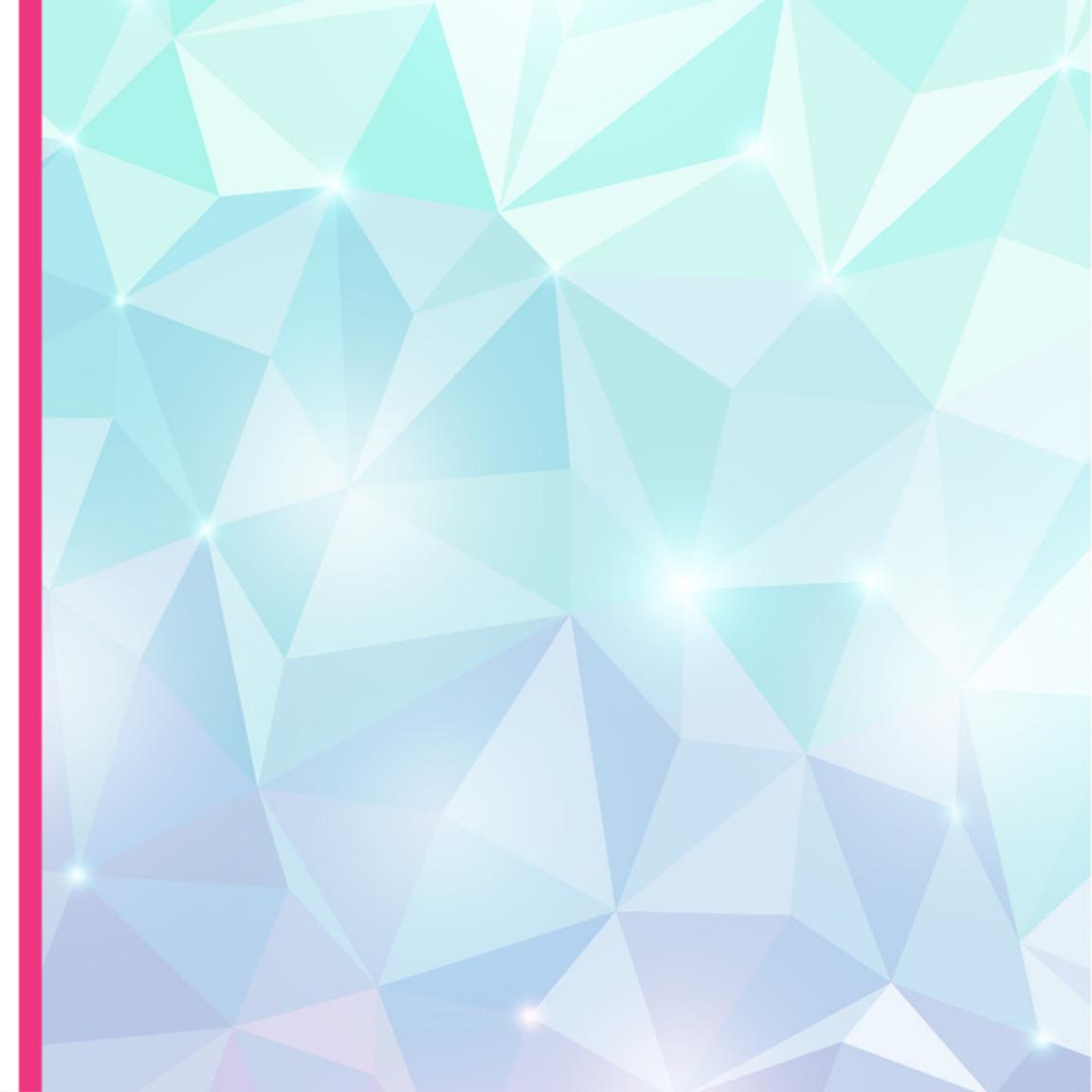


DATA INFRASTRUCTURE FOR SEPSIS RECOGNITION

Implementation & Monitoring

06/05/2019



SEPSIS WATCHER

Patient at risk of sepsis.

“...[T]he ‘**sepsis watcher protocol**’ allow[s] the clinical team more time to make a treatment decision. This means that for an hour, the child has more frequent vital sign checks and assessments. Then there’s a second sepsis huddle to reevaluate.”

- Fran Balamuth, MD, PhD

Source: <https://www.chop.edu/news/tick-tick-tick-defusing-sepsis-timebomb>

SEPSIS AVENGERS



Sharon Kaminski

33

Improvement Advisor



Paul Wildenhain

Data Analyst



Christian Minich

Data Engineer



Joe Mirizio

Data Developer

SEPSIS PROJECT OVERVIEW

STRATEGY

PHASE I

Improve
Treatment-Related
Metrics

PHASE II

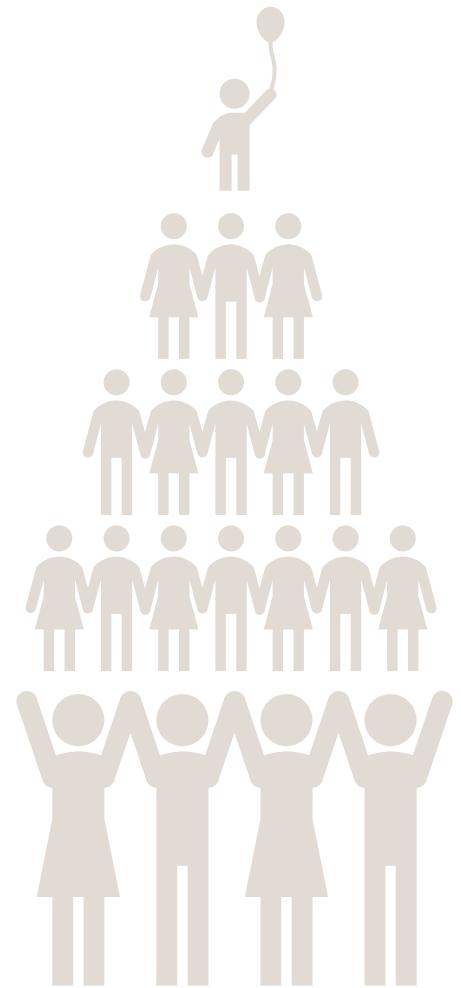
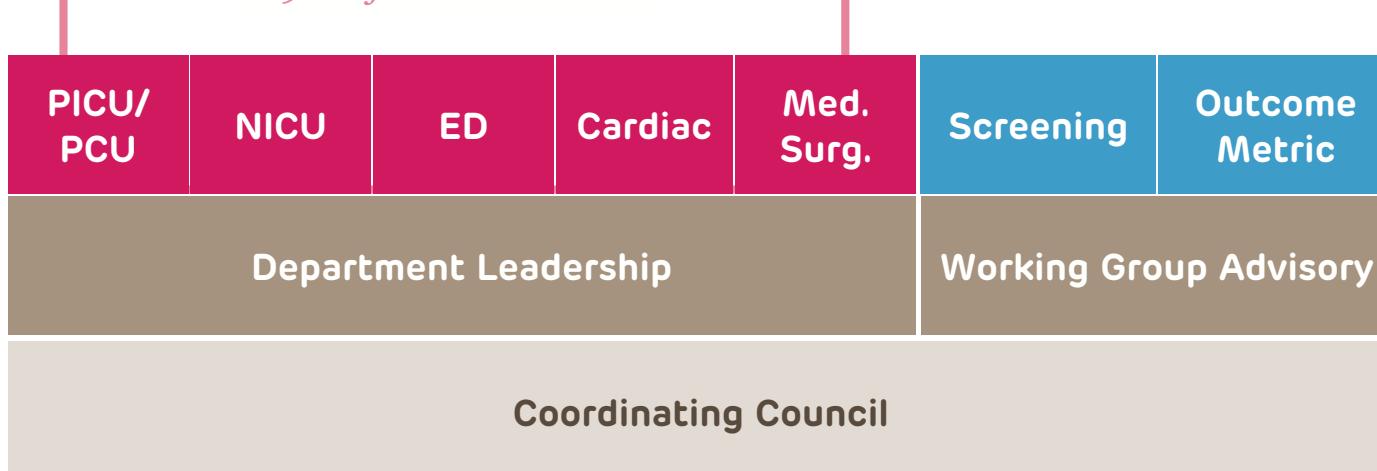
Measure
Sepsis
Outcomes

PHASE III

Improve
Early Sepsis
Recognition

ORGANIZATION

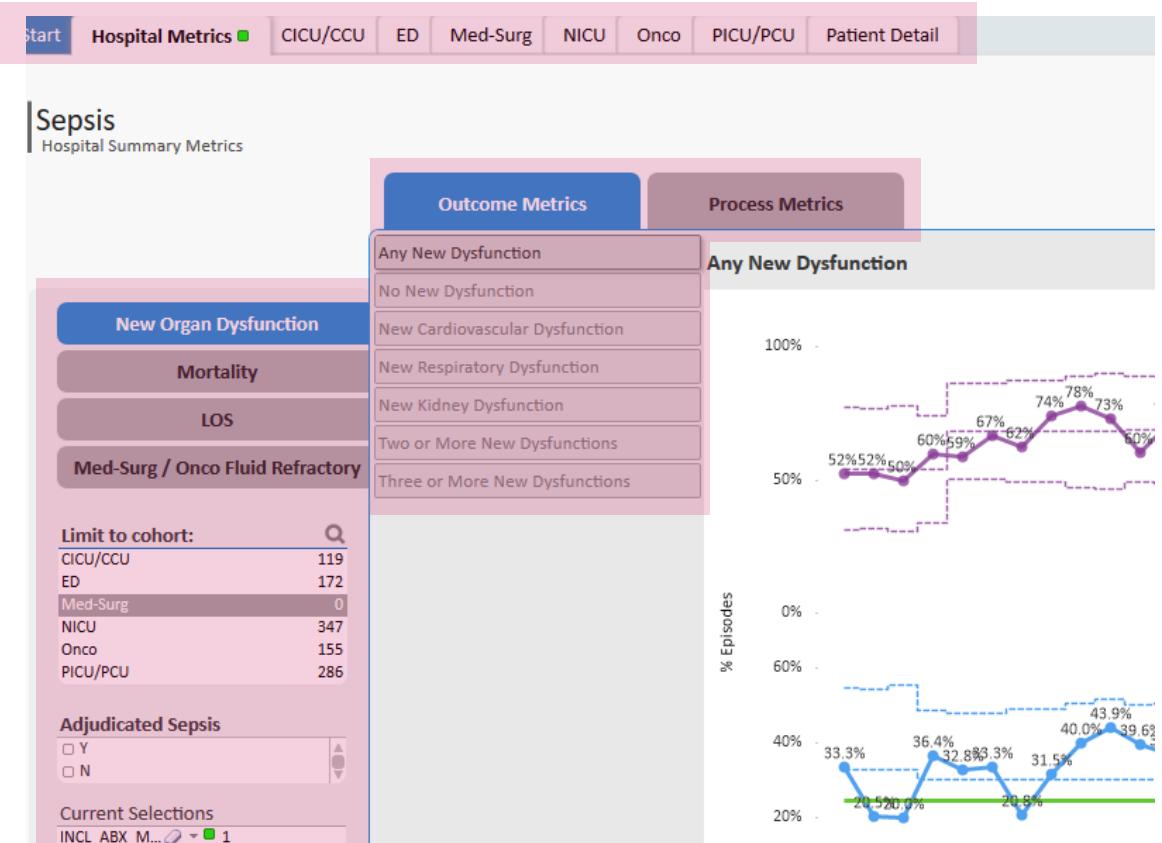
9 Projects & Metrics



DATA ECOSYSTEM • PAST

Monolithic Dashboard

- User agnostic
- Overloaded functionality
- 41 views x 4 filters



Ad Hoc Data Pulls

- Daily? Weekly? Monthly?

Rigid Data Mart

- 2 weeks for *minor* changes

5 RIGHTS OF CDS

1.  **Information** - Limited to teams needs
2.  **Intervention Format** - Push & pull
3.  **Person** - Not everyone gets email
4.  **Channel** - Email, dashboard, ad-hoc report
5.  **Time** - Pushed at appropriate frequency



Utilized this approach as a framework for data ecosystem.

Source: <https://healthit.ahrq.gov/ahrq-funded-projects/current-health-it-priorities/clinical-decision-support-cds/chapter-1-approaching-clinical-decision/section-2-overview-cds-five-rights>

DATA ECOSYSTEM • CURRENT

4 Automated Emails

- Pushed at appropriate frequency

3 Dashboard & Reports

- Ad-hoc analysis still important

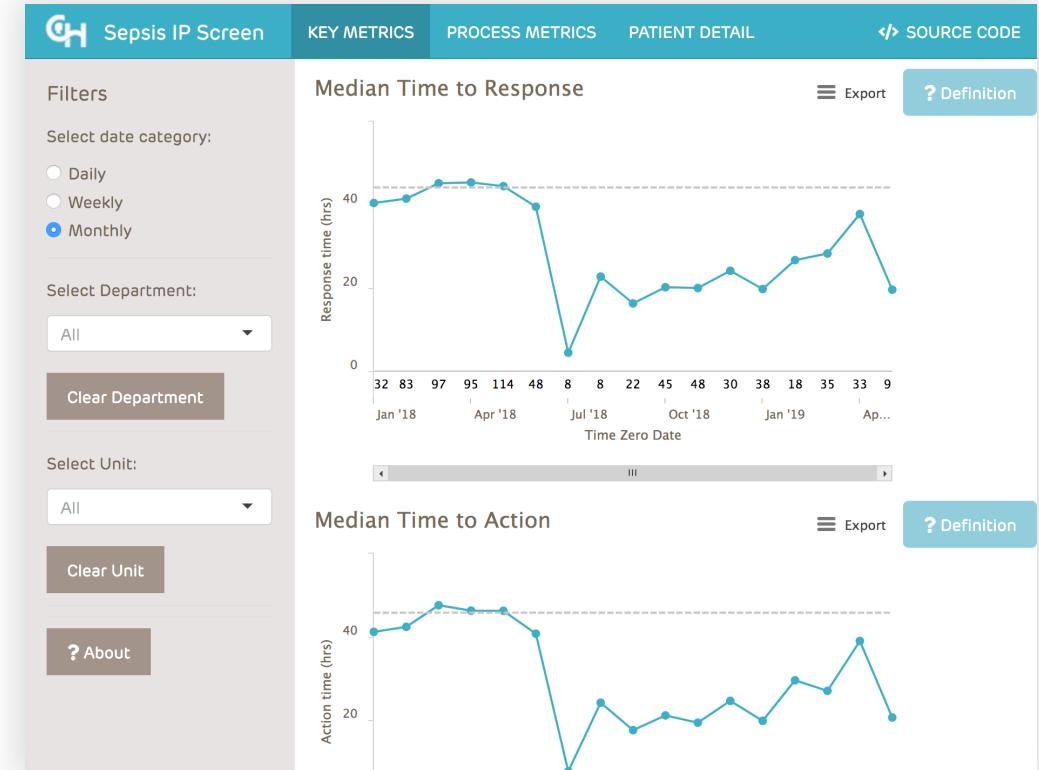
2 Data Pipelines

- REDCap → Data Warehouse

1 Sepsis Data Mart (“Automart”)

- Analyst-driven change

0 Data-Overloaded Physicians

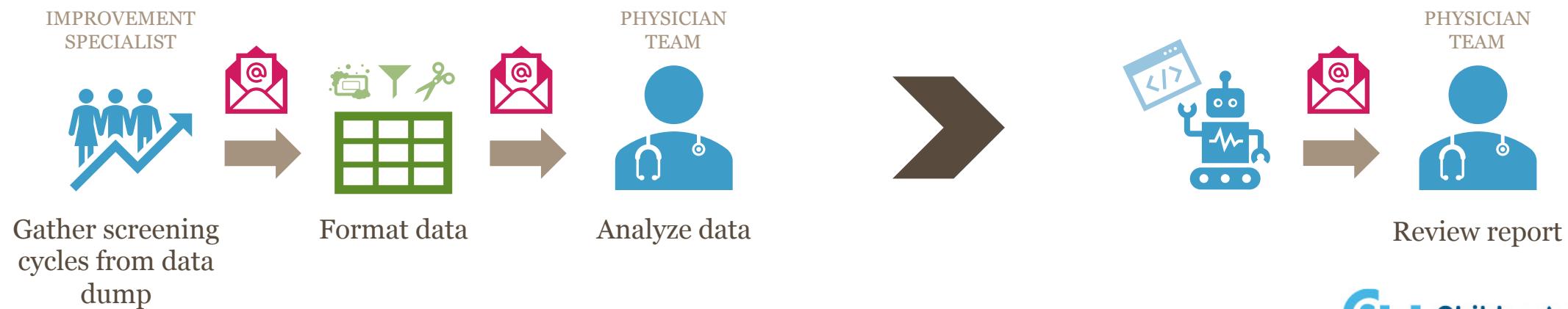


EXAMPLES



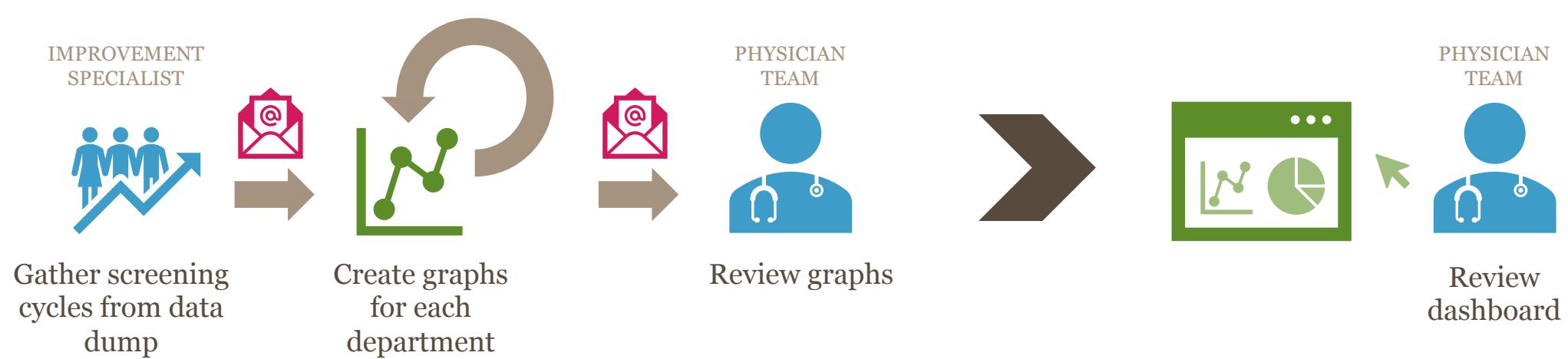
SEPSIS SCREENING REPORT

Identify adherence to sepsis screening best practice workflow for at-risk patients.



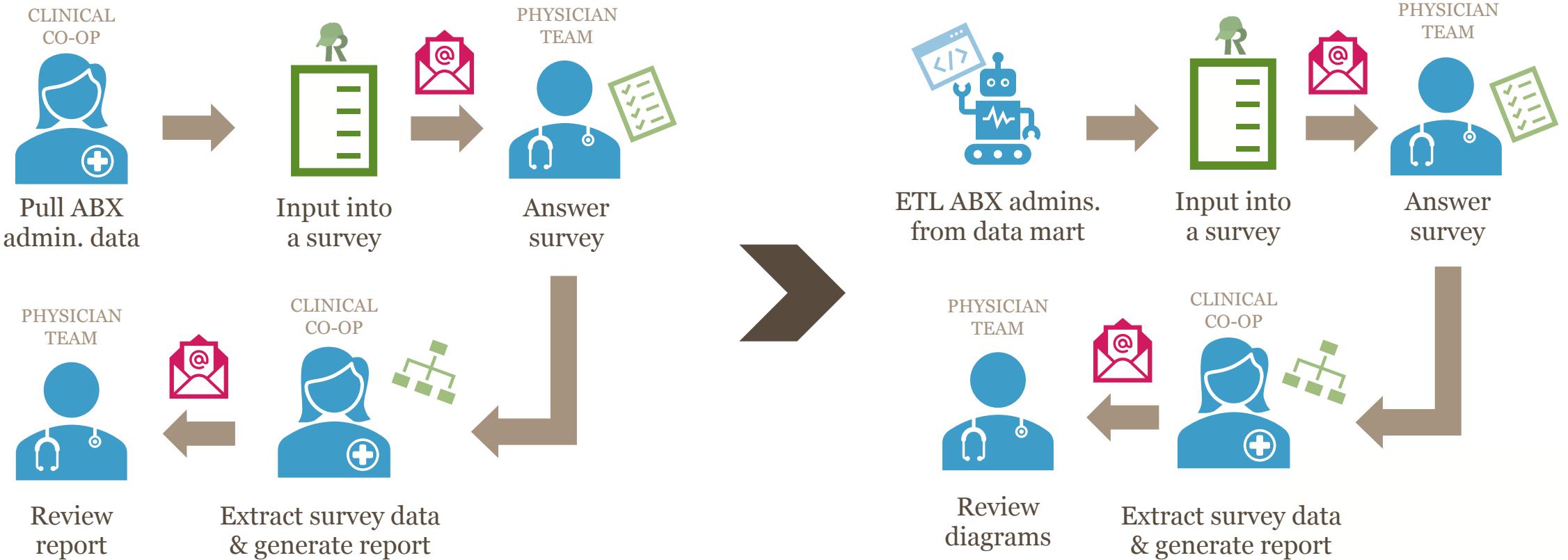
SEPSIS SCREENING METRICS

Display which departments adhere to the established guidelines around time-to-care.



LONG ANTIBIOTICS SURVEY

Explain barriers to timely sepsis care.
(i.e. Antibiotics administration)



INFRASTRUCTURE

STRATEGY

DATA PROBLEMS

- 👉 Analysts had one distribution tool
 - “Qlikview” synonymous for “published data”

Data mart changes go through IS

- Led analysts to push logic to BI tool

🤝 Project handoff

- Analyst have idiosyncratic practices
- Domain knowledge not easily transferable
- Sepsis project handed off 3 times



DATA PRINCIPLES

🔑 Analyst-owned data workflow

- Unassisted, end-to-end development

👩‍💻 Prescribed workflows for common problems

- Need to solve *[problem]* → use *[tool or process]*

🛠 Tools enforce best practices

- Version control
- Configuration-as-code
- Standardized packages

🔭 Enable discovery of analyst work

- Find each other's code & data artifacts

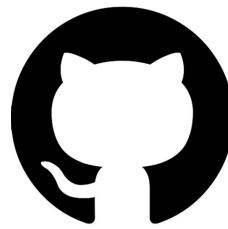


A close-up photograph of a person's face, which is intentionally blurred, creating a soft, out-of-focus effect. The person has dark hair and appears to be wearing a light-colored shirt.

INFRASTRUCTURE

TECHNOLOGY

GITHUB ENTERPRISE



Enterprise-hosted version control

🤝 Code collaboration

- Pull requests for reviews



Discoverability

- Search through team's code



APIs and extensibility

- Continuous Integration

A screenshot of a GitHub Enterprise repository page. The header shows the repository name 'CQI / FY16-Sepsis' and indicates it is private. It has 22 watchers, 1 star, and 1 fork. The main interface shows basic repository statistics: 733 commits, 7 branches, 9 releases, and 5 contributors. Below this, a list of recent commits is displayed, each with a small profile picture, the author's name, a brief description of the commit, and the time since the commit was made. The commits are as follows:

Author	Commit Description	Date
wildenhaip	Add skeleton for new NICU graph	Latest commit df70d0e 13 days ago
	Merge pull request #194 from CQI/development	a year ago
	Use procedure instead of procedure_order for CPT code	14 days ago
	update narrow chats markdown	a year ago
	catch up commit	a year ago
	Eliminating reg ex from short_med_name_better and committing updated ...	3 years ago
	catch up commit	a year ago
	Add skeleton for new NICU graph	13 days ago
	fix reference for msater_protocol_rule	2 years ago
	Add documentation for cefepime change	3 months ago
	update gitignore, add project	a year ago
	ad hoc analysis results and base datasets	3 years ago
	Update README.md	5 months ago
	update to sepsis data mart -- ED portion	3 years ago
	Updating Outcomes data for complete Q1 2016	3 years ago
	major screening + misc updates from development to master (#193)	2 years ago

JENKINS

Task automation platform



Generic task execution

- Automated emails
- Static database loads
- Continuous integration



Enforces version control



Configuration-as-code

- via Jenkinsfile

The screenshot shows the Jenkins Pipeline interface. At the top, there's a navigation bar with tabs for Pipelines, Executors, Administration, and Logout. Below the navigation is a search bar with the placeholder "scheduled". A "New Pipeline" button is located in the top right corner.

The main area displays a table of scheduled pipelines:

NAME	HEALTH	BRANCHES	PR
scheduled-tasks / Census Prediction for Daily Peak Census Validation2		-	
scheduled-tasks / Census_Prediction_Daily_Validation1		-	
scheduled-tasks / Census_Prediction_Daily_Validation2		-	
scheduled-tasks / adolescent-service-long-los		-	
scheduled-tasks / aki-email-patient-list		-	
scheduled-tasks / asp-pharma-rounds		-	
scheduled-tasks / auto-marts-monitor		-	
scheduled-tasks / bmt-immunizations-immune-info		-	
scheduled-tasks / bmt-immunizations-lab-12mo		-	
scheduled-tasks / bmt-immunizations-lab-8mo		-	
scheduled-tasks / bmt-immunizations-overdue-labs		-	
scheduled-tasks / bmt-immunizations-pat-info		-	
scheduled-tasks / bmt-immunizations-trans-info		-	
scheduled-tasks / card-arrest-email		-	
scheduled-tasks / cardiac-arrest-huddle-redcap		-	

A modal window titled "Jenkinsfile Generator" is open on the right side of the screen. It contains sections for "R Script", "Schedule", "Slack Notifications", "Network Drives", and "Vault Secrets".

The "R Script" section shows a file named "main.R" with the following content:

```
pipeline {  
    agent {  
        docker {  
            image 'ocqi/r-base'  
            args '-u root -v /mnt/DepartmentShare:/mnt/DepartmentShare'  
        }  
        triggers {  
            cron("09 * * 1")  
        }  
        stages {  
            stage('Run') {  
                steps {  
                    sh 'R ./main.R'  
                }  
            }  
        }  
        post {  
            success {  
                slackSend message: "${env.JOB_NAME} (${env.BUILD_NUMBER}) ran successfully (${env.BUILD_URL}|Open>)", color: 'good'  
            }  
            failure {  
                slackSend message: "${env.JOB_NAME} (${env.BUILD_NUMBER}) failed (${env.BUILD_URL}|Open>)", color: 'danger'  
            }  
        }  
    }  
}
```

The "Schedule" section shows a weekly trigger set to run at 09:00 AM on Monday.

The "Slack Notifications" section has an "ADD" button.

The "Network Drives" section lists drives: Q_DRIVE, ETL_ARCHIVE, QLVAAPDW1, DEPARTMENT_SHARE, and DEPARTMENT.

The "Vault Secrets" section has an "ADD" button and two entries: CDWPRD and CDWUAT.

DOCKER



Standard computing environment

✎ Maintained R images

- r
- r-base
- r-with-java

🚗 Database drivers & config

📦 Standard R packages

- e.g. tidyverse

The screenshot shows a GitHub repository page for 'CQI / r-docker'. The repository has 3 stars, 0 forks, and 0 pull requests. It contains files like Dockerfile, README.md, and r-packages.txt. A section titled 'r-base' describes it as the base R image for OCQI and lists features including unixODBC, Netezza Drivers, CDW ODBC DSNs, pando_c, R Libraries (CRAN), and various tidyverse packages.

r-base

The base R image for OCQI.

Features

In addition to providing base R functionality, this image adds the following:

- unixODBC
- Netezza Drivers
- CDW ODBC DSNs (CDWPRD , OCQI_PRD , CDWUAT , and OCQI_UAT)
- pando_c and development packages for R libraries
- R Libraries (CRAN)
 - i. tidyverse
 - ii. devtools
 - iii. rmarkdown
 - iv. odbc
 - v. REDCapR
 - vi. sendmailR
 - vii. EasyHTMLReport
 - viii. htmlTable
 - ix. readxl
 - x. writexl

RSTUDIO CONNECT



Publishing platform for R

- 👉 “Push button” publishing
 - via RStudio (IDE)

The screenshot shows the RStudio Connect interface. On the left, an RStudio IDE window displays an R Markdown file named 'impact_fy18.report.Rmd'. A modal dialog titled 'Publish to Server' is open, showing the file 'impact_fy18.report.Rmd' selected. On the right, the published content is displayed. It includes a section titled 'Who can view this document' set to 'Anyone - no login required', and another section for 'Who can change this document'. Below this, there is a 'Scheduled Admissions' dashboard. The dashboard has a header 'SCHEDULED ADMISSIONS BY BLOCK DAY'. It features a chart titled 'Scheduled Med/Surg Bed Admissions for May 1: 0' with a legend indicating green for Mon-Fri and orange for Saturday. Below the chart is a 'Select Weekday:' dropdown set to 'All'. To the right are two more charts for 'Med/Surg Bed June 2019' and 'Med/Surg Bed July 2019', each with a 5x5 grid of numbers representing daily admissions.

- 📦 Package management
 - via packrat

- 👮 User-authenticated content
 - Works with LDAP

AUTOMARTS



Automated data mart creation



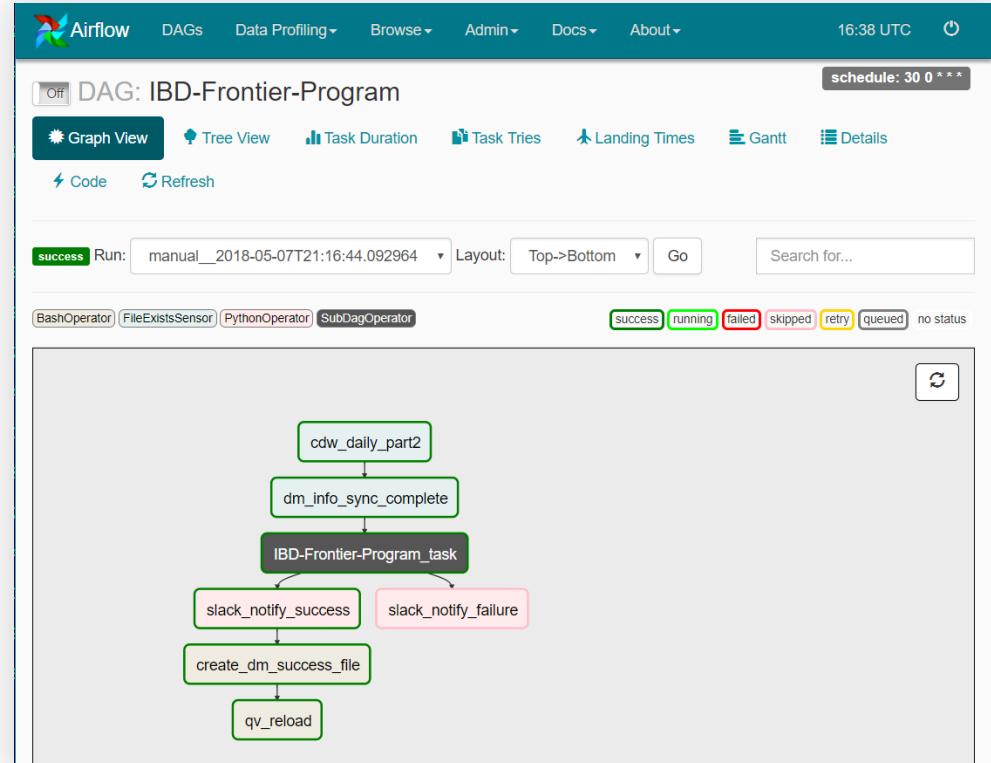
- Airflow jobs create data marts
 - DAGs scheduled daily



- Reads SQL from GitHub
 - Auto-discovered dependencies



- Feedback via Slack



DATA SOLUTIONS

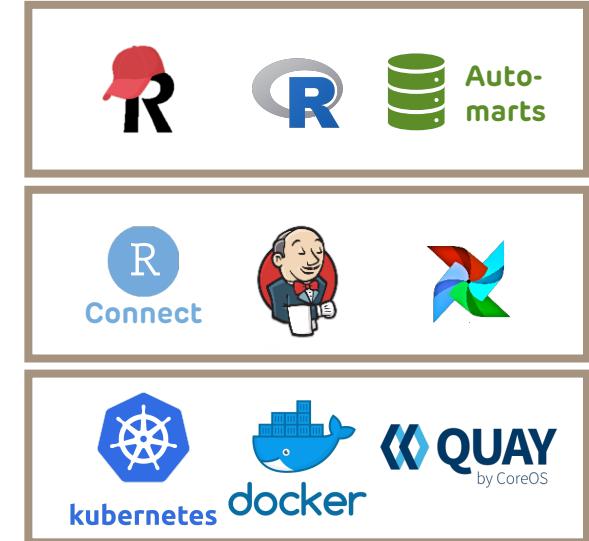
TECHNOLOGY RECIPES

📊 **Dashboard** =  +  Auto-marts

✉️ **Email** =  +  Auto-marts
=  +  Auto-marts + 

🏗 **Data Pipeline** =  +  Auto-marts +  + 

TECHNOLOGY STACK



NEXT STEPS



Create a catalog of data artifacts

- Diaspora of project repositories & different artifact types are difficult to search through



Build standard artifacts via automation

- Eliminate need to create routine data products (e.g. patient lists & metric SPC charts)



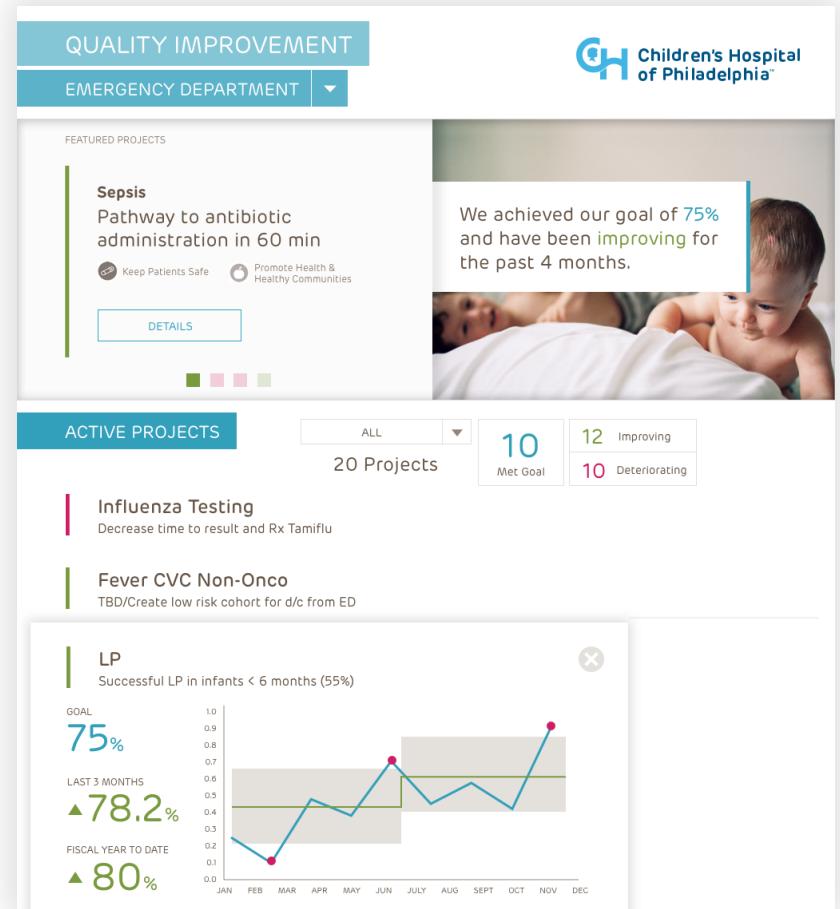
Develop operationalization strategy

- “You build it, you maintain it” – not sustainable



Continue productionizing infrastructure

- Harden & scale as needs grow



SO YOU WANT TO DO THIS...

Established centralized IS resources

- Servers to host applications
- SaaS & CaaS - *Quay, Kubernetes, Vault, etc.*

Investment in personnel required

- Cross-functional team:
Analysts, engineers, developers, & users
- 2-3+ data engineers & developers
- Full-stack analysts – *total workflow ownership*



SUMMARY

🚧 Sepsis data support was completely overhauled

- Saves clinical team 20+ hours every week

5 Cross-functional teams enable 5 Rights of CDS

- Developers design & build infrastructure →
- Analysts rapidly deploy reliable data artifacts →
- Improvement specialists & clinicians save time

💰 Large investment, but we'd do it again!

