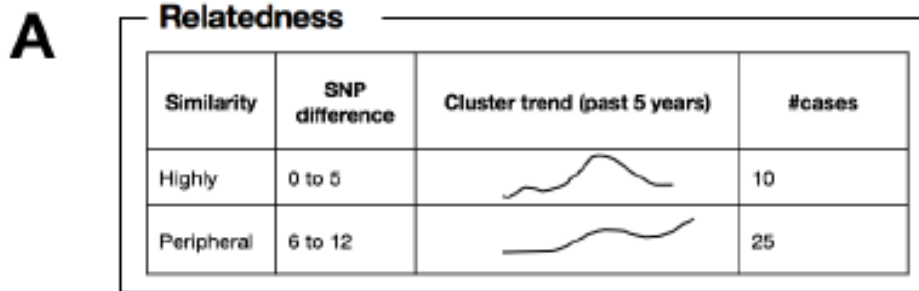
Mono-resistant
 Multidrug-resistant (MDR)
 Extremely Drug Resistant (XDR)

Drug	Prediction
Isoniazid	Resistant
Rifampin	Resistant
Ethambutol	Sensitive
Pyrazinimide	Sensitive

PHASE 3

DESIGN CHOICE QUESTIONNAIRE

Key finding #2: Generally strong preference patterns, consistent between clinicians and non-clinicians

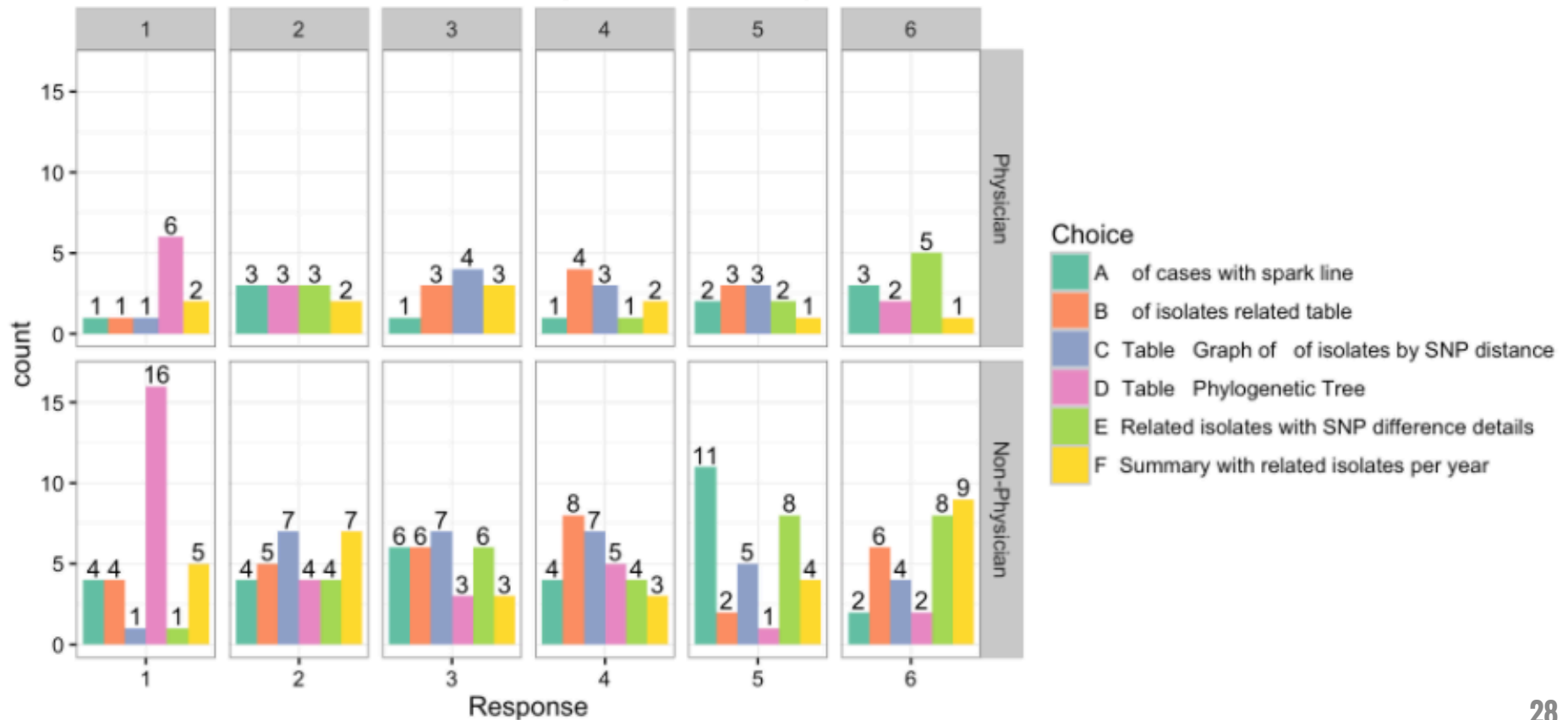


PHASE 3

DESIGN CHOICE QUESTIONNAIRE

Key finding #2: Generally strong preference patterns, consistent between clinicians and non-clinicians

Data on relatedness to other isolates clusters is presented below in a number of different formats Which do you find most interpretable

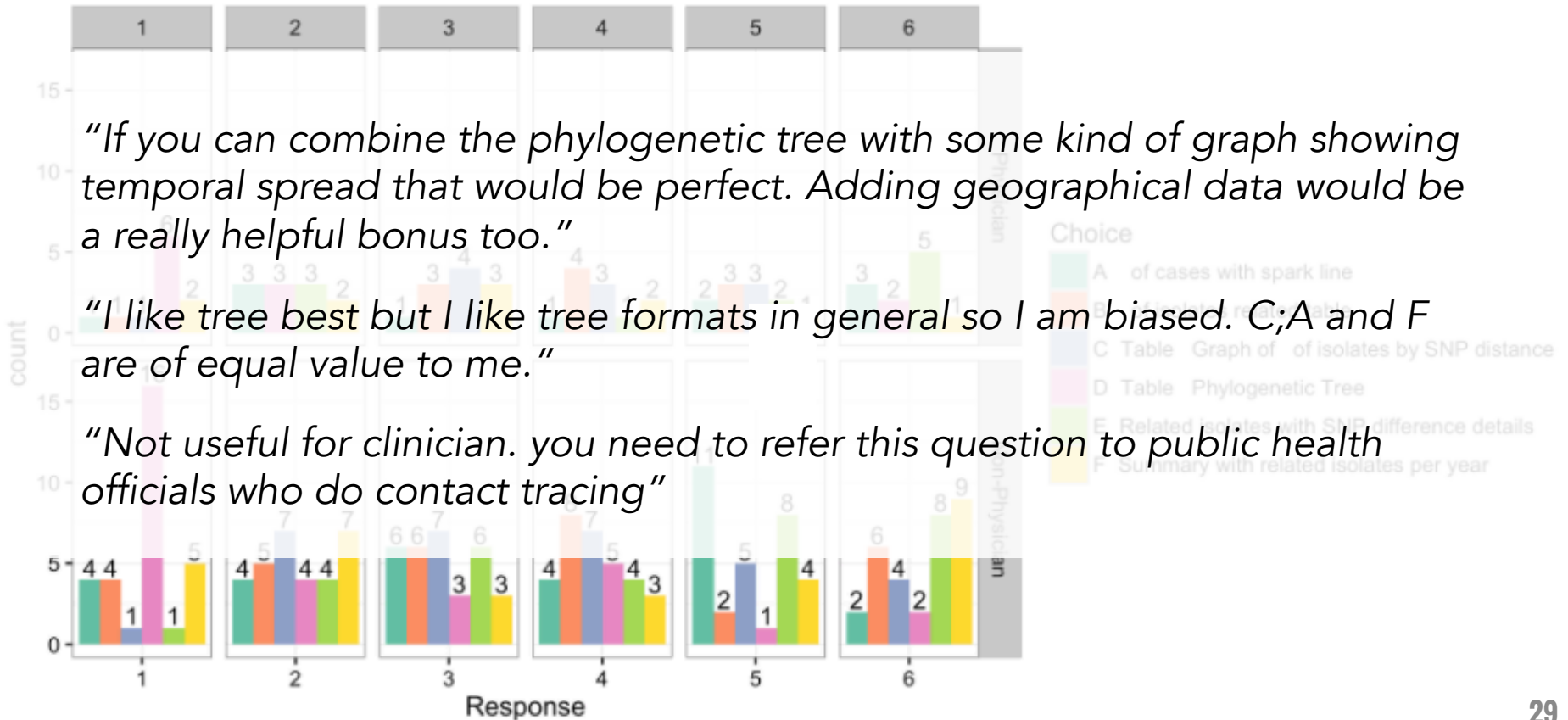


PHASE 3

DESIGN CHOICE QUESTIONNAIRE

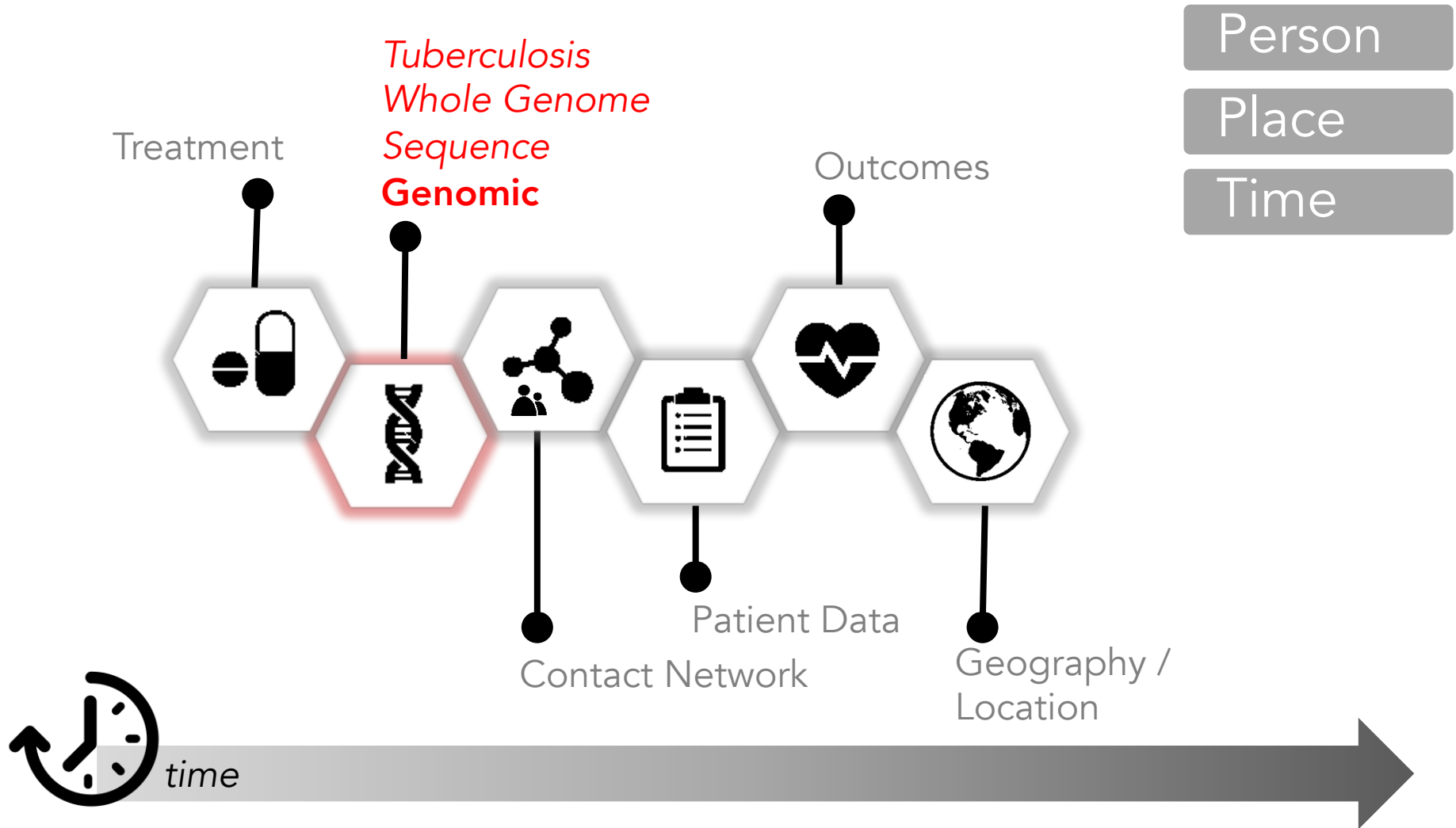
Key finding #2: Generally strong preference patterns, consistent between clinicians and non-clinicians

Data on relatedness to other isolates clusters is presented below in a number of different formats Which do you find most interpretable



**Problem & task data will be used to
construct more complex
visualizations in future**like my PhD work**

WHERE IS MY WORK HEADED?



EpiCOGS

<https://amcrisan.shinyapps.io/EpiCOGSDEMO/>

Load and View Data

Load Data

All Patients 118

Selected Patients 3

Feedback

Filter Patients

Location

Filter By: City

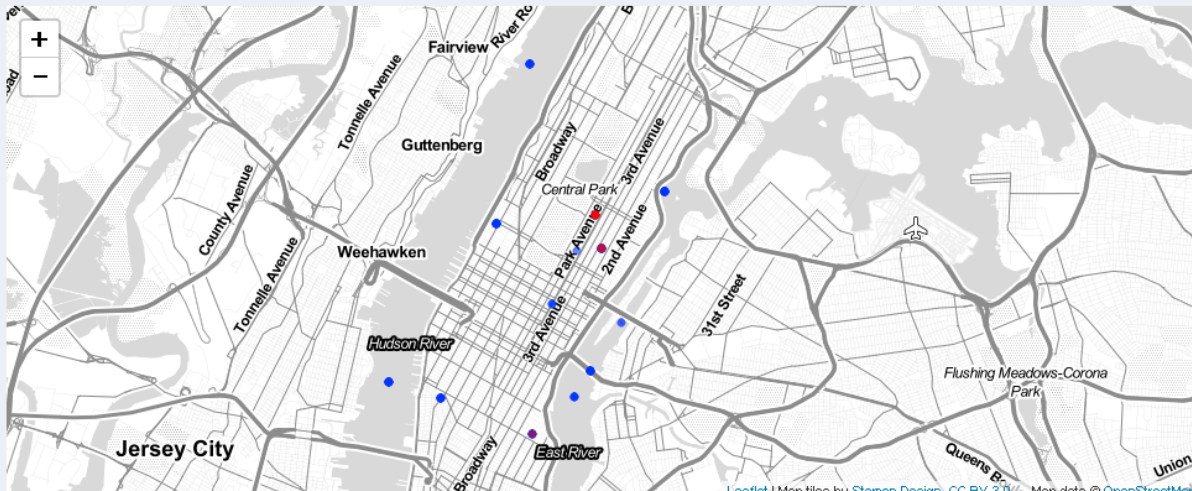
Filter By: State/Province

Demographics +

Diagnosis +

Treatment +

Outcome +



Show 10 entries

Search:

	City	State	Lat	Long	caseID	originOfBirth	gender	MethodOfDetection	tbType	diagnosisDate	t
1	New York	NY	40.74446	-73.96239	0002	1	M	Conact investigation	Active	2014-09-17	20
2	New York	NY	40.73901	-73.96682	0004	1	F	Pre-landing Surveillance	Active	2014-02-13	20
3	New York	NY	40.77706	-73.96111	0007	1	F	Screening Program	Active	2014-04-16	20
4	New York	NY	40.77012	-73.95928	0008	0	F	Unknown	Active	2014-01-28	20
5	New York	NY	40.73109	-73.97850	0009	0	F	Unknown	Active	2015-01-10	20
6	New York	NY	40.77533	-73.98853	0019	1	F	Screening Program	Active	2014-06-11	20
7	New York	NY	40.75850	-73.97310	0023	1	F	Other	Active	2014-12-12	20

DECOMPOSING VIS TO TWO LEVELS

PROBLEM & TASK BASED DESIGN

Working with stakeholders to solve relevant problems & provide workable solutions



ABSTRACTIONS & VISUAL ENCODINGS

Common terminology to describe & compare visualizations



IN CONCLUSION

- Data visualization can support decision making in diverse stakeholder groups
- Visual design, not just presence of visualization, matters
- Visualization is a research process in design
- Consider and evaluate alternative choices
- Stay tuned for future developments!

Contact Info



@amcrisan



<http://cs.ubc.ca/~acrisan>



acrisan@cs.ubc.ca

Thanks

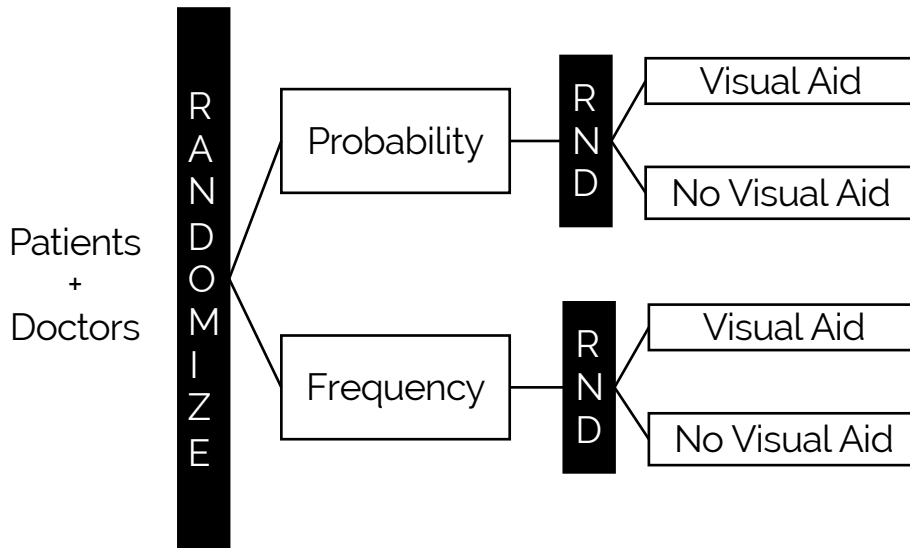
Dr. James Johnston, Dr. Maureen Mayhew, Dr. Victoria Cook, Nash Dahlla, Dr. Jason Wong, Dr. James Brooks, Johnathan Spence, Laura MacDougall, Michael Coss, Ciaran Aiken, and David Roth, Matthew Brehmer, Madison Elliott, Zipeng Liu, Dylan Dong, and Kimberly Dextras-Romagnino

Additional Slides

EXAMPLE : SHARED DECISION MAKING

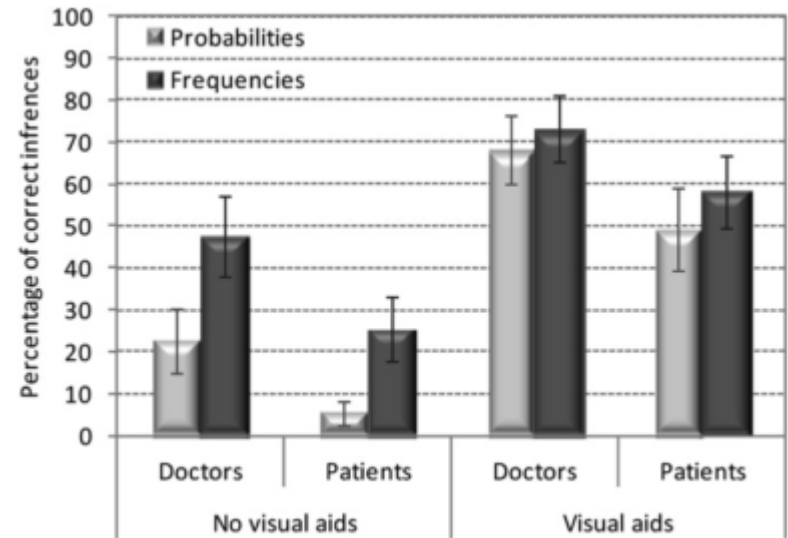
STUDY DESIGN

Quasi-randomized trial with four conditions
Outcome : correctly calculating the risk (essentially a math test)



RESULTS

Visualization improved comprehension of both doctors and patients
Visualization improved concordance between doctors and patients



DECOMPOSING VIS TO TWO LEVELS

PROBLEM & TASK BASED DESIGN

Working with stakeholders to solve relevant problems & provide workable solutions



ABSTRACTIONS & VISUAL ENCODINGS

Common terminology to describe & compare visualizations



PROBLEM & TASK BASED DESIGN

Why is data being visualized?

Different stakeholders have different needs!

Example

Problem:

Understanding Disease Dynamics

Tasks:

"How is a pathogen changing over time?" *"Are there clusters of disease?"*

?

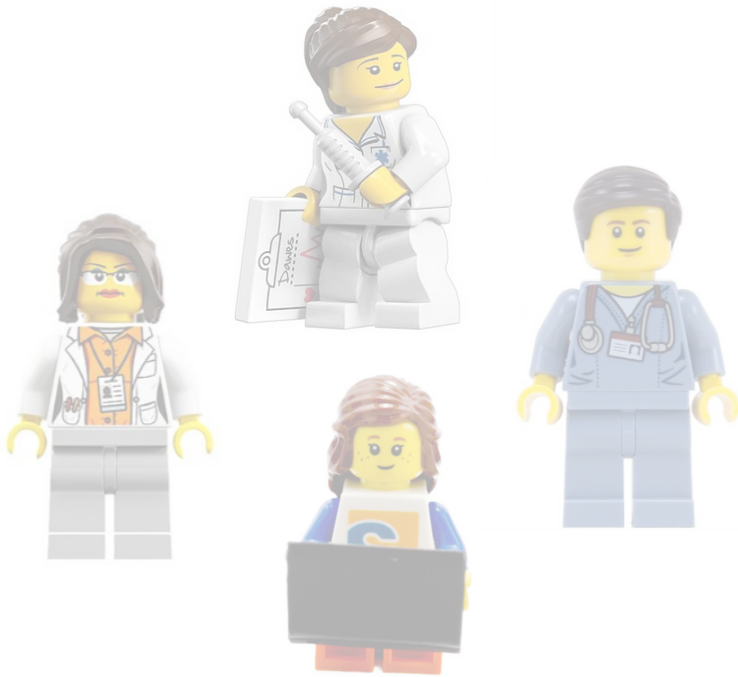
?



DECOMPOSING VIS TO TWO LEVELS

PROBLEM & TASK BASED DESIGN

Working with stakeholders to solve relevant problems & provide workable solutions



ABSTRACTIONS & VISUAL ENCODINGS

Common terminology to describe & compare visualizations

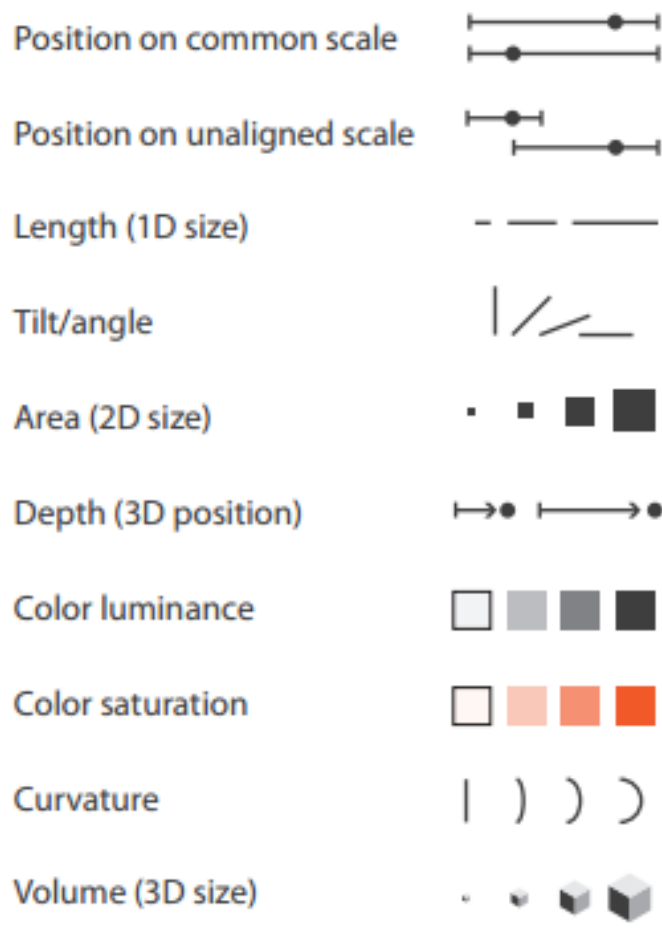


ABSTRACTIONS & VISUAL ENCODINGS

Decomposition Visualizations into geometric shapes & properties

Channels: Expressiveness Types and Effectiveness Ranks

➔ Magnitude Channels: Ordered Attributes



➔ Identity Channels: Categorical Attributes



DESCRIBING VISUALIZATIONS

Using geometric marks and their properties (channels)



Vertical Position



Vertical Position
Horizontal Position



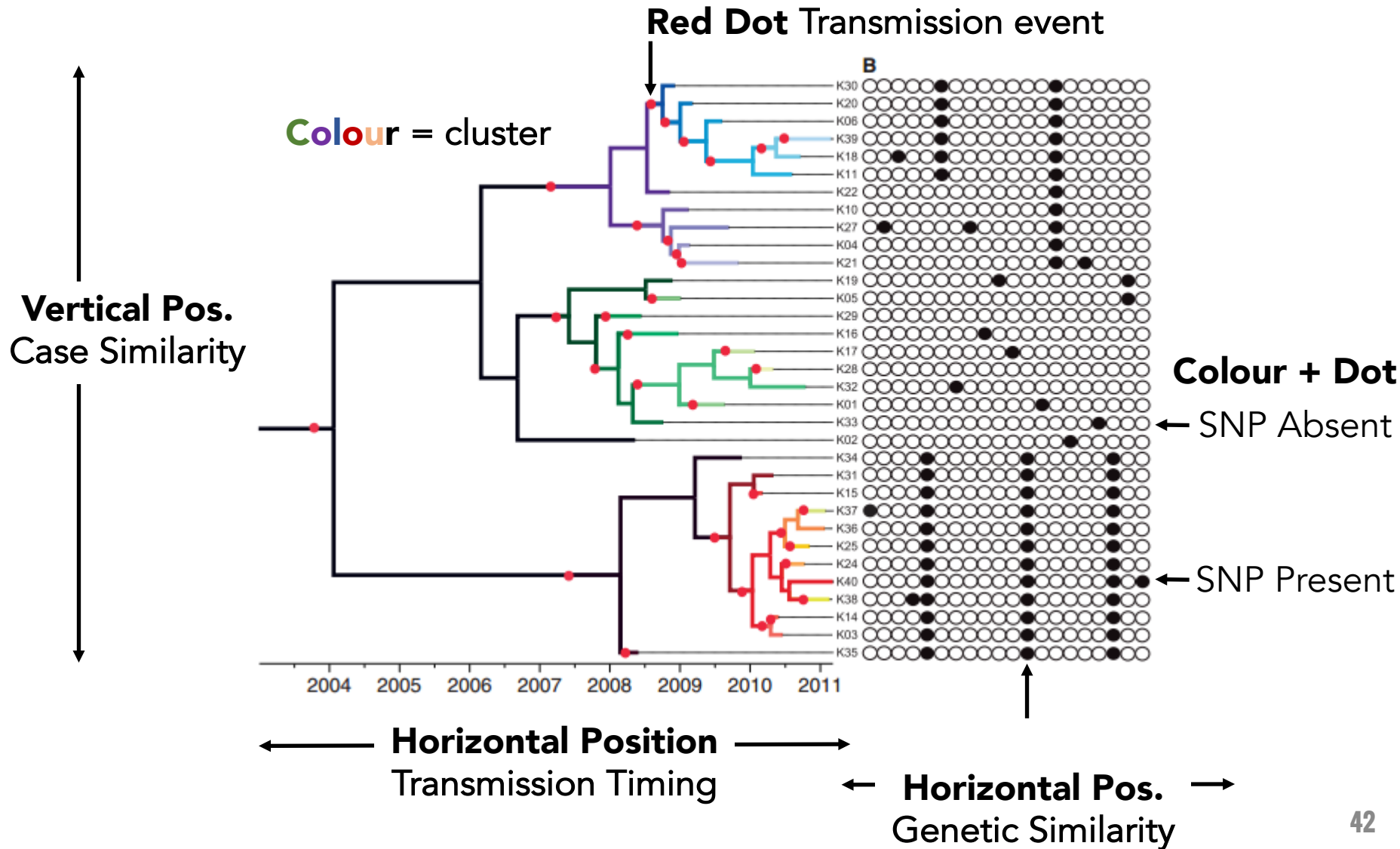
Vertical Position
Horizontal Position
Colour



Vertical Position
Horizontal Position
Colour
Size

DESCRIBING DISEASE DYNAMICS

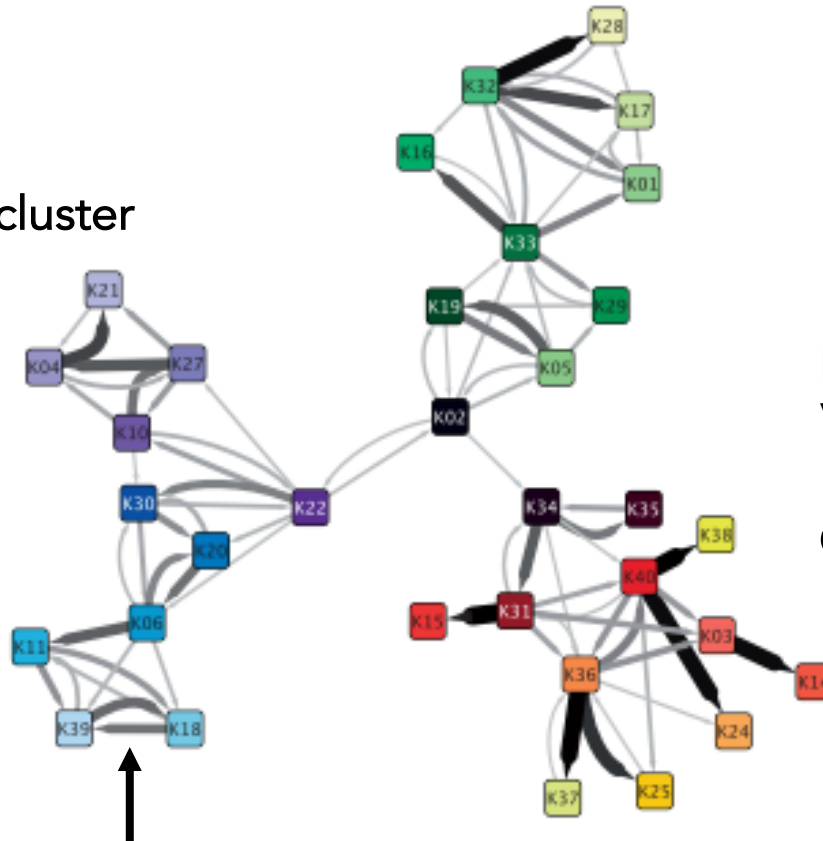
EXAMPLE 1: PHYLOGENETIC TREE + DOT PLOT



DESCRIBING DISEASE DYNAMICS

EXAMPLE 2: NETWORK DIAGRAM

Colour = cluster



**Horizontal +
Vertical Position**

Chain of transmissions
Case Similarity

Thickness of line

Probability of transmissions

DESCRIBING DISEASE DYNAMIC

COMPARING VISUALIZATIONS

What information does the visualization show?

How does the visualization show that information?

WHAT	HOW	TREE	NETWORK
Transmission	Horizontal pos.	■	
Timing	Colored Dot	■	
Transmission	Thickness		■
Confidence	Color		■
Case Similarity	Horizontal + Vertical pos.	■	■
	Black/White Dot	■	
	Colored Shape	■ (line)	■ (square)
SNP presence	Black/White Dot	■	

LINKING VIS ABSTRACTIONS TO BIOINFORMATIC ONTOLOGIES

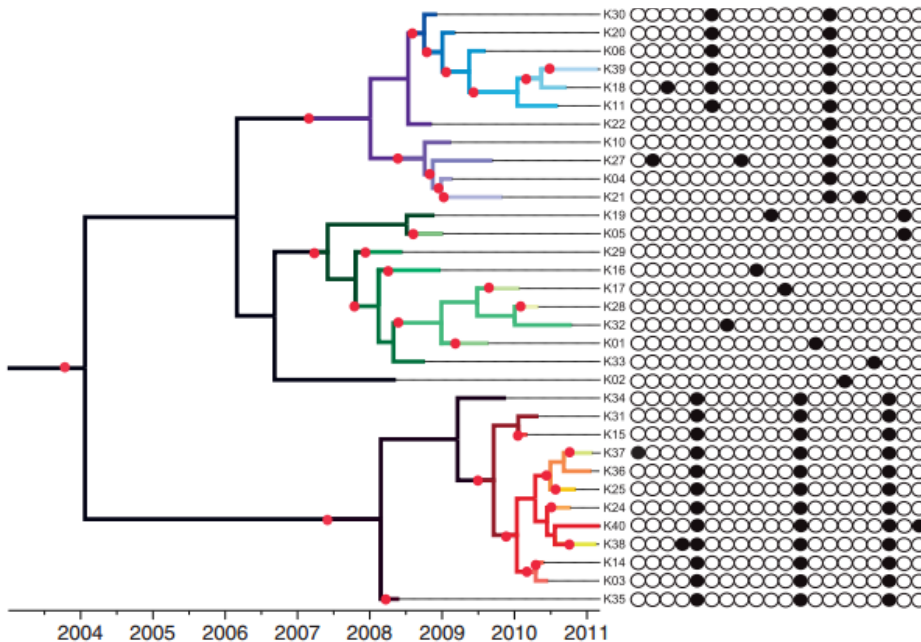
- Can connect ontologies to visualizations through abstractions
- *Suggest* visualizations based on available data
- Need to know *what kinds* of visualizations are suitable for different research questions
 - Currently working on this 🙌
 - Similar idea to www.vizhealth.org

PROBLEM & TASK BASED DESIGN

Why is data being visualized?

"How is a pathogen changing over time?"

Transmission timing & genetic similarity



"Are there clusters of disease?"

Transmission between & within clusters of related cases

