

IDEAS

OCCUPATION

AGE

METADATA

ID

STATUS

sick
exposed but well

HOSPITAL

WARD

A

B

C

G₁

F

E

D

FECES
EXPOSURESEX
EXPOSUREDROPLET
EXPOSURE

SOURCE

DATE OF
ONSET

INCUBATION

7 Days

NA

2017-10-22

2017-09-21

exposure

→ distribution of exposure?

other given data:

- map of hospital wards
- phylogenetic tree of IDs

FILTER

↳ relevance of metadata

- ID
- STATUS
- AGE → vulnerability?
- OCCUPATION → transmitters?
- HOSPITAL WARD

↳ phylogenetic tree has no branch lengths → cladogram

↳ status "exposed but well" has no "date of onset"

↳ "incubation" is hard to fit with exposure dates/onset date

↳ some patients have more than one mode of exposure

↳ map of hospital wards shows six wards but only 4 wards are present in the dataset (A, B, C, D)

CATEGORIZE

↳ transmission maps

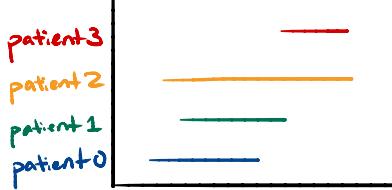
Source



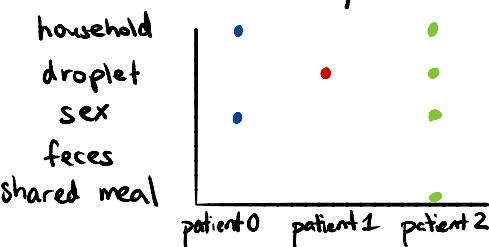
hospital wards

A → B

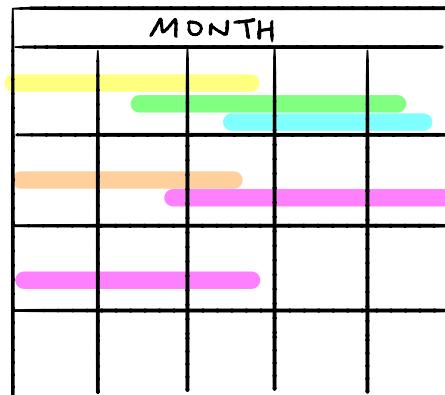
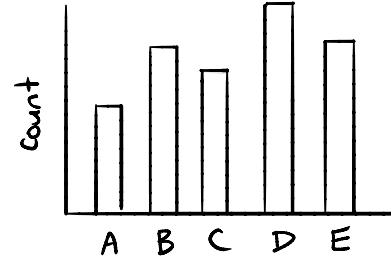
timeelines



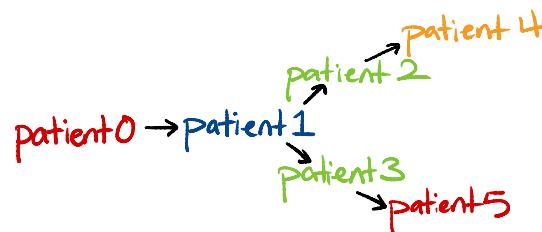
↳ exposure summary



COMBINE & REFINED



Legend: Ward A (red), Ward B (blue), Ward C (green), Ward D (orange), Ward E (purple), Ward F (pink)



BIOF 520: Identifying and Communicating the Source of the Zombie Outbreak

Diana Lin^{1,2}

¹Canada's Michael Smith Genome Sciences Centre, BC Cancer, Vancouver, BC, Canada

²Bioinformatics Graduate Program, University of British Columbia, Vancouver, BC, Canada

ABOUT MY VISUALIZATIONS

First, the summary bar chart is designed to help the user visually explore the data, which is often the first step in data analysis. The summary bar chart makes it easy for the user to get an overall feel of the data, and a composition breakdown of the dataset. As shown in the visualization, if you choose to look at the attribute exposure, and then colour by status, it is evident that while other modes of exposure do cause sickness, exposure by bite is the only mode where all victims become sick. **Second**, the exposure duration figure is intended to help to visualize the overlapping exposure durations, incubation periods, and onset dates. This is designed to help visualize the spread of the outbreak through time, and see when transmissions are occurring from patient to patient, or how the date of onset relates to date of exposure or incubation period. This figure is meant to help visualize the 3 sets of dates given in the dataset in a more linear fashion, and then overlaid onto a monthly calendar view, which is intended to be more intuitive for the layperson. **Third**, the directed transmission map combines all the data given into one visualization. The main framework of this visualization is the transmission map, which connects the nodes using a directed arrow representing transmission. The nodes (i.e. the patients) in this map can be coloured by attributes. This map utilizes the source and ID data given in the dataset, but the rest of the data can be accessed by hovering over nodes or edges. Clicking on certain attributes will highlight other nodes with the same value for that attribute. There is also a timeline below the map to show the speed of spread of the outbreak. Moving the slider will show the spread as a sort of time-lapse. (See OPERATIONS on Page 5 for more details). As shown in the visualization, the transmission map is coloured by ward, and the attributes selected are status: sick and exposure: bite. This visualization is designed to show all the information in a compact way, where certain highlighting and colouring aesthetics can be toggled. **I hypothesize that the source of the zombie outbreak is the rat, and that the mechanism of spread is exposure by bite. This hypothesis is supported by the third visualization, where the rat is at the root of the transmission map, and in the first visualization, where all bites result in sick patients.** However, not all sick patients are a result of exposure by bite. By looking at the third visualization and checking the exposure method for those patients, it is revealed that these sick patients have had sex with another patient who was bitten.

MY FEEDBACK

I did not receive much feedback due to lack of time for discussion, but it was mentioned that I should add a legend in the **second** visualization, and possibly incorporate the timeline into another visualization. A general suggestion given to all was to calculate new attributes from those given.

CHANGES

Following the feedback I received, I added a multiple selection checkbox to the second visualization, where multiple patient IDs can be selected for viewing. I also added a timeline in the **third** visualization-- one at the bottom, and an interactive slider to view the various time states of outbreak and watch the spread progress. I would also add an option to toggle the patient IDs to be replaced by occupation. Depending on the number of attributes given, I would also implement dropdown menus instead of sub-tabs and single selection boxes. (See DETAILS on Page 5). I would also add an additional column to the dataset, converting exposure start and end to exposure period.

BIOF 520: Identifying and Communicating the Source of the Zombie Outbreak

Diana Lin^{1,2}

¹Canada's Michael Smith Genome Sciences Centre, BC Cancer, Vancouver, BC, Canada

²Bioinformatics Graduate Program, University of British Columbia, Vancouver, BC, Canada

ABOUT MY VISUALIZATIONS

First, the summary bar chart is designed to help the user visually explore the data, which is often the first step in data analysis. The summary bar chart makes it easy for the user to get an overall feel of the data, and a composition breakdown of the dataset. As shown in the visualization, if you choose to look at the attribute exposure, and then colour by status, it is evident that while other modes of exposure do cause sickness, exposure by bite is the only mode where all victims become sick. **Second**, the exposure duration figure is intended to help to visualize the overlapping exposure durations, incubation periods, and onset dates. This is designed to help visualize the spread of the outbreak through time, and see when transmissions are occurring from patient to patient, or how the date of onset relates to date of exposure or incubation period. This figure is meant to help visualize the 3 sets of dates given in the dataset in a more linear fashion, and then overlaid onto a monthly calendar view, which is intended to be more intuitive for the layperson. **Third**, the directed transmission map combines all the data given into one visualization. The main framework of this visualization is the transmission map, which connects the nodes using a directed arrow representing transmission. The nodes (i.e. the patients) in this map can be coloured by attributes. This map utilizes the source and ID data given in the dataset, but the rest of the data can be accessed by hovering over nodes or edges. Clicking on certain attributes will highlight other nodes with the same value for that attribute. There is also a timeline below the map to show the speed of spread of the outbreak. Moving the slider will show the spread as a sort of time-lapse. (See OPERATIONS on Page 5 for more details). As shown in the visualization, the transmission map is coloured by ward, and the attributes selected are status: sick and exposure: bite. This visualization is designed to show all the information in a compact way, where certain highlighting and colouring aesthetics can be toggled. **I hypothesize that the source of the zombie outbreak is the rat, and that the mechanism of spread is exposure by bite. This hypothesis is supported by the third visualization, where the rat is at the root of the transmission map, and in the first visualization, where all bites result in sick patients.** However, not all sick patients are a result of exposure by bite. By looking at the third visualization and checking the exposure method for those patients, it is revealed that these sick patients have had sex with another patient who was bitten.

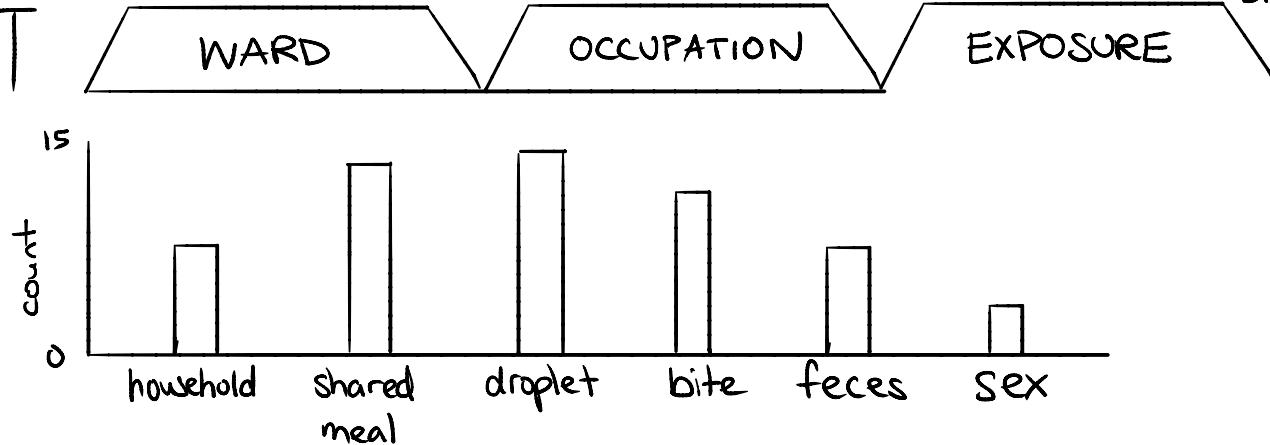
MY FEEDBACK

I did not receive much feedback due to lack of time for discussion, but it was mentioned that I should add a legend in the **second** visualization, and possibly incorporate the timeline into another visualization. A general suggestion given to all was to calculate new attributes from those given.

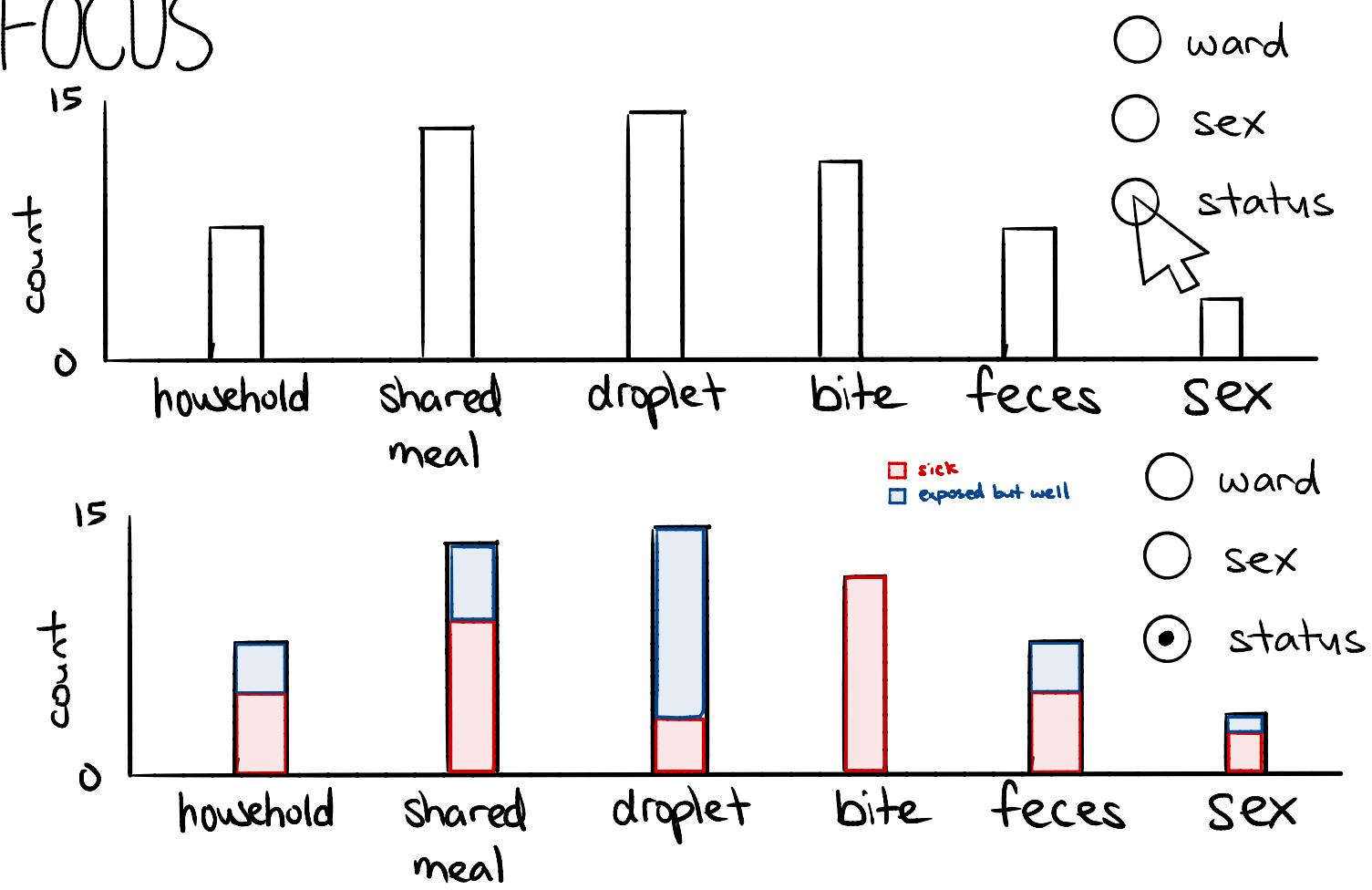
CHANGES

Following the feedback I received, I added a multiple selection checkbox to the second visualization, where multiple patient IDs can be selected for viewing. I also added a timeline in the **third** visualization-- one at the bottom, and an interactive slider to view the various time states of outbreak and watch the spread progress. I would also add an option to toggle the patient IDs to be replaced by occupation. Depending on the number of attributes given, I would also implement dropdown menus instead of sub-tabs and single selection boxes. (See DETAILS on Page 5). I would also add an additional column to the dataset, converting exposure start and end to exposure period.

LAYOUT



FOCUS



OPERATIONS

- each tab corresponds to a categorical attribute
- under each tab, another categorical attribute can be chosen by click
- there will be a more specific breakdown

DISCUSSION

Advantages:

- overview of counts for each attribute, with further breakdown
- see most common/rare values for each attribute

Disadvantages:

- cannot see all tabs at once (can't facet since x-axis differs)

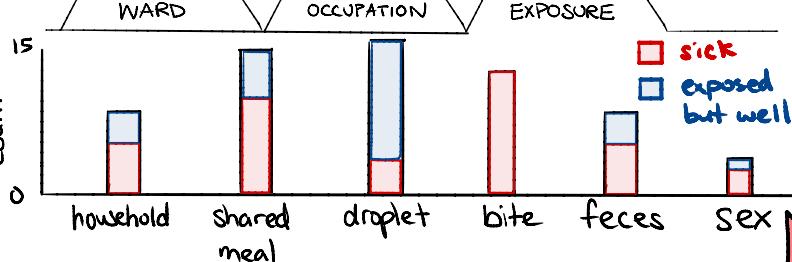
LAYOUT

OPERATIONS

DATA

BREAKDOWN

EXPOS



BREAKDOWN		
Ward	Sex	Status
●	●	●
●	●	●
●	●	●
●	●	●
●	●	●

DATA BREAKDOWN:

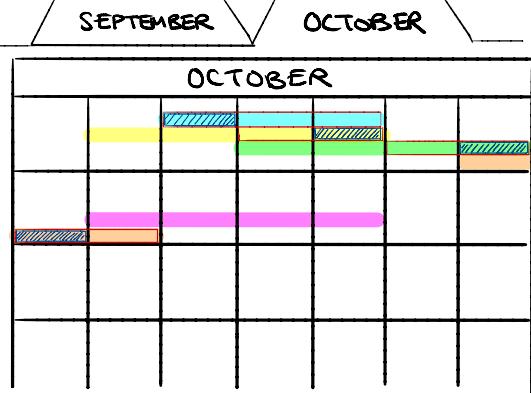
- there are sub tabs for each attribute
- under each subtab you can breakdown into one other attribute

EXPOSURE DURATIONS:

- subtabs are months in "calendar view"
- you can select multiple patients to view at once, and show incubation period or date of onset if applicable

OWN EXPOSURE DURATIONS

TRAN

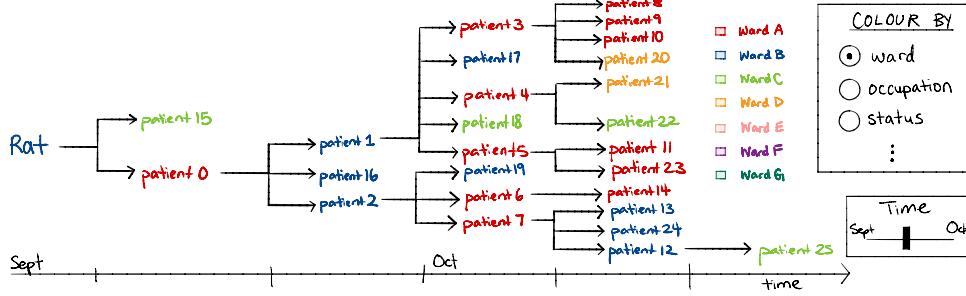


Patients

- patient 0
- patient 1
- patient 2
- patient 3
- patient 4
- patient 5
- patient 6
- patient 7
- patient 8
- patient 9
- patient 10
- patient 11
- patient 12
- patient 13
- patient 14
- patient 15
- patient 16
- patient 17
- patient 18
- patient 19
- patient 20
- patient 21
- patient 22
- patient 23
- patient 24
- patient 25

- incubation period
- date of onset

DURATIONS TRANSMISSION MAP



IMPLEMENTATION:

- an R shiny app dashboard with multiple tabs, allowing for user interaction
- html rendering so it can be viewed and interacted with in any browser
- each tab contains a different visualization aspect of the data
- subtabs under main tabs allow for visualization of individual attributes
- further breakdown of each subtab allows for simultaneous viewing of two attributes

MAIN TABS:

- Since this will be a webpage, main tabs can also be refreshed and have all three visualizations on one page

SUB TABS:

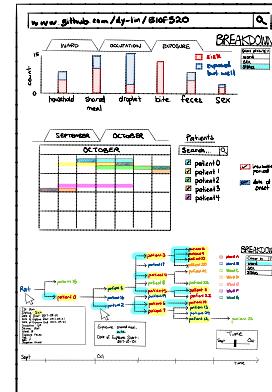
- Sub tabs can be converted to a dropdown menu if there are too many tabs (e.g. too many attributes)

SINGLE SELECTION BOXES

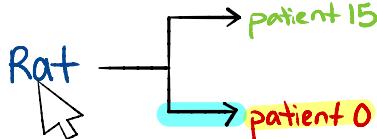
- single selection boxes can be converted to a dropdown menu as well in the case of too many options/attributes

MULTIPLE SELECTION CHECKBOXES

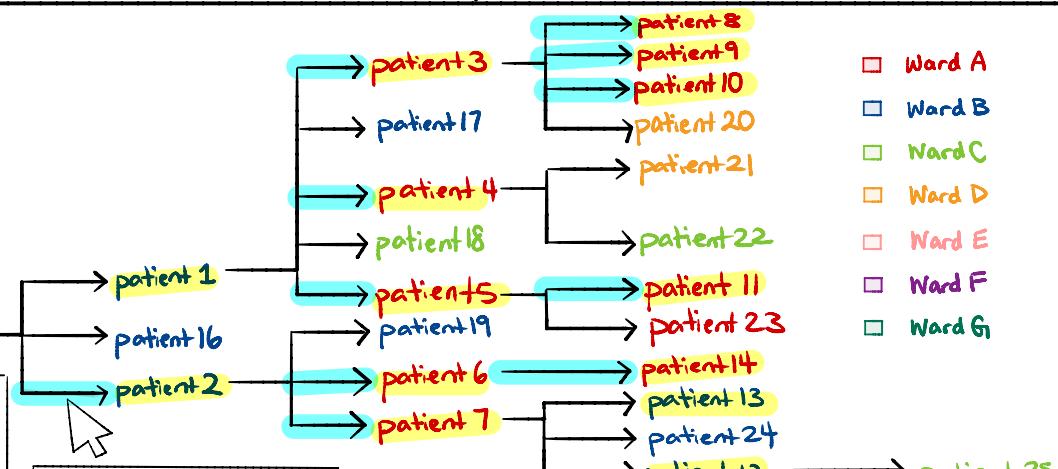
- a search bar will be added to search for options to select
- a "select all" checkbox will be added
- unselected attributes hidden in legend



FOCUS



ID: Rat
Status: Sick
Date of Onset: 2017-09-21
Date of Exposure Start: 2017-09-21
Date of Exposure End: 2017-09-21
Incubation: NA
Source: Rat
Ward: B
Exposure: Feces
Age: 1
Sex: F
Occupation: Animal



- Ward A
- Ward B
- Ward C
- Ward D
- Ward E
- Ward F
- Ward G

- ward
- occupation
- status

Time Sept Oct

Sept Oct

BIOF 520: Identifying and Communicating the Source of the Zombie Outbreak

Diana Lin^{1,2}

¹Canada's Michael Smith Genome Sciences Centre, BC Cancer, Vancouver, BC, Canada

²Bioinformatics Graduate Program, University of British Columbia, Vancouver, BC, Canada

ABOUT MY VISUALIZATIONS

First, the summary bar chart is designed to help the user visually explore the data, which is often the first step in data analysis. The summary bar chart makes it easy for the user to get an overall feel of the data, and a composition breakdown of the dataset. As shown in the visualization, if you choose to look at the attribute exposure, and then colour by status, it is evident that while other modes of exposure do cause sickness, exposure by bite is the only mode where all victims become sick. **Second**, the exposure duration figure is intended to help to visualize the overlapping exposure durations, incubation periods, and onset dates. This is designed to help visualize the spread of the outbreak through time, and see when transmissions are occurring from patient to patient, or how the date of onset relates to date of exposure or incubation period. This figure is meant to help visualize the 3 sets of dates given in the dataset in a more linear fashion, and then overlaid onto a monthly calendar view, which is intended to be more intuitive for the layperson. **Third**, the directed transmission map combines all the data given into one visualization. The main framework of this visualization is the transmission map, which connects the nodes using a directed arrow representing transmission. The nodes (i.e. the patients) in this map can be coloured by attributes. This map utilizes the source and ID data given in the dataset, but the rest of the data can be accessed by hovering over nodes or edges. Clicking on certain attributes will highlight other nodes with the same value for that attribute. There is also a timeline below the map to show the speed of spread of the outbreak. Moving the slider will show the spread as a sort of time-lapse. (See OPERATIONS on Page 5 for more details). As shown in the visualization, the transmission map is coloured by ward, and the attributes selected are status: sick and exposure: bite. This visualization is designed to show all the information in a compact way, where certain highlighting and colouring aesthetics can be toggled. **I hypothesize that the source of the zombie outbreak is the rat, and that the mechanism of spread is exposure by bite. This hypothesis is supported by the third visualization, where the rat is at the root of the transmission map, and in the first visualization, where all bites result in sick patients.** However, not all sick patients are a result of exposure by bite. By looking at the third visualization and checking the exposure method for those patients, it is revealed that these sick patients have had sex with another patient who was bitten.

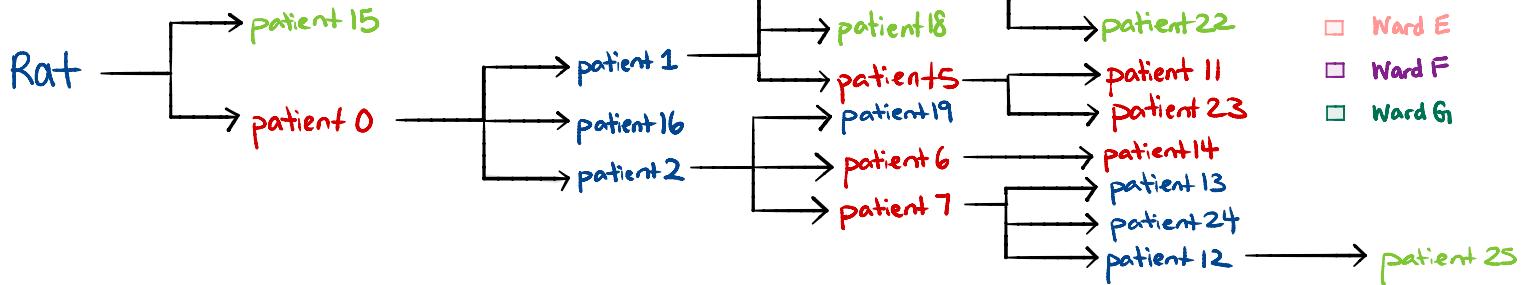
MY FEEDBACK

I did not receive much feedback due to lack of time for discussion, but it was mentioned that I should add a legend in the **second** visualization, and possibly incorporate the timeline into another visualization. A general suggestion given to all was to calculate new attributes from those given.

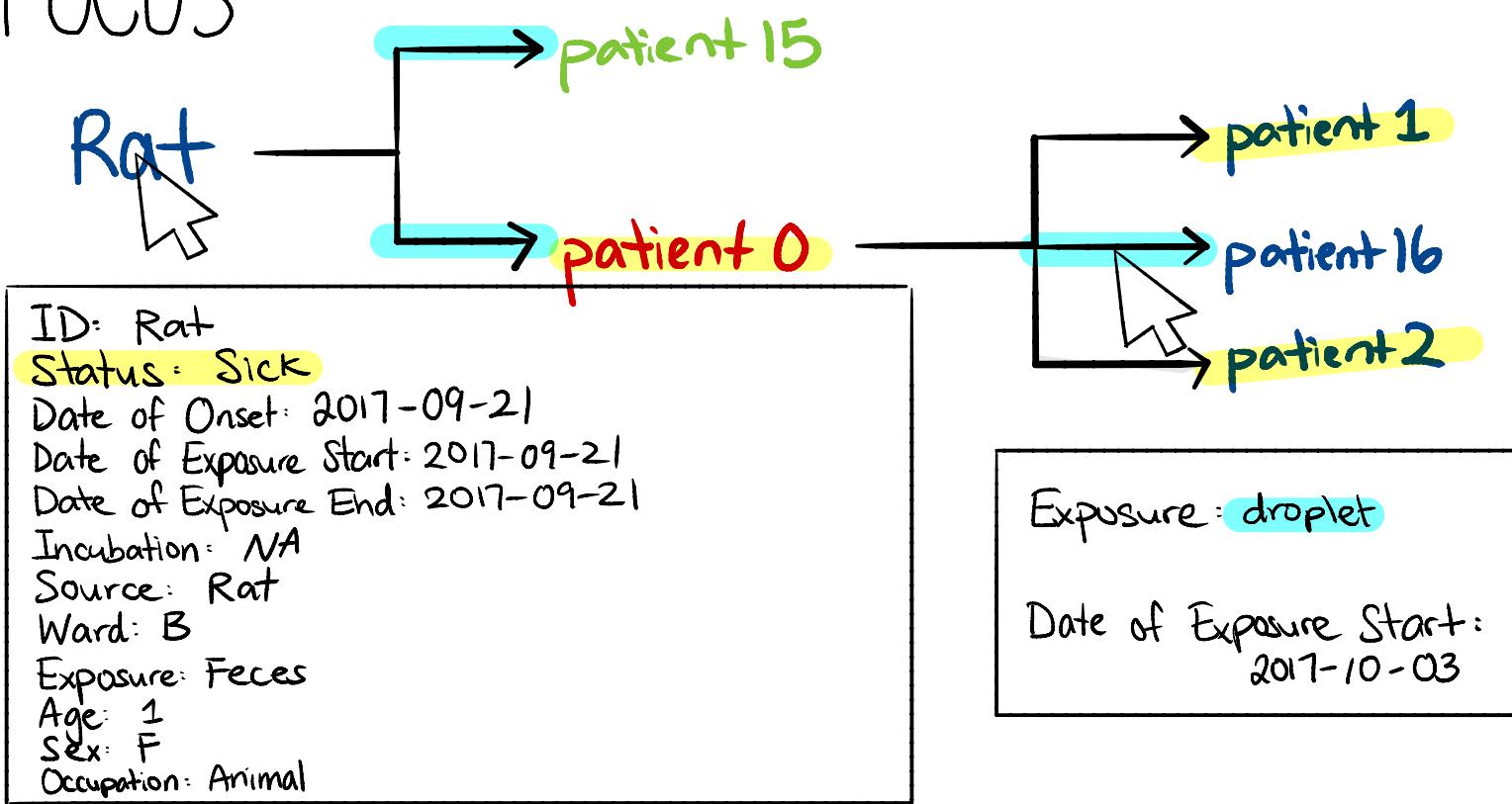
CHANGES

Following the feedback I received, I added a multiple selection checkbox to the second visualization, where multiple patient IDs can be selected for viewing. I also added a timeline in the **third** visualization-- one at the bottom, and an interactive slider to view the various time states of outbreak and watch the spread progress. I would also add an option to toggle the patient IDs to be replaced by occupation. Depending on the number of attributes given, I would also implement dropdown menus instead of sub-tabs and single selection boxes. (See DETAILS on Page 5). I would also add an additional column to the dataset, converting exposure start and end to exposure period.

LAYOUT



FOCUS



OPERATIONS

- hovering over each "arrow" of the transmission map reveals the mode of transmission/exposure
- hovering over each "tip" of the transmission map reveals all the metadata for that patient ID
- selecting an attribute from the pop-up window will highlight other "tips" with that same value for that attribute

DISCUSSION

Advantages:

- can see the ward-to-ward transmission at a glance
- only "relevant" metadata is shown unless hovering

Disadvantages:

- cannot see all exposure at once for whole map or just one mode

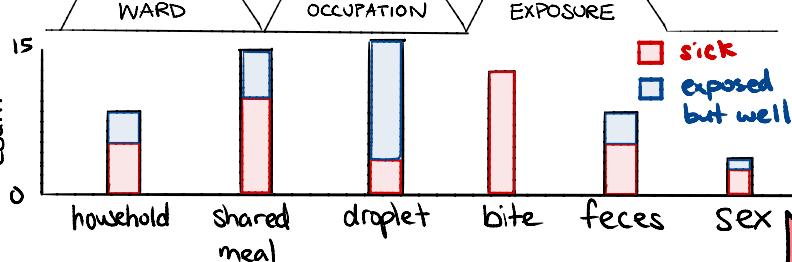
LAYOUT

OPERATIONS

DATA

BREAKDOWN

EXPOS



BREAKDOWN		
Ward	Sex	Status
Household	Male	Healthy
Shared meal	Female	Infected
Droplet	Male	Infected
Bite	Female	Infected
Feces	Male	Infected
Sex	Female	Infected

DATA BREAKDOWN:

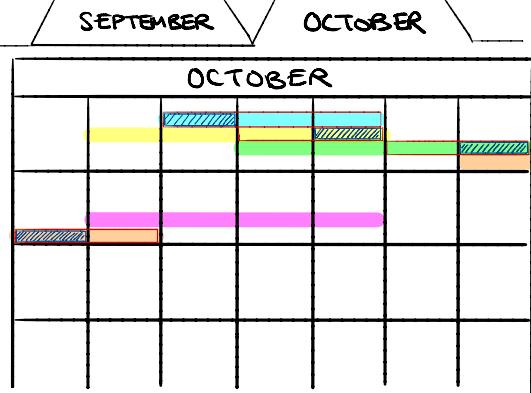
- there are sub tabs for each attribute
- under each subtab you can breakdown into one other attribute

EXPOSURE DURATIONS:

- subtabs are months in "calendar view"
- you can select multiple patients to view at once, and show incubation period or date of onset if applicable

OWN EXPOSURE DURATIONS

TRAN



Patients

- patient 0
- patient 1
- patient 2
- patient 3
- patient 4
- patient 5
- patient 6
- patient 7
- patient 8
- patient 9
- patient 10
- patient 11
- patient 12
- patient 13
- patient 14
- patient 15
- patient 16
- patient 17
- patient 18
- patient 19
- patient 20
- patient 21
- patient 22
- patient 23
- patient 24
- patient 25

- incubation period
- date of onset

DETAILS

IMPLEMENTATION:

- an R shiny app dashboard with multiple tabs, allowing for user interaction
- html rendering so it can be viewed and interacted with in any browser
- each tab contains a different visualization aspect of the data
- subtabs under main tabs allow for visualization of individual attributes
- further breakdown of each subtab allows for simultaneous viewing of two attributes

MAIN TABS:

- Since this will be a webpage, main tabs can also be refreshed and have all three visualizations on one page

SUB TABS:

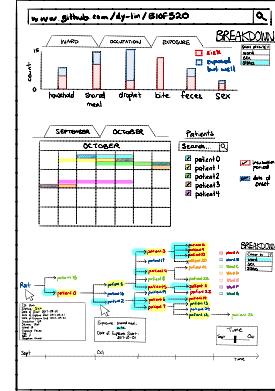
- Sub tabs can be converted to a dropdown menu if there are too many tabs (e.g. too many attributes)

SINGLE SELECTION BOXES

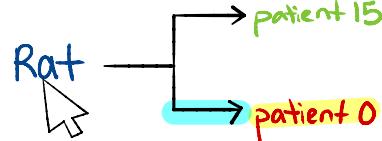
- single selection boxes can be converted to a dropdown menu as well in the case of too many options/attributes

MULTIPLE SELECTION CHECKBOXES

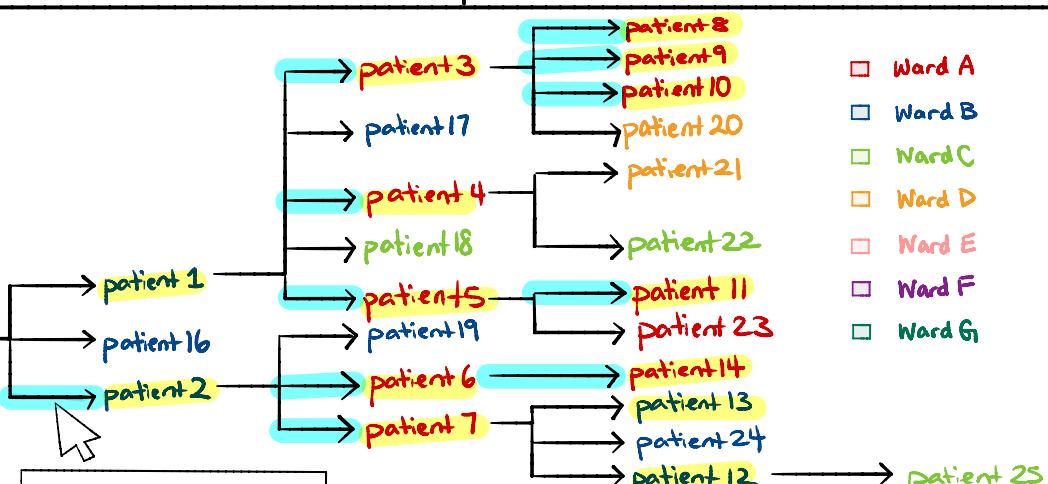
- a search bar will be added to search for options to select
- a "select all" checkbox will be added
- unselected attributes hidden in legend



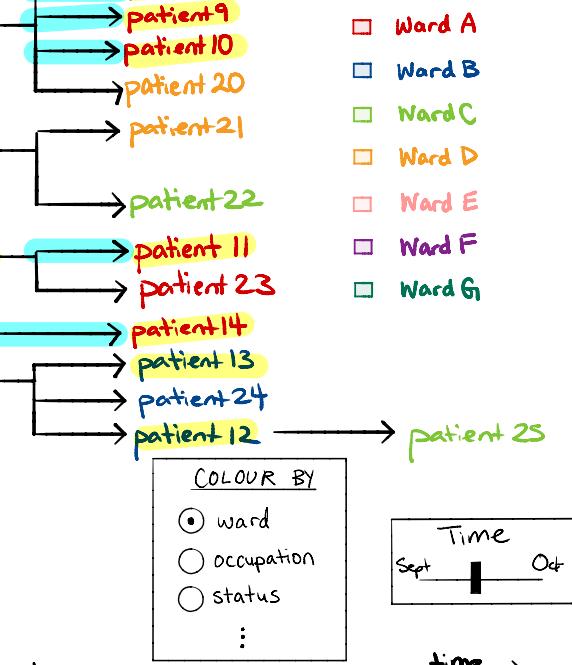
FOCUS



ID: Rat
Status: Sick
Date of Onset: 2017-09-21
Date of Exposure Start: 2017-09-21
Date of Exposure End: 2017-09-21
Incubation: NA
Source: Rat
Ward: B
Exposure: Feces
Age: 1
Sex: F
Occupation: Animal



Sept + Oct



time