How can
population data science
big health data
make an impact on health(care) system(s)?

Andriy Koval

Health System Impact Fellow

CIHR IRSC
Canadian Institutes of Health Research
Instituts de recherche en santé du Canada

British Columbia
BSERVATORY
Population & Public Health

THE UNIVERSITY OF BRITISH COLUMBIA

vancouver island health authority

Plan for today

- A sketch to set the stage
- 5 things
- 2 graphs
- 5 things again

20 years ago

EHR Developers:

- Make your records digital and bask in clinically actionable knowledge!

Public:

- OK

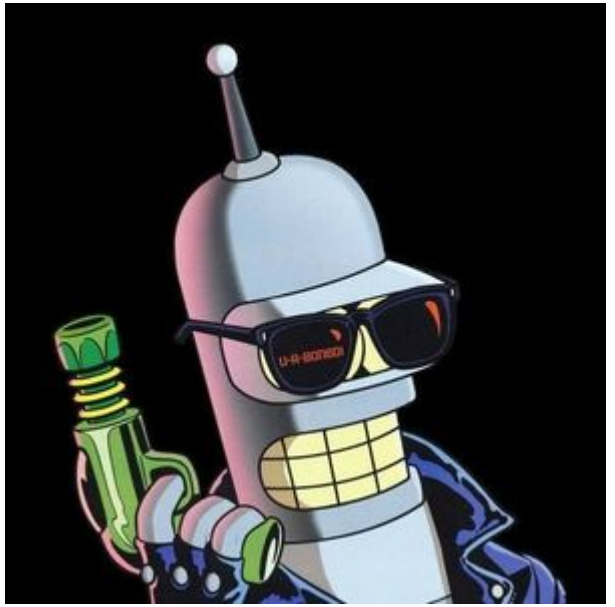
Today

Public:

- So what about basking in clinically actionable knowledge?

EHR Developers:

- Err… It turns out, EHR data is usually complex, frequently unstructured, and most of the time missing. And while making sense of it is often problematic and unreliable, do not worry, artificial intelligence will come to our rescue!



http://villains.wikia.com/wiki/Bender_Bending_Rodríguez?

The promise of big clinical data remains largely UNFULFILLED

How can
population data science
big health data
make an impact on health(care) system(s)?

What features of data science
What creative skills
What concepts

are important* to help fulfill this promise?

*in my opinion

How can
population data science
big health data
make an impact on health(care) system(s)?

Stated less optimistically:

IF there is any impact, what* would be responsible for it?

*in my opinion

Plan for today

- A silly sketch to set the stage
- 5 things
- 2 graphs
- 5 things again

TL;DR

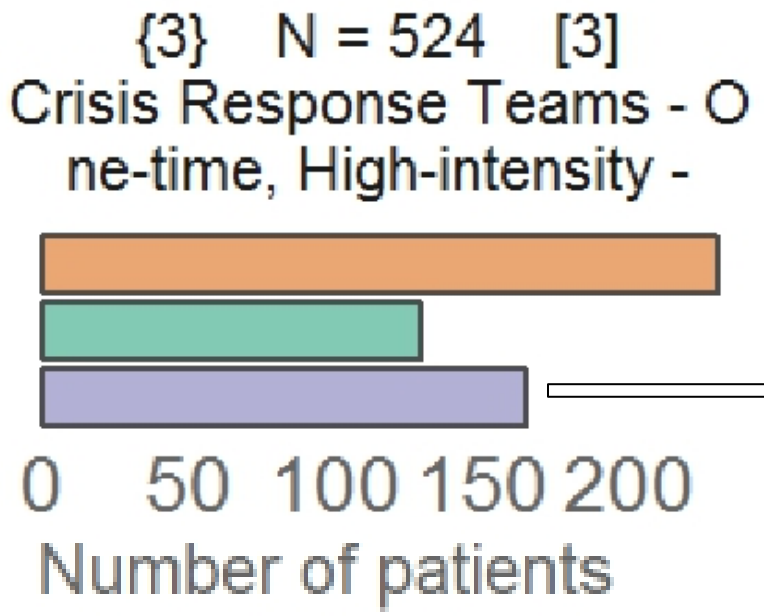
Hone these skills: this is applied health data science*

Concept	Skill	Central Idea	Expected Functional Benefits
Script Composition	Intelligent Programming	Programming script *as a* Medium of communication	Scripts could be read by humans. Less buggy and more transparent code.
File Architecture	Workflow Architecture	Reproducible project *as a* Non-linear narrative	Easier to remember and to collaborate. More reliable reproduction of analyses.
Display production	Data Visualization	Information display *as a* Scripted function	Presents results in approachable way. Easy to adjust script to other purposes.
Digital Publishing	Self Publishing	Dynamic report *as a* Platonic form	.R + .Rmd = html + pdf + .docx Greater control over appearance. Stronger online presence
Model Expression	Statistical Modeling	Model *as a* Simplification of reality	Improved team communication Deeper understanding of the phenomenon

* In my opinion

Plan for today

- A silly sketch to set the stage
- 5 things
- 2 graphs
- 5 things again

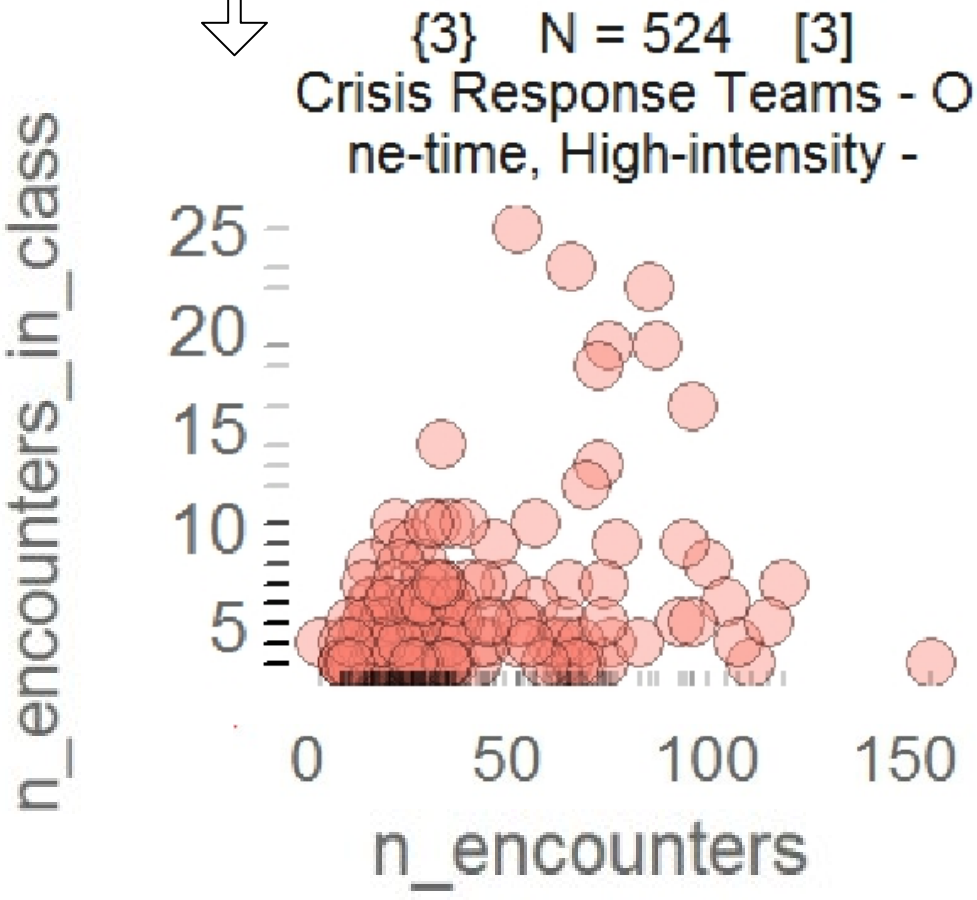


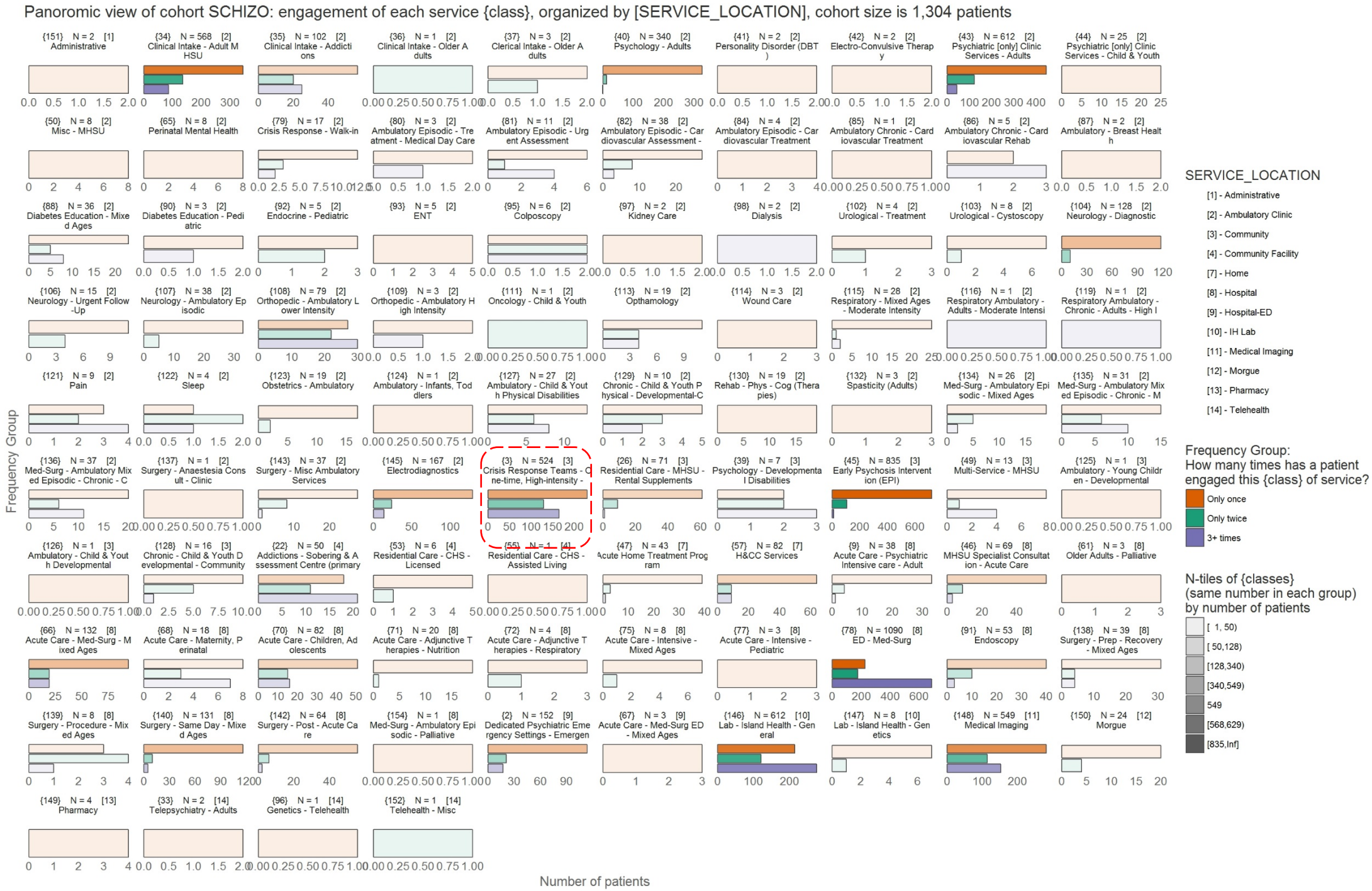
Consider that there is a cohort of patients (N = 1,304) who have been identified as having a form of SCHIZOPHRENIA.

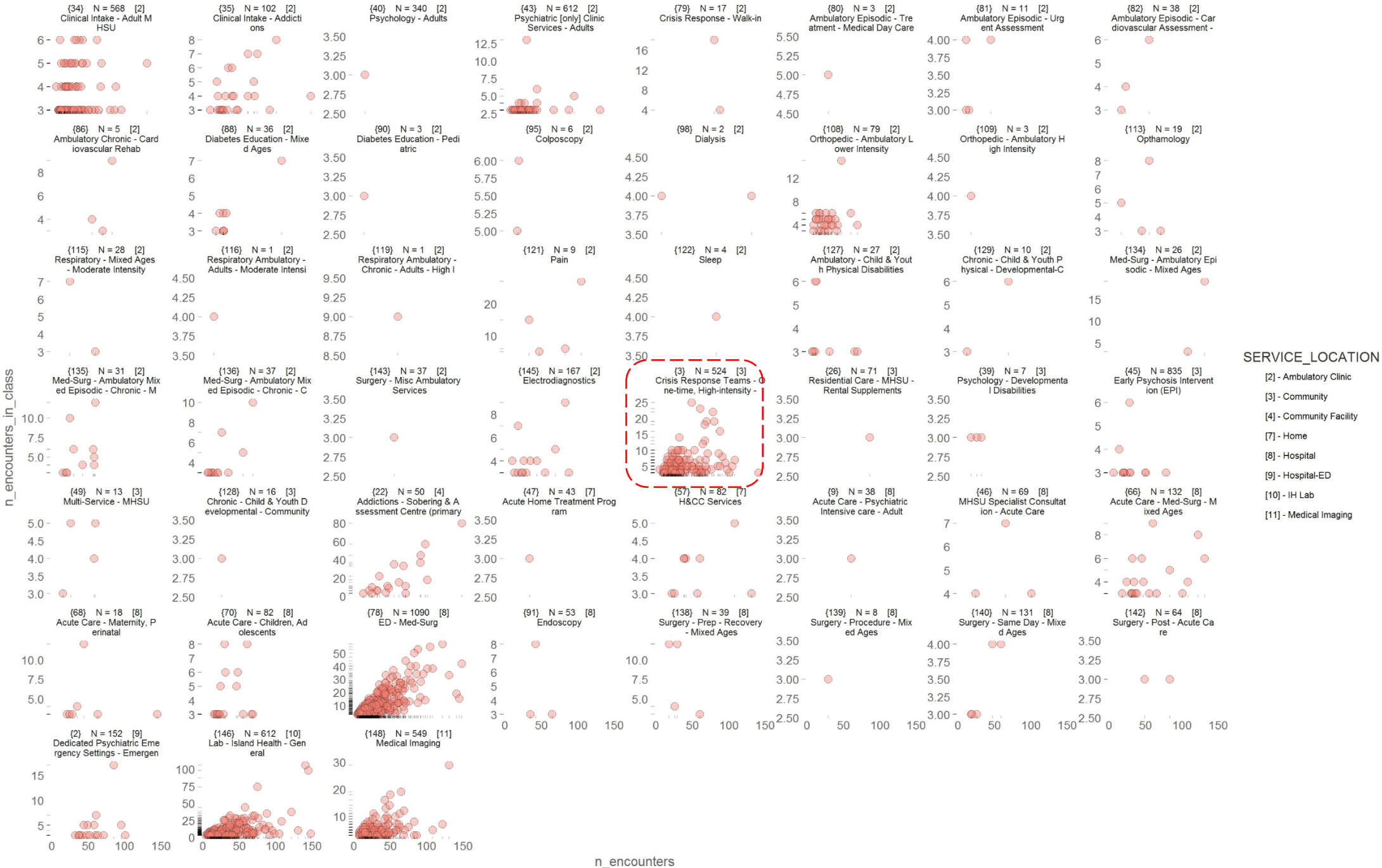
How were they using health services?
Which classes?
How often?

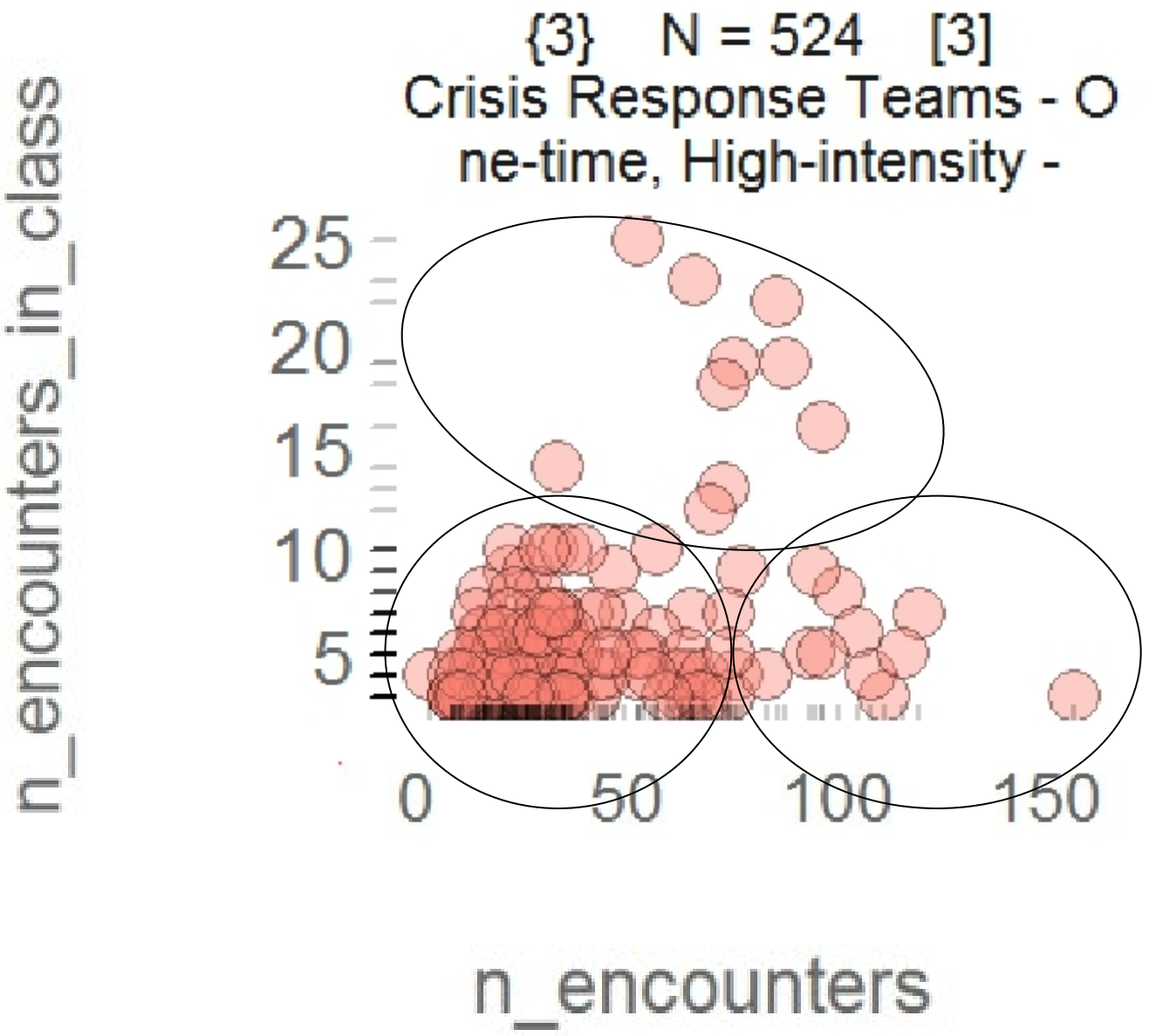
Frequency Group:
How many times has a patient engaged this {class} of service?

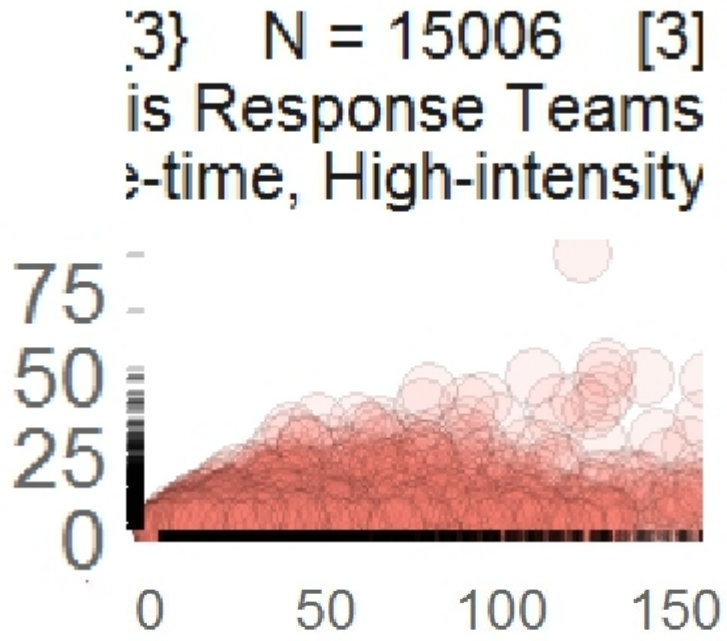
- Only once
- Only twice
- 3+ times




Panoromic view of cohort SCHIZO: engagement of each service {class}, organized by [SERVICE_LOCATION], cohort size is 1,304 patients
{151} N = 2 [1] Administrative
{34} N = 568 [2] Clinical Intake - Adult MHSU
{35} N = 102 [2] Clinical Intake - Addictions
{36} N = 1 [2] Clinical Intake - Older Adults
{37} N = 3 [2] Clerical Intake - Older Adults
{40} N = 340 [2] Psychology - Adults
{41} N = 2 [2] Personality Disorder (DBT)
{42} N = 2 [2] Electro-Convulsive Therapy
{43} N = 612 [2] Psychiatric [only] Clinic Services - Adults
{44} N = 25 [2] Psychiatric [only] Clinic Services - Child & Youth
{50} N = 8 [2] Misc - MHSU
{65} N = 8 [2] Perinatal Mental Health
{79} N = 17 [2] Crisis Response - Walk-in
{80} N = 3 [2] Ambulatory Episodic - Treatment - Medical Day Care
{81} N = 11 [2] Ambulatory Episodic - Urgent Assessment
{82} N = 38 [2] Ambulatory Episodic - Cardiovascular Assessment -
{84} N = 4 [2] Ambulatory Episodic - Cardiovascular Treatment
{85} N = 1 [2] Ambulatory Chronic - Cardiovascular Treatment
{86} N = 5 [2] Ambulatory Chronic - Cardiovascular Rehab
{87} N = 2 [2] Ambulatory - Breast Health
{88} N = 36 [2] Diabetes Education - Mixed Ages
{90} N = 3 [2] Diabetes Education - Pediatric
{92} N = 5 [2] Endocrine - Pediatric
{93} N = 5 [2] ENT
{95} N = 6 [2] Colposcopy
{97} N = 2 [2] Kidney Care
{98} N = 2 [2] Dialysis
{102} N = 4 [2] Urological - Treatment
{103} N = 8 [2] Urological - Cystoscopy
{104} N = 128 [2] Neurology - Diagnostic
{106} N = 15 [2] Neurology - Urgent Follow-Up
{107} N = 38 [2] Neurology - Ambulatory Episodic
{108} N = 79 [2] Orthopedic - Ambulatory Lower Intensity
{109} N = 3 [2] Orthopedic - Ambulatory High Intensity
{111} N = 1 [2] Oncology - Child & Youth
{113} N = 19 [2] Opthamology
{114} N = 3 [2] Wound Care
{115} N = 28 [2] Respiratory - Mixed Ages - Moderate Intensity
{116} N = 1 [2] Respiratory Ambulatory - Adults - Moderate Intensi
{119} N = 1 [2] Respiratory Ambulatory - Chronic - Adults - High I
{121} N = 9 [2] Pain
{122} N = 4 [2] Sleep
{123} N = 19 [2] Obstetrics - Ambulatory
{124} N = 1 [2] Ambulatory - Infants, Toddlers
{127} N = 27 [2] Ambulatory - Child & Youth Physical Disabilities
{129} N = 10 [2] Chronic - Child & Youth Physical - Developmental-C
{130} N = 19 [2] Rehab - Phys - Cog (Therapies)
{132} N = 3 [2] Spasticity (Adults)
{134} N = 26 [2] Med-Surg - Ambulatory Episodic - Mixed Ages
{135} N = 31 [2] Med-Surg - Ambulatory Mixed Episodic - Chronic - M
{136} N = 37 [2] Med-Surg - Ambulatory Mixed Episodic - Chronic - C
{137} N = 1 [2] Surgery - Anaestesia Consult - Clinic
{143} N = 37 [2] Surgery - Misc Ambulatory Services
{145} N = 167 [2] Electrodiagnostics
{3} N = 524 [3] Crisis Response Teams - One-time, High-intensity -
{26} N = 71 [3] Residential Care - MHSU - Rental Supplements
{39} N = 7 [3] Psychology - Developmental Disabilities
{45} N = 835 [3] Early Psychosis Intervention (EPI)
{49} N = 13 [3] Multi-Service - MHSU
{125} N = 1 [3] Ambulatory - Young Children - Developmental
{126} N = 1 [3] Ambulatory - Child & Youth Developmental
{128} N = 16 [3] Chronic - Child & Youth Developmental - Community
{22} N = 50 [4] Addictions - Sobering & Assessment Centre (primary
{53} N = 6 [4] Residential Care - CHS - Licensed
{55} N = 1 [4] Residential Care - CHS - Assisted Living
{47} N = 43 [7] Acute Home Treatment Program
{57} N = 82 [7] H&CC Services
{9} N = 38 [8] Acute Care - Psychiatric Intensive care - Adult
{46} N = 69 [8] MHSU Specialist Consultation - Acute Care
{61} N = 3 [8] Older Adults - Palliative
{66} N = 132 [8] Acute Care - Med-Surg - Mixed Ages
{68} N = 18 [8] Acute Care - Maternity, Perinatal
{70} N = 82 [8] Acute Care - Children, Adolescents
{71} N = 20 [8] Acute Care - Adjunctive Therapies - Nutrition
{72} N = 4 [8] Acute Care - Adjunctive Therapies - Respiratory
{75} N = 8 [8] Acute Care - Intensive - Mixed Ages
{77} N = 3 [8] Acute Care - Intensive - Pediatric
{78} N = 1090 [8] ED - Med-Surg
{91} N = 53 [8] Endoscopy
{138} N = 39 [8] Surgery - Prep - Recovery - Mixed Ages
{139} N = 8 [8] Surgery - Procedure - Mixed Ages
{140} N = 131 [8] Surgery - Same Day - Mixed Ages
{142} N = 64 [8] Surgery - Post - Acute Care
{154} N = 1 [8] Med-Surg - Ambulatory Episodic - Palliative
{2} N = 152 [9] Dedicated Psychiatric Emergency Settings - Emergen
{67} N = 3 [9] Acute Care - Med-Surg ED - Mixed Ages
{146} N = 612 [10] Lab - Island Health - General
{147} N = 8 [10] Lab - Island Health - Genetics
{148} N = 549 [11] Medical Imaging
{150} N = 24 [12] Morgue
{149} N = 4 [13] Pharmacy
{33} N = 2 [14] Telepsychiatry - Adults
{96} N = 1 [14] Genetics - Telehealth
{152} N = 1 [14] Telehealth - Misc
Frequency Group
Number of patients
SERVICE_LOCATION
[1] - Administrative
[2] - Ambulatory Clinic
[3] - Community
[4] - Community Facility
[7] - Home
[8] - Hospital
[9] - Hospital-ED
[10] - IH Lab
[11] - Medical Imaging
[12] - Morgue
[13] - Pharmacy
[14] - Telehealth
Frequency Group:
How many times has a patient engaged this {class} of service?
Only once
Only twice
3+ times
N-tiles of {classes}
(same number in each group)
by number of patients
[1, 50)
[50,128)
[128,340)
[340,549)
549
[568,629)
[835,Inf]


Panoromic view of cohort SCHIZO: engagement of each service {class}, organized by [SERVICE_LOCATION], cohort size is 1,304 patients
{34} N = 568 [2] Clinical Intake - Adult M HSU
{35} N = 102 [2] Clinical Intake - Addicti ons
{40} N = 340 [2] Psychology - Adults
{43} N = 612 [2] Psychiatric [only] Clinic Services - Adults
{79} N = 17 [2] Crisis Response - Walk-in
{80} N = 3 [2] Ambulatory Episodic - Tre atment - Medical Day Care
{81} N = 11 [2] Ambulatory Episodic - Urg ent Assessment
{82} N = 38 [2] Ambulatory Episodic - Car diovascular Assessment -
{86} N = 5 [2] Ambulatory Chronic - Card iovascular Rehab
{88} N = 36 [2] Diabetes Education - Mixe d Ages
{90} N = 3 [2] Diabetes Education - Pedi atric
{95} N = 6 [2] Colposcopy
{98} N = 2 [2] Dialysis
{108} N = 79 [2] Orthopedic - Ambulatory L ower Intensity
{109} N = 3 [2] Orthopedic - Ambulatory H igh Intensity
{113} N = 19 [2] Opthamology
{115} N = 28 [2] Respiratory - Mixed Ages - Moderate Intensity
{116} N = 1 [2] Respiratory Ambulatory - Adults - Moderate Intensi
{119} N = 1 [2] Respiratory Ambulatory - Chronic - Adults - High I
{121} N = 9 [2] Pain
{122} N = 4 [2] Sleep
{127} N = 27 [2] Ambulatory - Child & Yout h Physical Disabilities
{129} N = 10 [2] Chronic - Child & Youth P hysical - Developmental-C
{134} N = 26 [2] Med-Surg - Ambulatory Epi sodic - Mixed Ages
{135} N = 31 [2] Med-Surg - Ambulatory Mix ed Episodic - Chronic - M
{136} N = 37 [2] Med-Surg - Ambulatory Mix ed Episodic - Chronic - C
{143} N = 37 [2] Surgery - Misc Ambulatory Services
{145} N = 167 [2] Electrodiagnostics
{3} N = 524 [3] Crisis Response Teams - O ne-time, High-intensity -
{26} N = 71 [3] Residential Care - MHSU - Rental Supplements
{39} N = 7 [3] Psychology - Developmenta l Disabilities
{45} N = 835 [3] Early Psychosis Intervent ion (EPI)
{49} N = 13 [3] Multi-Service - MHSU
{128} N = 16 [3] Chronic - Child & Youth D evelopmental - Community
{22} N = 50 [4] Addictions - Sobering & A ssessment Centre (primary
{47} N = 43 [7] Acute Home Treatment Prog ram
{57} N = 82 [7] H&CC Services
{9} N = 38 [8] Acute Care - Psychiatric Intensive care - Adult
{46} N = 69 [8] MHSU Specialist Consultat ion - Acute Care
{66} N = 132 [8] Acute Care - Med-Surg - M ixed Ages
{68} N = 18 [8] Acute Care - Maternity, P erinatal
{70} N = 82 [8] Acute Care - Children, Ad olescents
{78} N = 1090 [8] ED - Med-Surg
{91} N = 53 [8] Endoscopy
{138} N = 39 [8] Surgery - Prep - Recovery - Mixed Ages
{139} N = 8 [8] Surgery - Procedure - Mix ed Ages
{140} N = 131 [8] Surgery - Same Day - Mixe d Ages
{142} N = 64 [8] Surgery - Post - Acute Ca re
{2} N = 152 [9] Dedicated Psychiatric Eme rgency Settings - Emergen
{146} N = 612 [10] Lab - Island Health - Gen eral
{148} N = 549 [11] Medical Imaging
n_encounters_in_class
n_encounters
SERVICE_LOCATION
[2] - Ambulatory Clinic
[3] - Community
[4] - Community Facility
[7] - Home
[8] - Hospital
[9] - Hospital-ED
[10] - IH Lab
[11] - Medical Imaging


{3} N = 524 [3]
Crisis Response Teams - O
ne-time, High-intensity -
n_encounters_in_class
25
20
15
10
5
0
50
100
150
n_encounters


3} N = 15006 [3]
is Response Teams
-time, High-intensity
75
50
25
0
0
50
100
150

IHACRU / **bcopph-mhsu-surveillance** Private

Watch 4 | Star 0 | Fork 0

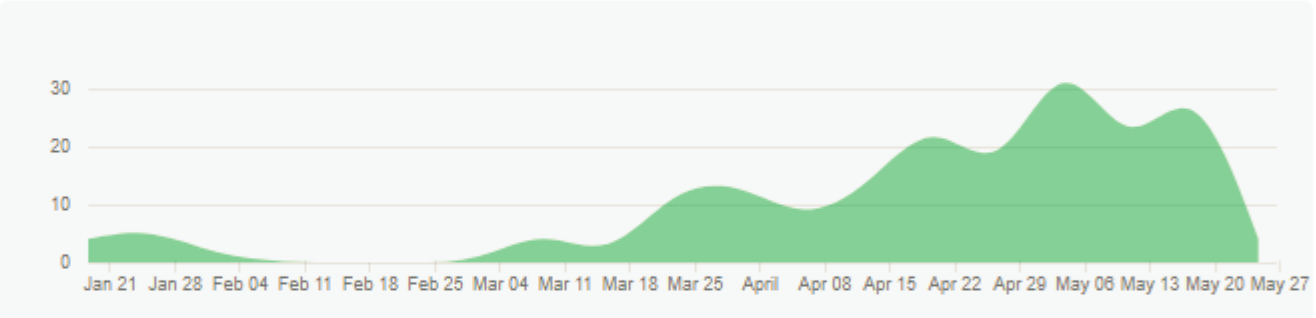
Code | Issues 5 | Pull requests 0 | Projects 0 | Wiki | Insights | Settings

- Pulse
- **Contributors**
- Traffic
- Commits
- Code frequency
- Dependency graph
- Network
- Forks

Jan 21, 2018 – May 29, 2018

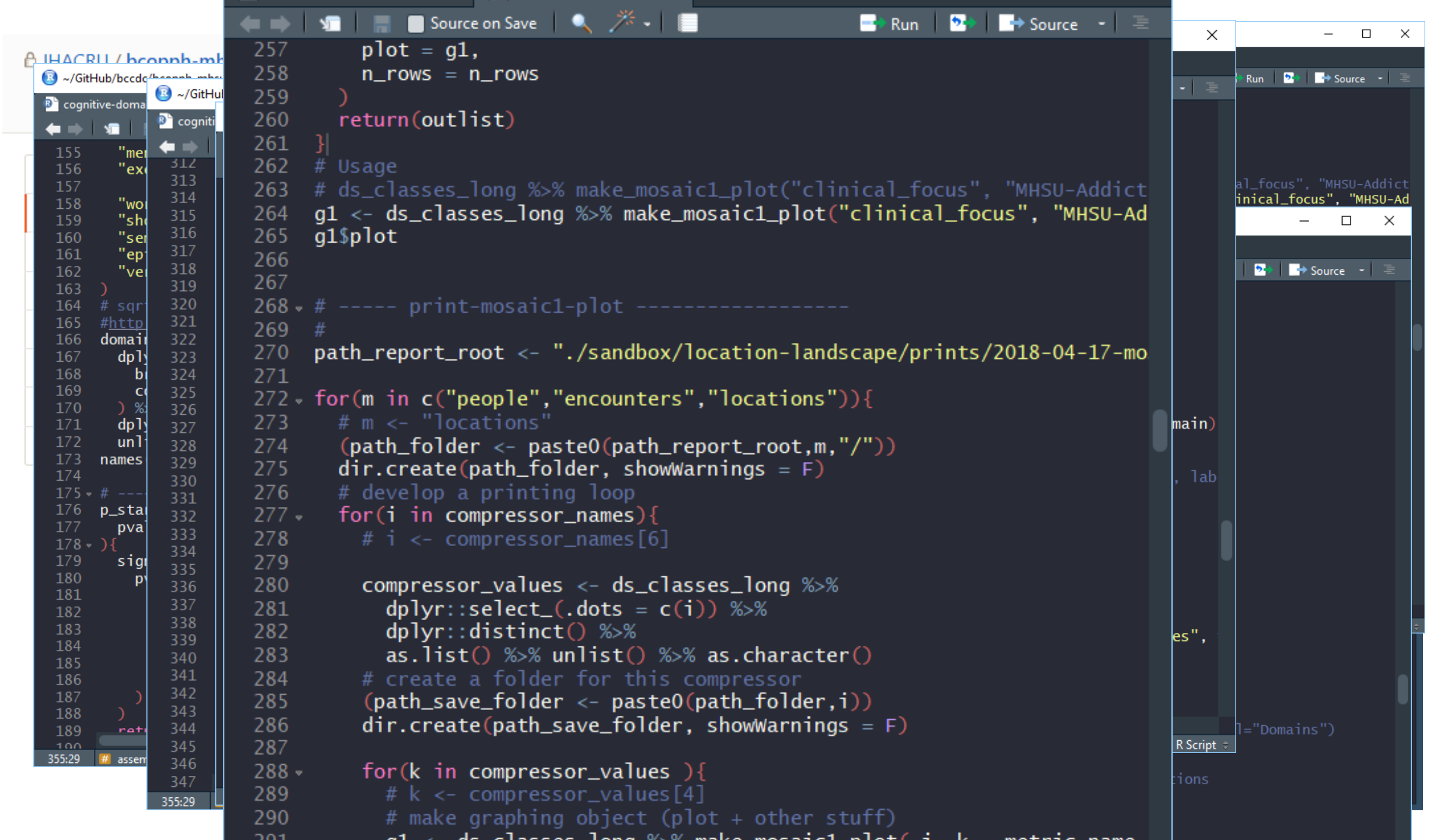
Contributions: Commits

Contributions to master, excluding merge commits



andkov #1

193 commits 1,857,369 ++ 1,765,443 --


Source on Save
Run
Source
plot = g1,
n_rows = n_rows
)
return(outlist)
}
# Usage
# ds_classes_long %>% make_mosaic1_plot("clinical_focus", "MHSU-Addict
g1 <- ds_classes_long %>% make_mosaic1_plot("clinical_focus", "MHSU-Ad
g1$plot
----- print-mosaic1-plot ------------------
#
path_report_root <- "./sandbox/location-landscape/prints/2018-04-17-mo
for(m in c("people","encounters","locations")){
# m <- "locations"
(path_folder <- paste0(path_report_root,m,"/"))
dir.create(path_folder, showWarnings = F)
# develop a printing loop
for(i in compressor_names){
# i <- compressor_names[6]
compressor_values <- ds_classes_long %>%
dplyr::select_(.dots = c(i)) %>%
dplyr::distinct() %>%
as.list() %>% unlist() %>% as.character()
# create a folder for this compressor
(path_save_folder <- paste0(path_folder,i))
dir.create(path_save_folder, showWarnings = F)
for(k in compressor_values){
# k <- compressor_values[4]
# make graphing object (plot + other stuff)
R Script

Concept	Skill	Central Idea	Expected Functional Benefits
Script Composition	Intelligent Programming	Programming script *as a* Medium of communication	Scripts could be read by humans. Less buggy and more transparent code.
File Architecture	Workflow Architecture	Reproducible project *as a* Non-linear narrative	Easier to remember and to collaborate. More reliable reproduction of analyses.
Display production	Data Visualization	Information display *as a* Scripted function	Presents results in approachable way. Easy to adjust script to other purposes.
Digital Publishing	Self Publishing	Dynamic report *as a* Platonic form	.R + .Rmd = html + pdf + .docx Greater control over appearance. Stronger online presence
Model Expression	Statistical Modeling	Model *as a* Simplification of reality	Improved team communication Deeper understanding of the phenomenon

Acknowledgements

THE UNIVERSITY OF BRITISH COLUMBIA

British Columbia OBSERVATORY Population & Public Health

vancouver island health authority

CIHR IRSC
Canadian Institutes of Health Research
Instituts de recherche en santé du Canada

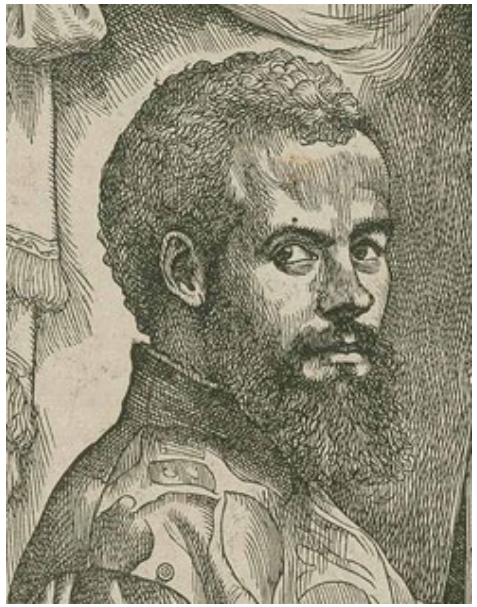
Health System Impact Fellowship

Kate Smolina

Ken Moselle

Appendix

Shoulders to stand on



Andreas Vesalius

John Tukey

Edward Tufte

Hadley Wickham

Model Manifestations

Tabular

	id	time	attend
1	1	0	1
2	1	1	6
3	1	2	2
4	1	3	1
5	1	4	1
6	1	5	1
7	1	6	1
8	1	7	1
9	1	8	1
10	1	9	1
11	1	10	1
12	1	11	1

Algebraic

$$y_{it} = \beta_0 + \beta_1 time_t + \varepsilon_{it}$$

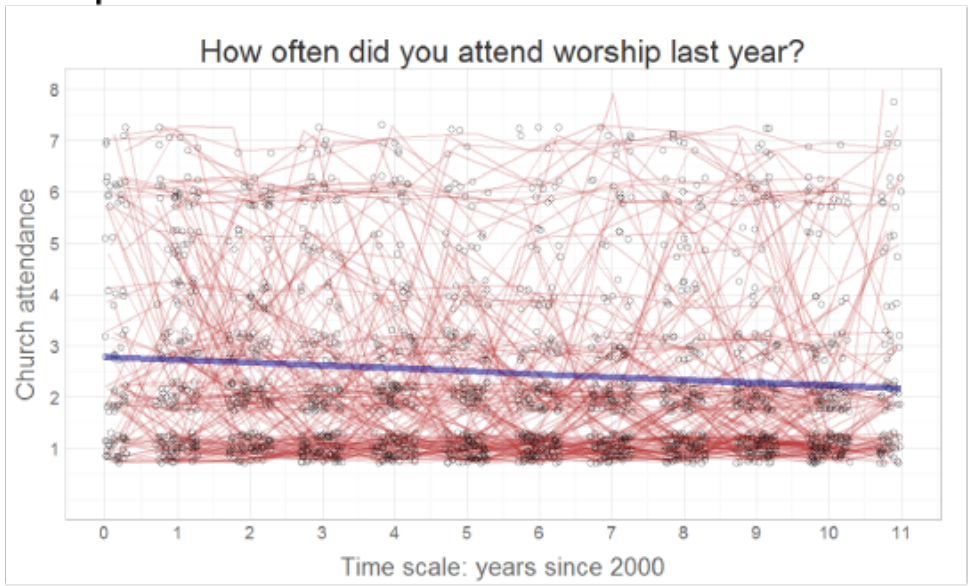
$$\beta_0 = \gamma_{00}$$

$$\beta_1 = \gamma_{10}$$

Semantic

In 2000 respondents attended church less than once a month (2.79) and gradually declined in their attendance since (.06 per year).

Graphical



Syntactic

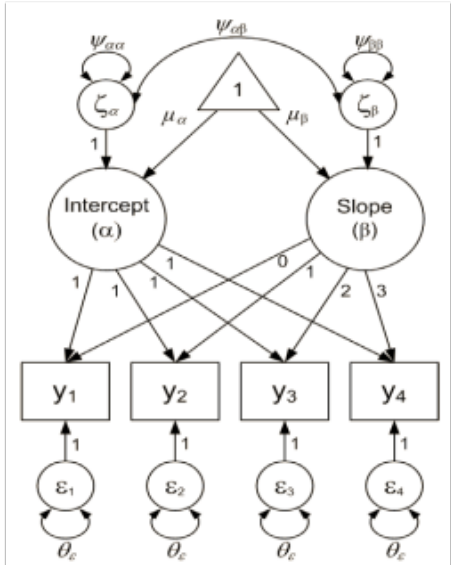
```
nlme::gls(attend ~ 1 + time, data=dsM)
```

```
Coefficients:
               Value Std.Error t-value p-value
(Intercept)  2.7882   0.07774   35.86       0
time        -0.0566   0.01197   -4.73       0
```

Numeric

	modelB
logLik	-3719
deviance	7438
AIC	7444
BIC	7461
df.resid	1858
N	1860
p	2
ids	155

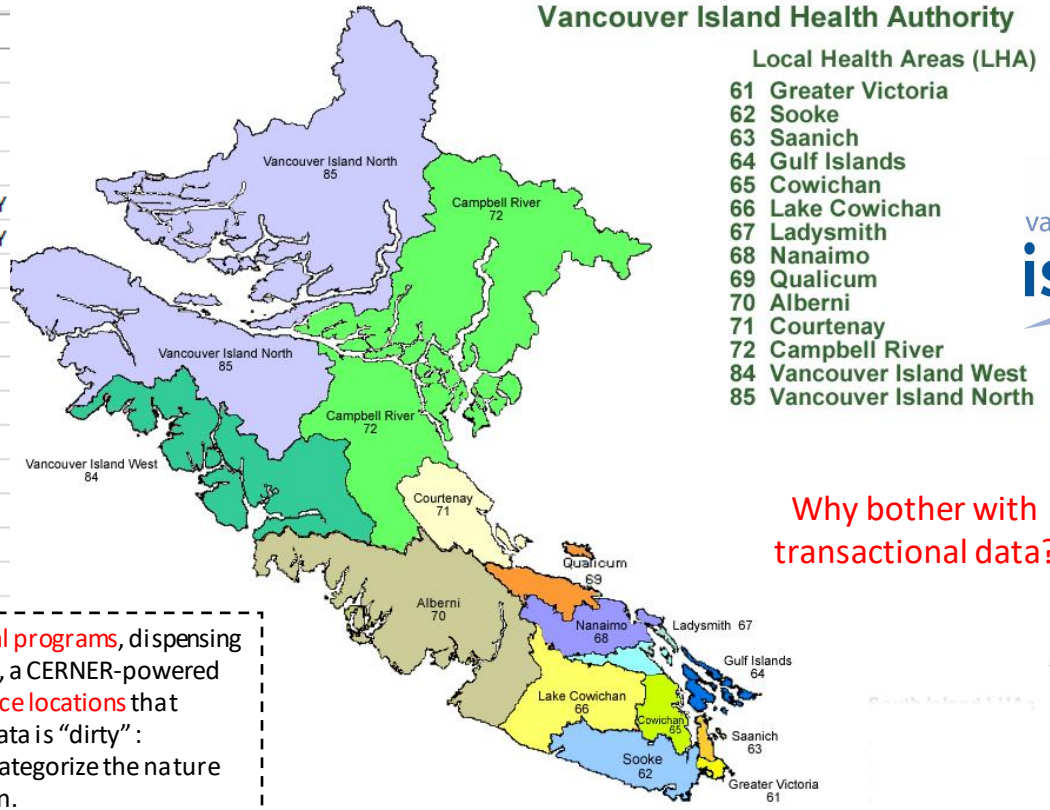
Schematic



Vancouver Island Health Authority

Examples of service locations

	Unit_Name
110	Respiratory Therapy-CDH
111	Respiratory Therapy
112	Respiratory Therapy Clinic-CVH
113	Respiratory Therapy Clinic-CRG
114	RJH - ROYAL BLOCK EXT 3 RESPIRATORY
115	RJH - ROYAL BLOCK EXT 4 RESPIRATORY
129	House 1
130	Mt. Doug Apts
131	House 2
132	A Wing-Cowichan
133	House 3
134	House 1
135	House 2
136	House 3
137	House 4
138	House 4



vancouver island health authority

VIHA is a collection of 2300+ individual programs, dispensing health services. Implemented in 2007, a CERNER-powered EHR spans the entire universe of service locations that constitute VIHA. However, location data is "dirty": frequently it's hard to make out and categorize the nature of the services offered at each location.

Why bother with transactional data?

Ken Moselle

Source: https://midwifery.ubc.ca/student-portal/midwifery-clinical-placements/about/vancouver-island-health/

2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018

Vancouver Island Health Authority

Research relying ONLY on Acute + ED data sources may be overlooking a substantial portion of patients' clinical history.

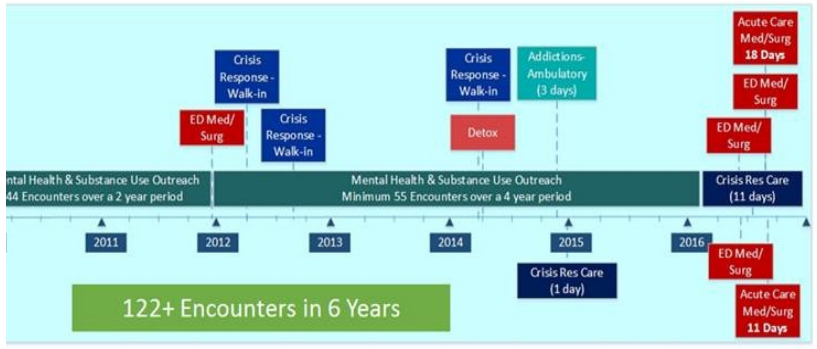
Patients' interaction with health services may offer a glimpse into a brewing crisis that otherwise would have been missed by administrative data.

I join Ken to create a learning system for mapping the full-spectrum cross-continuum space of health services at VIHA onto a smaller set of descriptive labels: Clinical Context Coding Scheme (CCCS)

Summary – Comparison of Views of Patient M (High Need/High Risk Mental Health & Substance Use)

View based on ED + Acute Care data

View based on administrative plus transactional/clinical encounter data (3Ts data)



By Ken Moselle, PhD, R.Psych.
February 21, 2017

Island Health

Here is why

Ken Moselle

2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018

github/ihacru/**ihacru-location-census**

Clinical Context Coding Scheme (CCCS) organizes the processes of (1) identifying service features of each location and (2) grouping individual locations into a class of locations based on similarity of these features.

1	Unit_Name
110	Respiratory Therapy-CDH
111	Respiratory Therapy
112	Respiratory Therapy Clinic-CVH
113	Respiratory Therapy Clinic-CRG
114	RJH - ROYAL BLOCK EXT 3 RESPIRATORY (old S3)
115	RJH - ROYAL BLOCK EXT 4 RESPIRATORY (old S3)
129	House 1
130	Mt. Doug Apts
131	House 2
132	A Wing-Cowichan
133	House 3
134	House 1
135	House 2
136	House 3
137	House 4
138	House 4
139	A Wing-Cowichan
140	A Wing
141	Sandringham Community Residential

```
ds %>% unique_sums(c("location_class_code","location_class_description")) %>% arrange(desc(n_people)) %>% neat()
```

location_class_code	location_class_description	n_encounters	n_people	n_locations
78	ED - Med-Surg	887402	141754	95
148	Medical Imaging	648411	119550	52
146	Lab - Island Health - General	833549	109088	134
57	H&CC Services	131292	79030	8
66	Acute Care - Med-Surg - Mixed Ages	127659	57117	107
34	Clinical Intake - Adult MHSU	73417	46158	21
140	Surgery - Same Day - Mixed Ages	53449	34363	24
145	Electrodiagnostics	89601	32366	10
142	Surgery - Post - Acute Care	45732	31765	35
135	Med-Surg - Ambulatory Mixed Episodic - Chronic - Mixed Ages	186901	26455	33
91	Endoscopy	38387	26006	17
138	Surgery - Prep - Recovery - Mixed Ages	34812	22639	29
37	Clerical Intake - Older Adults	37194	21907	5
43	Psychiatric [only] Clinic Services - Adults	32934	20858	32
16	Time-limited Ambulatory Treatment Services - Mental Health - Adults (secondary level)	27356	20785	29

vancouver island health authority

Ken Moselle

Health Service Location
N = 2300+

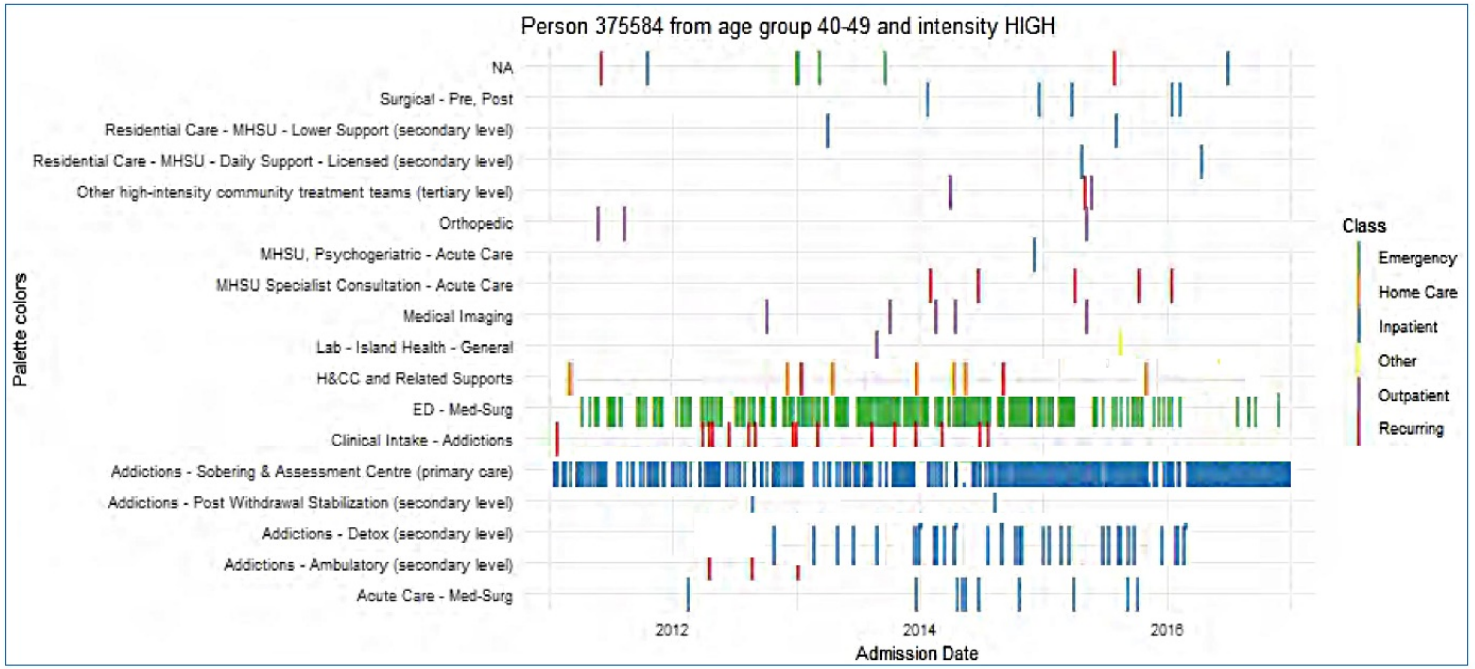
→

Location Class
N = 150+

2007 2008 2009 2010 2011 2012 2013 2014 2015 **2016** 2017 2018

Aggregating individual *service locations* into homogeneous *location classes* allowed manageable rendering of clinical landscapes: a complete history of person's interaction with health services. This image forms the focal point of my methodological interest: I want to express mathematically and evaluate statistically constellations of person's engagements with healthcare system.

Severe addiction



vancouver island health authority

Ken Moselle

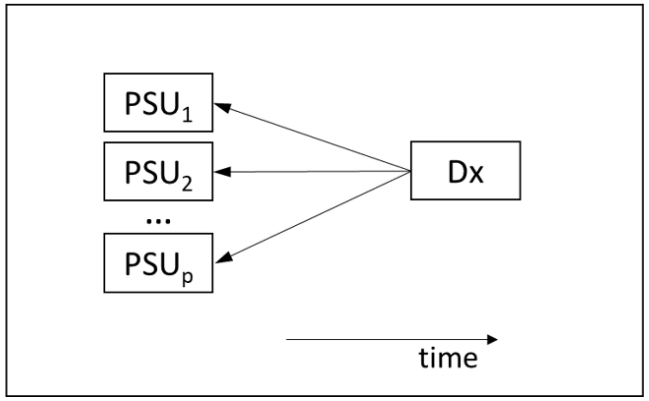
This is a fictional composite visualization based on data from several patients, cut and reassembled (Photoshop) to create an image that is representative of a single individual patient 'journey' through the array of secondary and tertiary services, but not actually reflecting at a row level the data of any patient.

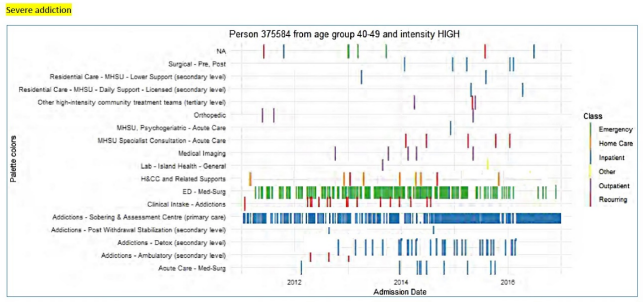
2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018

CIHR

Health System Impact Fellowship: Project 1

Stratifications of Clinical Histories





Question: Do individuals with certain diagnoses/event tend to have similar patterns of service utilization?

Premise: Transactional records of secondary and tertiary health services of Island Health are linked with substance use profile from MHSU-MRR profile, emergency room, and acute care records to assemble a data frame for estimating and training statistical models for identifying patterns of service use (PSU) related to specific health outcomes.

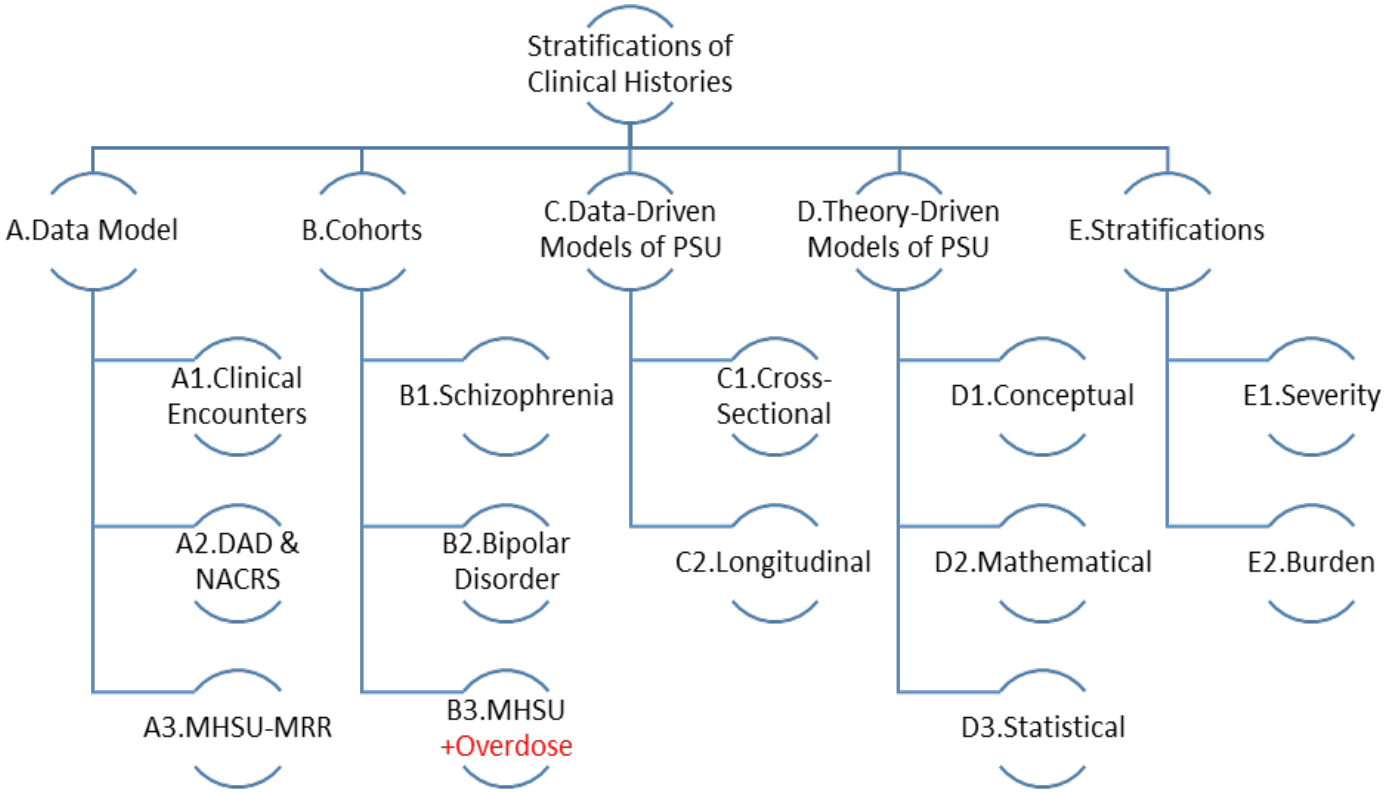
Applied Objective: Demonstrate *clinical heterogeneity of diagnostically homogeneous cohorts* by describing the variability in their clinical histories.

Methodological Question: How can we stratify patients on severity of condition and burden of disease based on their clinical history?

Building on this work, Kate, Ken, and I have designed two projects for my CIHR Health System Impact Fellowship to support BCOPPH in enhancing surveillance of chronic and/or mental health and substance use (MHSU) conditions

Kate Smolina

2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018



Work Break Down (WBD) structure of the proposed project

2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018