## 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)





# 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



## PROPOSAL

# Visualization of public health data can improve knowledge sharing and decision making in

infectious disease prevention and control

# WHY VISUALIZATION?

Least Understa	ndable			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



Zikmund-Fisher (2013). A demonstration of "less can be more" in risk graphics.

### **EXAMPLE OF GUIDANCE : WWW. VIZHEALTH.ORG**

AVERAGE LIFE EXTENSION

Total average:

Risk over time

Casa counts



SEVERE SKIN DISORDER

YORK MOK

## **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

### Steps for visual design

Partner with a group of stakeholders that have a problem

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

- 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]

- 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**

- 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

- 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

### Steps for visual design AN ITERAVTIVE PROCESS

- 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

Samp	le Det	ails
------	--------	------

Sequencing	Oxford	Date received in	
Location		Lab	
Local Lims Specimen ID	12.0610882	Run date	01/01/20151008
Guuid	b7aa98e0-3612-4c0	o-a47b-471e0e78c72	d

Organism Identification	
Predicted/closest match	
TBCOMP/microti.	100%
TBCOMP	100%
TBCOMP/TB	96.77%
TBCOMP/tuberculosis-canettii	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.



++

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?

Response	Chart	Percentage	Count
There aren't any barriers		0.0%	0
Additional laboratory data is needed		33.3%	2
Timeliness of results being provided (too slow)		83.3%	5
Results provided over multiple unconnected documents		83.3%	5
Difficulty interpreting lab results		50.0%	3
Lab data is not routinely provided		16.7%	1
Lab data is not routinely linked to patient data		50.0%	3
Other, please specify		16.7%	1
		Total Responses	6

### **PHASE 2** DESIGN SPRINT



## **PHASE 2** DESIGN SPRINT



### Participants: 42 Goal: Compare control (existing report) with options developed in the design sprint

<text><text><text><text><text><text><text><text><text><text></text></text></text></text></text></text></text></text></text></text>	manual manua manual manual manual manual manual manual	Patient Infe	ormation				Epidemia	logic Sumr	nary		
wind wind wind wind wind wind wind wind	marking       marking       marking         name       marking       marking       marking         marking       marking       marking       marking       marking         marking       markin	Patient Hanse	Bob Johnson		Sample Type	Oputum.	Aletheolology-7	alients are automa	feally assignments also	ters based upon based	ger sigle suckelde
margin       margin       margin       margin         margin       margin       margin       margin         margin       margin       margin       margin         margin       margin       margin       margin         margin       margin       margin       margin       margin         margin       margin       margin       margin       margin       margin         margin <td>main minimum       main minimum         main minimum       main minimum         statement       main minimum         main minimum       main minimum         statement       main minimum         main minimum       main minimum</td> <td>Patient ID</td> <td>123456769</td> <td></td> <td>Sample Site</td> <td></td> <td>bed and an or a</td> <td>ananences casses</td> <td>right and a second second</td> <td>no second to ask</td> <td>erenerative paper.</td>	main minimum       main minimum         main minimum       main minimum         statement       main minimum         main minimum       main minimum         statement       main minimum         main minimum       main minimum	Patient ID	123456769		Sample Site		bed and an or a	ananences casses	right and a second second	no second to ask	erenerative paper.
water     water     water       Burding and public fills water     water water     water       Burding and public fills water     water water     water       Burding and public fills water     water     water       Burding and public fills water     water     water       Burding and public fills water     water     water       Burding and public fills water     water     water       Burding and public fills water     water     water       Burding and public fills water     water     water       Burding and public fills water     water     water       Burding and public fills water     water     water       Burding and public fills water     water     water       Burding and public fills water     water     water       Burding and public fills water     water     water       Burding and public fills water     water     water       Burding and public fills water     water     water       Burding and public fills water     water     water       Burding and public fills water     water     water       Burding and public fills water     water     water       Burding and public fills water     water     water       Burding and public fills water     water     water       Burd	<ul> <li>where we have the same way and w</li></ul>	Patient Dull	01-01-1800		Sample Date	01-01-1908	The specim	in beiongs to a j	previously existin	g cluster	
Summary of Principal       Beington analysis of the sources in this, high pair is the spectration in the spectration is the spectration in the spectration is the spectration in the spectration is the s	Summary of Principal     Image: Summary of Summary	Location	Owland		Specimen ID	120404789	-	SNP			Mex
Summary of relatings         Building and purphysical frequencies of the generative frequencies of	Summary of rotating       Between the state of some when the spectra state state state state state state of some when the spectra state state state state of some when the spectra state sta	P	of Findings				Siniatty	diference	Cluster	Yeed gasit 5 year	ni) (iki
Base base is the standards of sources in the specific horizon.       Base base is the standards of sources in the specific horizon.       The specific horizon and specific horizon.       The specific horizon. <td>Base base is the transformed is constraints. Reference, there, the constraints are constraints.       Degrad Base constraints are constraints are constraints.       The main constraints are constraints.       The main constraints are constraints.       The main constraints.</td> <td>Based upon a</td> <td>or renoring e</td> <td>a sector</td> <td>n's companie data. Prince</td> <td>aliand has manufastianiam</td> <td>Highly</td> <td>0 to 5</td> <td></td> <td></td> <td>2</td>	Base base is the transformed is constraints. Reference, there, the constraints are constraints.       Degrad Base constraints are constraints are constraints.       The main constraints are constraints.       The main constraints are constraints.       The main constraints.	Based upon a	or renoring e	a sector	n's companie data. Prince	aliand has manufastianiam	Highly	0 to 5			2
Constraining and matches of account with sinking operation flocking.       Charging and matches of account with sinking operation flocking.       Charging and matches of account with sinking operation flocking.       Charging and matches of account with account with a schedule with	mode being as a statute of loads at the statute protect fielding.           mode being as a statute of loads at the statute statute statute at the statute at the statute statute at the statute statute at the statute statute at the statute statutestatutestatute statute statute statute statute statutestatutestatu	Subseculosis	that is predicte	d to be set	istant to 2 antibiotics #e	serviceld, Rifempiro), This		-		-	
Biggrool B         Monitor Links and the hole down one as any queet of traduction and interphotodocol and int	Biggins II     Consistence of the state	case belongs	to a cluster of	cases with	similar genemic finding	6.	Perpheral	6 to 12	/	$\sim$	6
$\label{eq:product} Product P$	Markan Strandskinger Standskinger Stand	Diagnosis									
The sportner was consider to be net drug resident (RDR) TB. Example and the folders: <u>The sportner</u> for a constrained to the net of the sportner (RDR) (RDR) <u>The sportner</u> for a constrained to the net of the sportner (RDR) (RDR) <u>The sportner</u> for a constrained to the net of the sportner (RDR) (RDR) <u>The sportner</u> for a constrained to the net of the sportner (RDR) (RDR) (RDR) <u>The sportner</u> for a constrained to the net of the sportner (RDR)	The sponse was consider to be net-decing ensidered (ADM) TR. Some and a procession of the net-decing ensidered (ADM) TR. The the the the the net-decing ensidered (ADM) TR. The the the the the net-decing ensidered (ADM) TR. The the the the the the the the the the t	The specimen	was specialed	as myceb	acterium tuberculosis		The whole-game greater than 6.7	me sepuence analy million with 85.475	tois of the isolate was in mapping and a server	omidenet <u>ISGNOLAR</u> age of BLOPS .	LEY is the number of r
Durge Netwich         Neutrol 1         Converter Service         Neutrol 1         Converter Service           Status         1         Gase: pQL Area Add Charge: EDST.           Druch Missing         2         -           Status         7         -           Status         7         -	Dage Network     Periodic Torum     Contraction Torum     Contraction Torum       Statistical Statistical Statistical Contraction Contrection Contraction Contrection Contraction Contraction Co	Treatment Menablogy Dr	ug sensitivites ere	n podited.	using the genomic sequence of	data il accordance to the method	ī				
Normaline Terrory in National Deveload 1         Gene Lab. Koret Aud Change 1917 Expension Service in Service J         Normaline Service In Service J         Normaline In	Normaline         Normaline <t< th=""><th>Treatment Menacoly: Dr reported to public The specimen Summary of s</th><th>ug sensitivities and shell paper cell was consider 1 ensitive finding</th><th>in produted. to be <b>mail</b>ti</th><th>ung the percent accuracy drag resistant (NDR) 1</th><th>ate e acosteror o te nethad</th><th>- Comment</th><th>15</th><th></th><th></th><th></th></t<>	Treatment Menacoly: Dr reported to public The specimen Summary of s	ug sensitivities and shell paper cell was consider 1 ensitive finding	in produted. to be <b>mail</b> ti	ung the percent accuracy drag resistant (NDR) 1	ate e acosteror o te nethad	- Comment	15			
Impute         Name         Name         Search (obs. And Add tage) DATA.           Detech (all of the Add tage) DATA.         Impute         Impute <td>Impute     Name     I     Seet (ob. Anne. And ComplexitY)       Deviation Some     Z     -       Deviation Some     Z     -   Partners of 2  Partners of 2</td> <td>Treatment Mehasogy Dr reported in public The specimen Summary of a Drugs</td> <td>ug sensitivities een viterigegen ool was consider t ensitive finding Prediction</td> <td>e pedited. to be <b>mail</b>tic</td> <td>any the periodic sequence of the generation of t</td> <td>alls in accordance to the method TB.</td> <td>Comment</td> <td>la</td> <td></td> <td></td> <td></td>	Impute     Name     I     Seet (ob. Anne. And ComplexitY)       Deviation Some     Z     -       Deviation Some     Z     -   Partners of 2  Partners of 2	Treatment Mehasogy Dr reported in public The specimen Summary of a Drugs	ug sensitivities een viterigegen ool was consider t ensitive finding Prediction	e pedited. to be <b>mail</b> tic	any the periodic sequence of the generation of t	alls in accordance to the method TB.	Comment	la			
Contraction         Samuel         Z         -           Chr         Samuel         Z         -         -           Chr         Samuel         Z         -         -           All         Samuel         Z         -         -           All         Samuel         Z         -         -           All         Samuel         Z         -         -             Page 1073         Page 1073         Page 1073         Page 1073	Literative Paralismic Sensitive         J         -           Char         South         Z         -           Char         South         Z         -           A         South         Z         -           A         South         Z         -           A         South         Z         -           A         South         Z         -	Treatment Menescoy Dr reported in public The specimen Summary of s Drugs	La sensitivites est the agence: was consider t ensitive finding Prediction Prediction	e protined. to be mailing Status	drug the genomic sequence of drug resistant (MDR) 1 Comment Generical, Anten Acid	ada e accontence to the method TB.	Comment	15			
Projection         Section         Projection         Section         Projection         Projection <td>Type Content     Second of the content o</td> <td>Treatment Menacoy: Dreading The spectree Summary of s Drags Isoniatid Riterpin</td> <td>ug sensitivities aus shed paper and was consider th ansitive findings Prediction Resident Resident</td> <td>o produted. to be mailed Status 1</td> <td>drug the generatic sequence of the generation of the generation of the generation of the second seco</td> <td>data in acconducto to the method TB. Changer 52/87 Changer 5575</td> <td>Comment</td> <td>15</td> <td></td> <td></td> <td></td>	Type Content     Second of the content o	Treatment Menacoy: Dreading The spectree Summary of s Drags Isoniatid Riterpin	ug sensitivities aus shed paper and was consider th ansitive findings Prediction Resident Resident	o produted. to be mailed Status 1	drug the generatic sequence of the generation of the generation of the generation of the second seco	data in acconducto to the method TB. Changer 52/87 Changer 5575	Comment	15			
And Section     Frequencies       All Section     Frequencies	Land     January     January       30     Sentul     Z       31     Sentul     Z         Page 1 of 2	Treatment Menaboy: Dr appelled in public Summary of s Drugs Drugs Bernland Rhampin Etherstuhl	sg sensibilites an anter paper est was consider t anstitue finding Precision Precision Resistant Sensitive Sensitive	er produted.	any the percent sequence - drug resistant (MDR) 1 Conversers Gene tails, Anire Acid Lense rpob, Anire Acid	dala in Accordance to the method TB. Changer 50/87 Changer 50/87	Comment Mission 2 Marca 2 Marca	is —			
Image:	Image: Property of 2         Automating         <	Treatment Menacogy Dr appelled in public Summary of s Drugs Drugs Bernland Rhampin Ethersteld Pyrachemide men	ug sensibilites an electroperatorial was consider t ensible finding Precision Resistent Resistent Sensitive Sensitive	e peditor. b be multi- i status 1 1 2 2	any the promit requests drug resistant (MDR) Commerce Gener Kall, Anire Add Commerce Gener pobl. Anire Add Commerce Commerce	dala in Accordance to the method TB. Change: 52/5T (Change: 52/5T	Comment Memory 2 Mary 2 Mary	ta ———			
ren i neven ( n i neven )	Page 1 of 2	Treatment Menuacy: Dreyenter Summary of s Dregs Bernladd Ritemin Enerstadd Pyschersch Opt Se	vas consider to bulgapin at vas consider to ensitive finding Precisant Resistant Sensitive Sensitive Sensitive	e poditko. b be multi status 1 2 2 2	drag the percent concerns of drag resistant (NDR) 1 Comment Semi kalls, Anim And Gene (nob. Anim And - -	das a actoritarios to the restluce TB. Changes 10/017 (Changes 10/75.	Comment Manager 1 Mary 1 Mary	ta ———			
Page 1 of 2	Page 1 of 2	Treatment Menualogy Dr wywitel is public The specimen Summary of s Drugs Remitted Rhampin Ethershell Prachersch Oct Ski	uga sonothities and hump appear and wass consider 1 ensitive finding Precision Resistant Bensitive Sensitive Sensitive Sensitive Sensitive	e podited.	drug the percent requestor i drug resistant (MDR) 1 Convert Generation, Antine Acid Gener, 1908, Antine Acid - - - -	Ada in accordance to the method TB. Georgen 50/07 (charges 50/75.	Comment Network 1 Mary 2 Administration	ta y Dr. Jako Dra	8	Spates	<b>A</b>
	45.4 M	Treatment Mehazing: Dresperind In public The spectrum summary of s Drugs Bernland Riterapin Etheritude Pyrachemide Skil Skil Skil	cg sensitivities are short-paper on: was consider t ensitive finding Prediction Prediction Prediction Sensitive Sensitive Sensitive Sensitive Sensitive	er predicted.	atig the percent expander drag resistant (MDR) Comment Bane kalls, Anten Ante Anten kalls, Anten Ante - - - - -	das e acusteros to the methad TB. Changer 15/07 Changer 15/07	Comment Messare 2 May 1 2 May 1 Pastan	Y Dr. Jahr Drai Laboratory D	Di Vendiar	Ngatas Ma	D-0-Set

Mycobac	terial Ge	nome \$	Sequ	encing Re	SUITS England	DIACON		-		
INVERTIGATION	108.004800			INTENT D	THEFT	Children	uala u	ELAILO	1. 1991	~
BRD-GAT	LAN 1992	OTHER !	M	LOCATION	Aycobecterium subenculaeia Aucobecterium subenculaeia			1005	¢.	
	1.018 1999	1,0110,011		LOUGHERT				475	47%	
EMPLE TYPE	89,75,87			PPACE PGL	1.0011008	Manufacture of control to the			375	
REPORTING LAB	047080			REPORT DATE	1.0441008	in province				
CUMBARY						TREAT	MENT (	DETAILS		
aumani						Drup	Gene	Mutation	Cutalog	Coveneo
The spectrum h	om Beb Johans circlet and cities	en is pesitive in polis. It balance	or Myce	bacterium tobereul bastes of opposically	color: It is predicted to be	Inomiacid	Aut5	53/IST	Mykrobe v2	47x
				and of the control	rowing here.	Ritampin	rpoli	55311.	Walker et al.	38×
DIAGNOBIE	×۸					EPIDEN	NOLOG	Y DETA	LS	
The specimen is	positive for My	odectorium	Merce	icua la		1000		No.	24	Datamore
						2115,A		19-6	- 0	
TREATMEN	п Л —			EPIDENIOL	A year	2010,6		2014		
Eased on predi-	ted artibiotic as	rait/fes. th		This isolate belong	p to a cluster of @genetically	210.A 210				
individual has m	ruitidrug-realisti	ent (MER) TS.		Nialed cases, sug	gesting recent hansnisator.	2010.8		2010	2	
First-Line Drug						272.4		10.0	1.0	
isoniadid.	Provisiont (	Nati 50117)				1000		10.15		
Reample	Plesistant (	111685 Bogs			_	0110.8				
Dhambutol	Sensitive				3	2410/C		20.0		
Pytazinimide	Sensitive					2810,D		2010		
Second-Line D	1001			10× 100	20.204.201					
Sneptomycin	Sensitive									
Operforment	Sensitive			COMMENTS						
Offenation	Sensitive			The second parts	encounter taken the initial					
Montheastin	Sensitive			CONTRACTOR NAME	september and setting					

Sensitive Sensitive

Sensitive Sensitive

ALTHORNEO BY DR. JOIN SHITT

COTON LANOWOOT OF

Kanamych

LOOIL LIME ID	12.0010862	BUUD .	\$7.44860-0012-408-	
RUR DATE	1.0011000	RATERINA	R.LUMINAMISED	
101ML REH08	4.73M	MAPPED PERCEIPTO	4.70M (80-471)	
REPERENCE DENCME	DOTAL NOT	0.000		

46/47 reads

3636 reads

龄 Public Hea England	e leport	Report Table Laborations Reviewed By				
Patient D	etails		Requeste	r Details		
Patient Name	Bels Johnson		Reporter	ir.hui		
Parket ID	12366189			1294 Smith / 2 Elemengham, UK		
Patient Out	00-01-1908		Game to			
Leader	Oxford	Calleral Large to				
Sample D	surum	Sample In	*	11146478		



Relatedness		
	Likely Island Sex than 5.5M Difference	Peoplety Related (5-50-5NP Differences)
Number of lockstee	1	*

362

Public Health

England

For further information on related lociates and existing dicators, please contact the Public Health lab at 125-456-380

rt issued B	y: OXFORD	Report Date: 1 JAN 1900	
	Γ Δ	PATIENT INFORM	ATION
D	8	Name: Bob Advesor Birth Date: 1 Jan 1900 Location: Similingham	Identifier: 12345670 Sample Date: 1.Jan 1 Gender: M

MYCOBACTERIAL GENOME SEQUENCING REPORT

SPECIES IDENTIFIED BY SEQUENCING Ø 1005. identical to Myselbertwhen tedevooles/s



Ью

5

6



EPIDEMICLOGICAL RELATIONSHIPS Relarge to a chester of 8 genetically related cases, suggesting recen ranamitation



COMMENTS

high-puality data her analysis

Sequenced & Aug 2016 on an Illumina MiSeq, picking & 73M made, 4.70M (VP.475) mapped to the #31W (NC000962.2) relevance percent.

The sample was sequenced twice; the initial sequencing run did out provide



ing

totype	2		
	Prodiction	Gene (	Nutation
u	Reistant	luf5	8481

Rid, respire

#### Sequence Quality

The whole generate sequence analysis of the induits was considered <u>IEEE (SIAL IT</u> as the number of reads was granter the 6.7 million with 10.47% mapped and a converge of 91.09%.

#### **Reviewer Comments**

No additional commant

#### Authorization

lignation		Print Name	Gr. Julia Brailt
Date	41-01-0900	Posities	Lab Director



Public Health England

**Technical Details** 

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

#### Resistatupe

The resistivitype describes the mutations that are predicted to confer dwg resistance

Owag	Gene	Mutation	Grang	Cevenage	Support
Isoriaatid	Aut5	58157	Mykrole v2	iCx.	46/17 reads
Risepin	and a	55.571.	Willerstal	394	38/08 mich

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

		-	holate	Ser	SMP Oletanee
z			2945_4	294.5	3
8			2014_4	205.6	
9	2		2943,4	294.3	8
ł			3018_8	294.0	2
ê	1		204.2_4	294.5	10
			2012_8	294.5	9
	9	2041 2912 2810 2014 2815	3012_0	294.5	10
		Yest	2012_0	204.5	

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



**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

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



#### В

#### Relatedness

	Likely Related (less than 5 SNP Difference)	Possibly Related (6-30 SNP Differences)
Number of isolates	2	6





#### — Relatedness

Ε

Isolate Name	SNP difference
2015_A	3
2014_A	4
2013_A	8
2013_B	7
2012_A	10
2012_B	9
2012_C	10
2012_D	9

Data on relatedness to other isolates clusters is presented below in a

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

number of different formats Which do you find most interpretable 2 3 5 6 15. Physician 10 6 Choice 5 5 3 3 3 33 of cases with spark line A 3 2 111 of isolates related table В count Table Graph of of isolates by SNP distance С 16 D Table Phylogenetic Tree 15 Е Non-Physician F 10 -9 8 8 8 667 6 6 5 5 5 5 5 -44 4 3 3 2 2 3 5 6 Response

- Related isolates with SNP difference details
- Summary with related isolates per year

**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

"If you can combine the phylogenetic tree with some kind of graph showing temporal spread that would be perfect. Adding geographical data would be a really helpful bonus too."

"I like tree best but I like tree formats in general so I am biased. C;A and F are of equal value to me."

"Not useful for clinician. you need to refer this question to public health officials who do contact tracing"



## 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	( <sup>6</sup> )		2/	¥./		Fairviev	River			KA	H		
Load Data	Pen	-10			ele Aleen	*/ */		Z					
All Patients 118			11		Zom	Guttenberg	Temperor de la contractica de		enneme				
Selected Patients 3			4	IL Avenu	A)	AT .	Central P	Park	9			-/	$\sum$
💬 Feedback				Avenue	Weehav	wken		short due				X	
Filter Patients	4			Long	Am		Avenue	$\wedge$	31st C	S Y	T	-19-	Y
- Location		Jos	$\mathbf{>}$	3///			P.			X	Flushing Me	adows-Corona	
Filter By: City	7	4	Jersev (	City		HB	S Brenzing	A	R	$\leq$	9		
New York			774	T.A			HT /		X	Leaflet   Map tiles by Stamen De	sign, CC BY 3.0	Map data © Oper	Union . StreetMap
	Snov	N 10 •	entries								Search:		
Filter By: State/Province		City	State	Lat 🔷	Long 🔶	caseID 🔶	originOfBir	th 🔶	gender 🔶	MethodOfDetection	tbType	diagnosisD	ate 🔶 t
Filter By: State/Province	1	City New	State	Lat 🔶	Long 🔶	caseID 🔶	originOfBir	th 🔶	gender 🔶	MethodOfDetection	tbType	diagnosisD	ate 🔶 t
Filter By: State/Province	1	City New York	State NY	40.74446	Long 🔷 -73.96239	caseID 🔶	originOfBir	th 🔶 1	gender 🔶 M	MethodOfDetection	tbType Active	diagnosisD 2014-09-17	ate 🔶 t 20
Filter By: State/Province	1	City New York New York	State NY NY	Lat ♦ 40.74446 40.73901	Long ♦ -73.96239 -73.96682	caseID ♦ 0002 0004	originOfBirt	th ∲ 1 1	gender ≑ M F	MethodOfDetection Conact investigation Pre-landing Surveillance	tbType Active Active	diagnosisD 2014-09-17 2014-02-13	ate
Filter By: State/Province	1 2 3	City New York New York	State NY NY NY NY	Lat 40.74446 40.73901 40.77706	Long ÷ -73.96239 -73.96682 -73.96111	caseID ♦ 0002 0004 0007	originOfBirt	th 🔶 1 1	gender 🔶 M F F	MethodOfDetection Conact investigation Pre-landing Surveillance Screening Program	tbType       Active       Active       Active	diagnosisD 2014-09-17 2014-02-13 2014-04-16	ate ∲ t 20 20 20
Filter By: State/Province         Demographics         Diagnosis         Treatment         Outcome	1 2 3 4	City New York New York New York	State NY NY NY NY	Lat 40.74446 40.73901 40.77706 40.77012	Long -73.96239 -73.96682 -73.96111 -73.95928	caseID ♦ 0002 0004 0007 0008	originOfBirt	th ∳ 1 1 1 0	gender 🔶 M F F	MethodOfDetection Conact investigation Pre-landing Surveillance Screening Program Unknown	tbType       Active       Active       Active       Active       Active	diagnosisD 2014-09-17 2014-02-13 2014-04-16 2014-01-28	ate ∳ t 20 20 20 20
Filter By: State/Province         Demographics         Diagnosis         Treatment         Outcome	1 2 3 4 5	City New York New York New York New York	State NY NY NY NY NY NY	Lat 40.74446 40.73901 40.77706 40.77012 40.73109	Long ( -73.96239 -73.96682 -73.96111 -73.95928 -73.97850	<ul> <li>caseID (</li> <li>0002</li> <li>0004</li> <li>0007</li> <li>0008</li> <li>0009</li> </ul>	originOfBirt	th ≑ 1 1 1 0 0	gender 🔶 M F F F F	MethodOfDetection         Conact investigation         Pre-landing         Surveillance         Screening Program         Unknown         Unknown	tbType       Active       Active       Active       Active       Active       Active       Active	diagnosisD 2014-09-17 2014-02-13 2014-04-16 2014-01-28 2015-01-10	ate (1) 1 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2
Filter By: State/Province	1 2 3 4 5 6	City New York New York New York New York	State NY NY NY NY NY NY	Lat 40.74446 40.73901 40.77006 40.77012 40.73109	Long -73.96239 -73.96682 -73.96111 -73.95928 -73.97850 -73.98853	<ul> <li>caseID</li> <li>0002</li> <li>0004</li> <li>0007</li> <li>0008</li> <li>0009</li> <li>0019</li> </ul>	originOfBir	th \$	gender M F F F F F F F F F F F F F F F F F F	MethodOfDetection         Conact investigation         Pre-landing         Surveillance         Screening Program         Unknown         Screening Program	tbType       Active       Active       Active       Active       Active       Active       Active       Active       Active	diagnosisD 2014-09-17 2014-02-13 2014-04-16 2014-01-28 2015-01-10 2014-06-11	ate ♦ t 20 20 20 20 20 20 20 20 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



### **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!



# **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





# **DESCRIBING VISUALIZATIONS**

Using geometric marks and their properties (channels)



### **DESCRIBING DISEASE DYNAMICS** EXAMPLE 1: PHYLOGENETIC TREE + DOT PLOT



### **DESCRIBING DISEASE DYNAMICS** EXAMPLE 2: NETWORK DIAGRAM



### **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		
	Horizontal + Vertical pos.		
Case	Black/White Dot		
ommaney	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



