

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

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

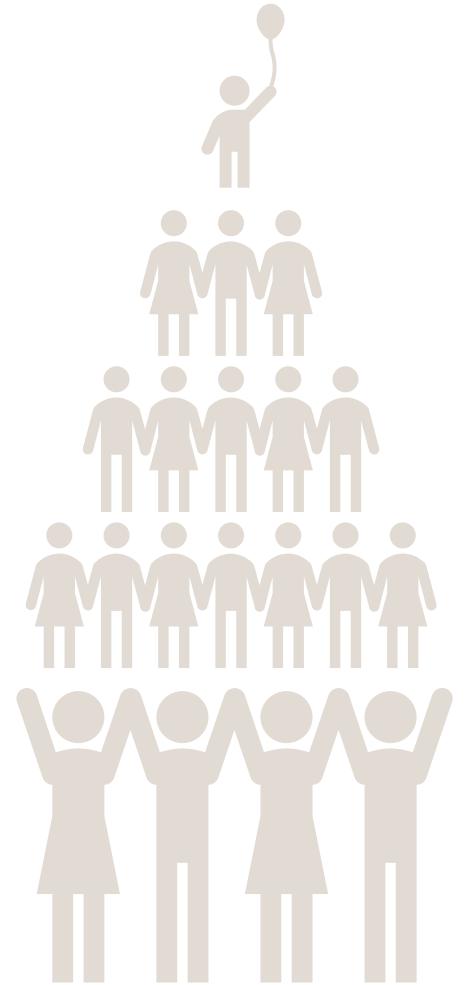
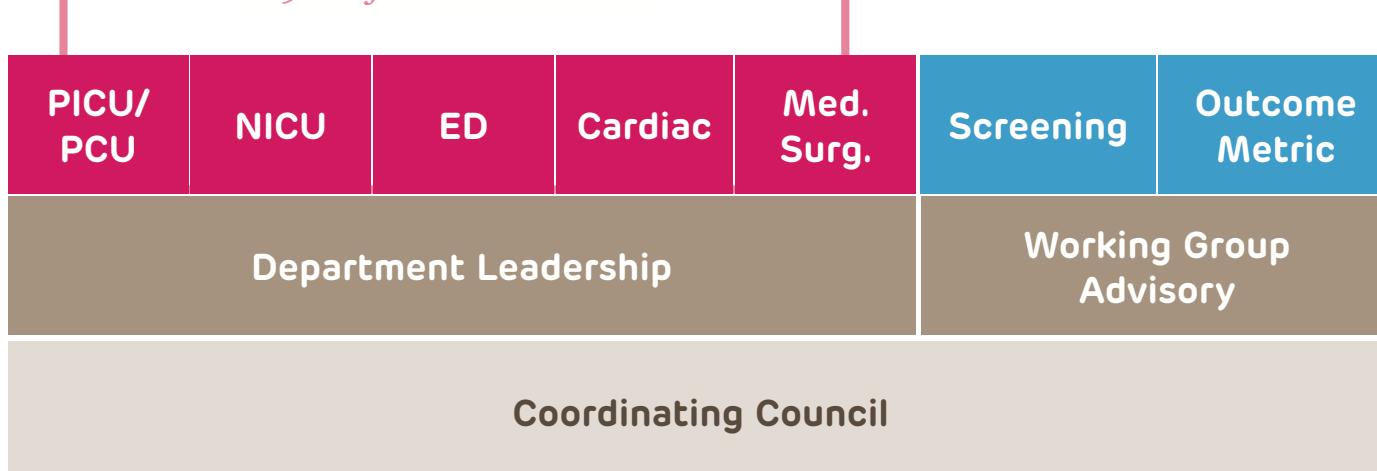
Measure
Sepsis
Outcomes

PHASE III

Improve
Early Sepsis
Recognition

ORGANIZATION

9 Projects & Metrics



DATA ECOSYSTEM • PAST

Monolithic Dashboard

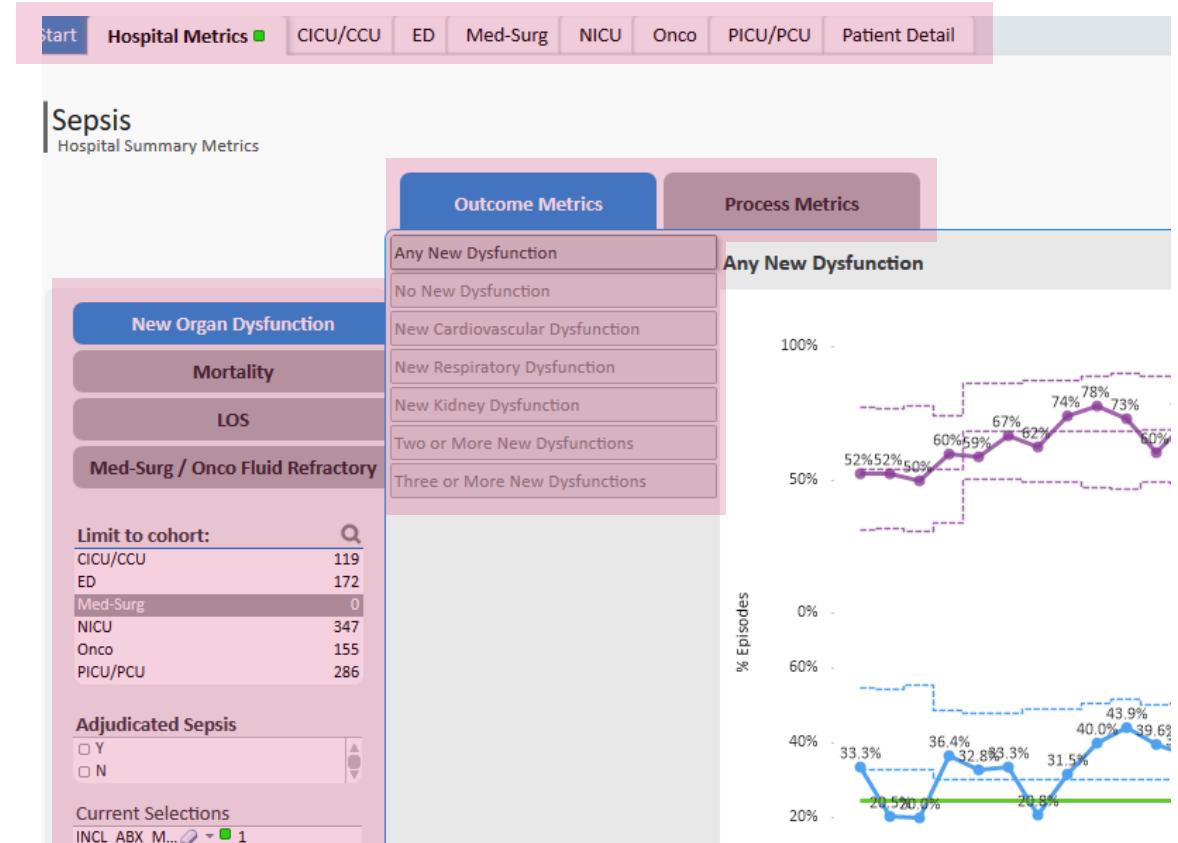
- User agnostic
- Overloaded functionality
- 41 views x 4 filters

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

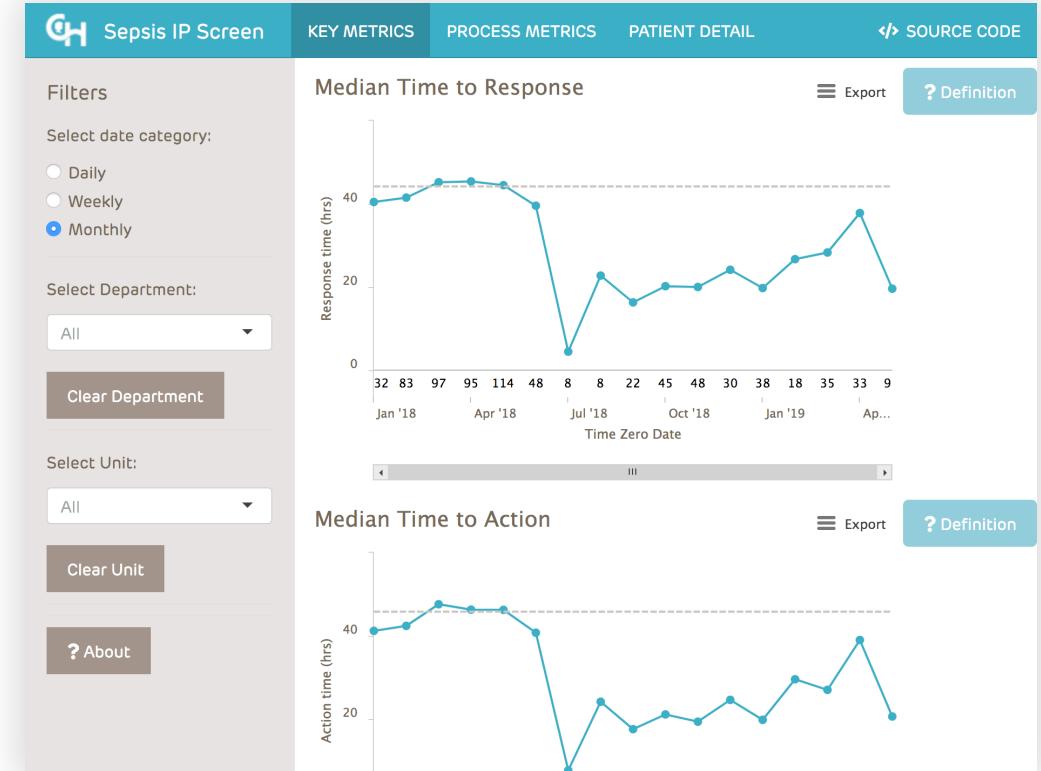
2 Data Pipelines

- REDCap → Data Warehouse

1 Sepsis Data Mart (“Automart”)

- Analyst-driven change

0 Data-Deprived Physicians

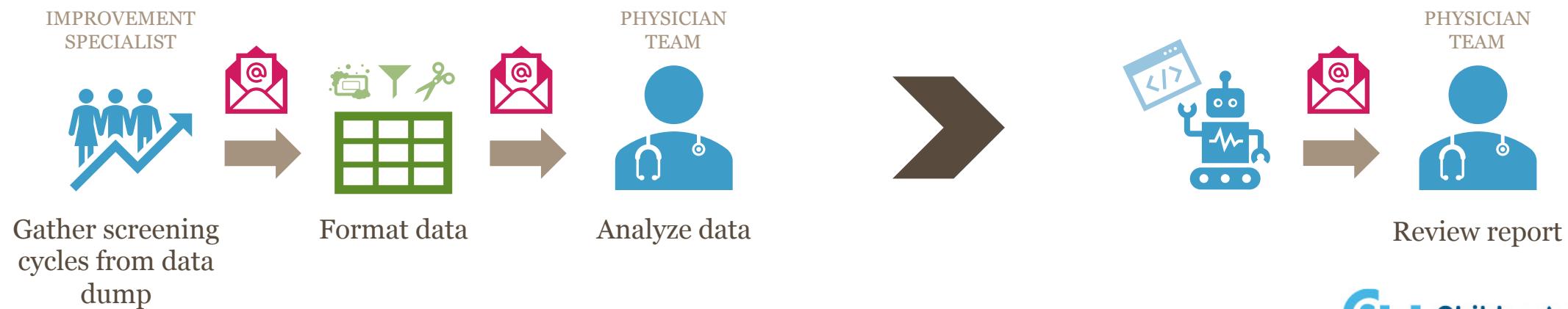


EXAMPLES



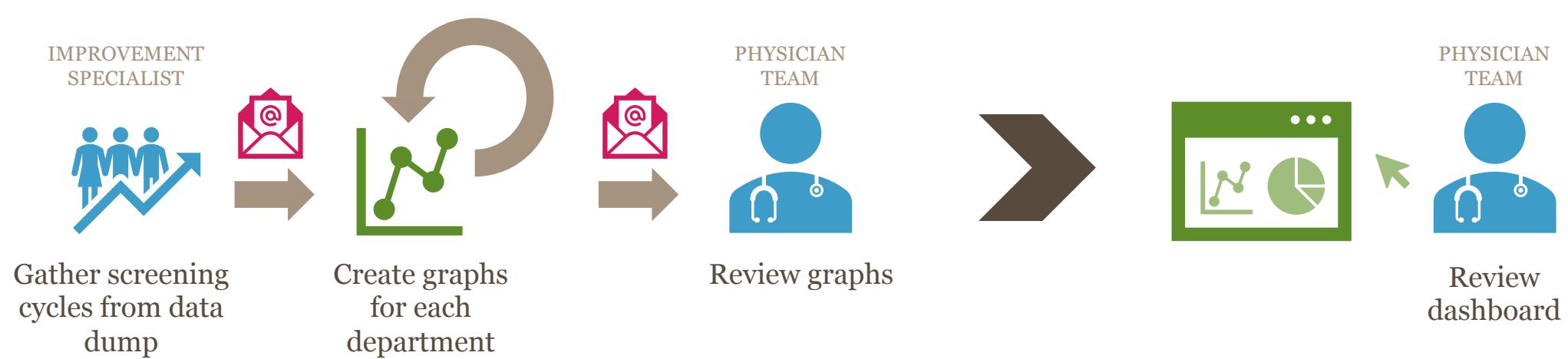
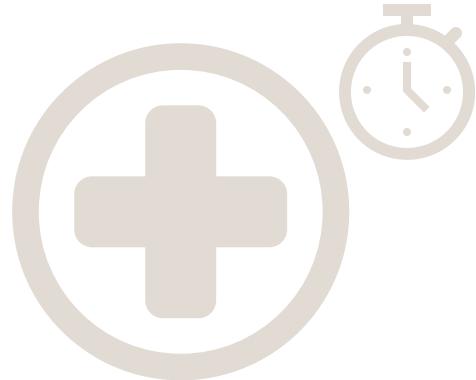
SEPSIS SCREENING REPORT

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



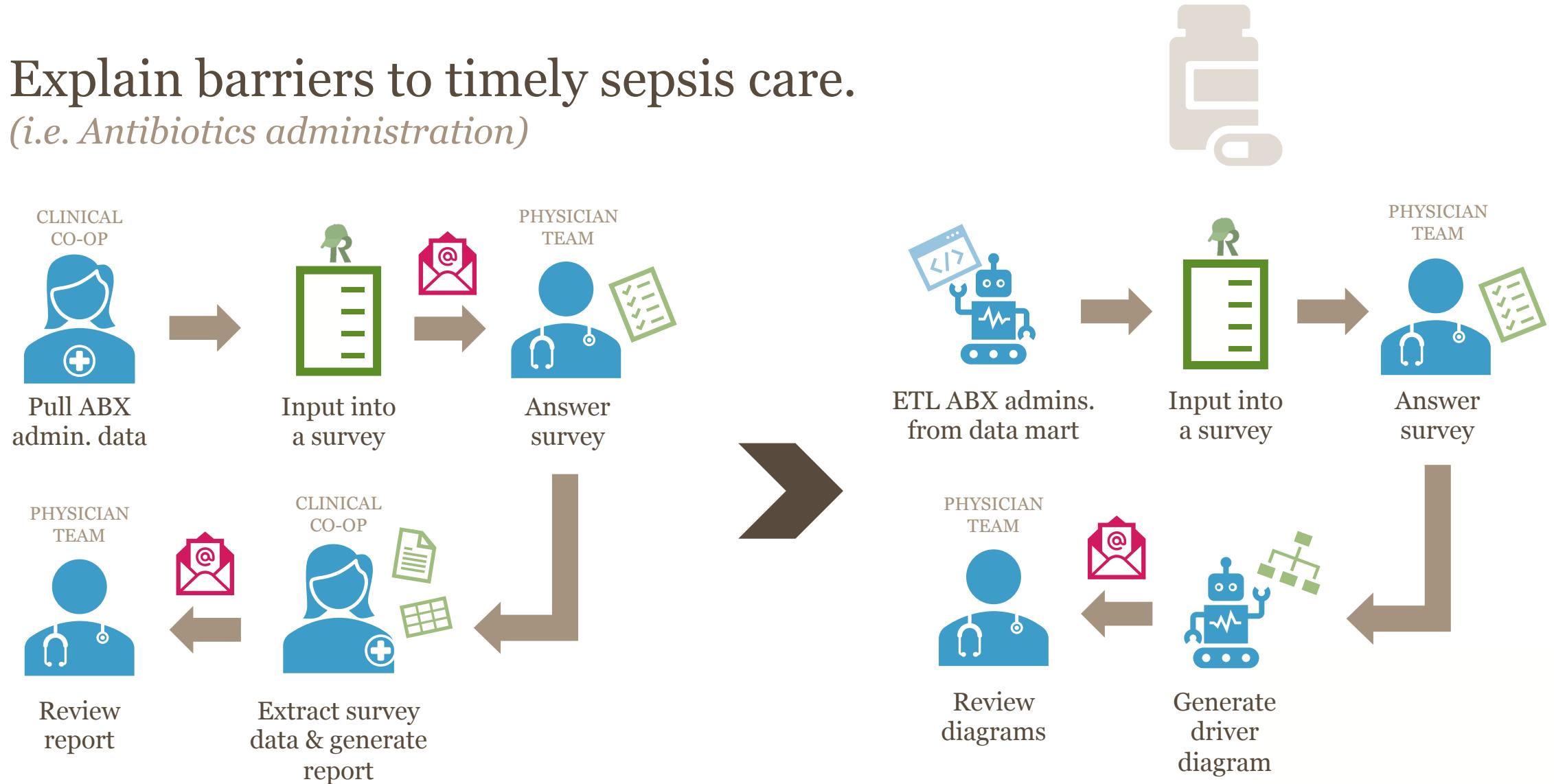
SEPSIS SCREENING DASHBOARD

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

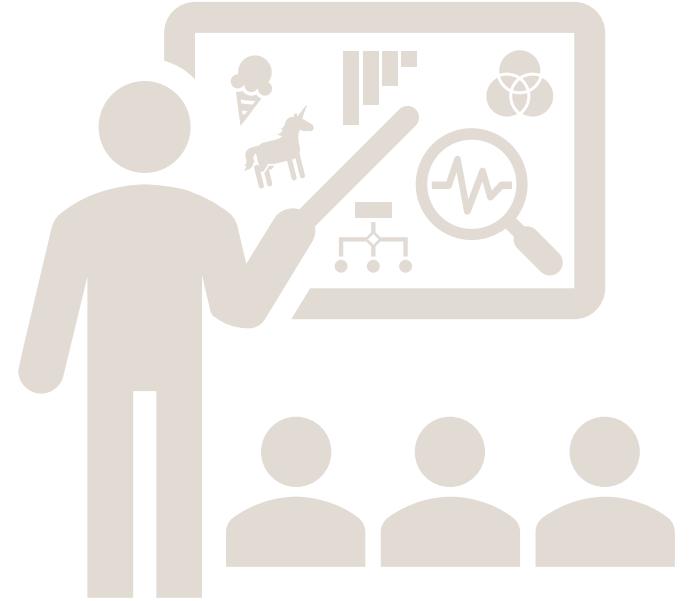
- Need to solve *[problem]* → use *[tool or process]*

🛠 Tools enforce best practices

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

🔭 Enable discovery of analyst work

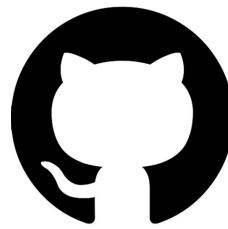
- Find each other's code & data artifacts



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 repository is named 'CQI / FY16-Sepsis' and is marked as private. The page shows 733 commits, 7 branches, 9 releases, and 5 contributors. The commit history lists various changes made by users like 'wildenhaip' and 'Code'. The commits are dated from 13 days ago to 2 years ago.

Commit	Description	Date
wildenhaip Add skeleton for new NICU graph	Latest commit df70d0e 13 days ago	
Ad Hoc Analyses	Merge pull request #194 from CQI/development	a year ago
Code	Use procedure instead of procedure_order for CPT code	14 days ago
Outcomes	update narrow chats markdown	a year ago
PICU_STAT	catch up commit	a year ago
Recognition Pediatrics Replication...	Eliminating reg ex from short_med_name_better and committing updated ...	3 years ago
Sepsis Reconition	catch up commit	a year ago
Sepsis Response	Add skeleton for new NICU graph	13 days ago
Sepsis data mart	fix reference for msater_protocol_rule	2 years ago
.gitignore	Add documentation for cefepime change	3 months ago
FY16-Sepsis.Rproj	update gitignore, add project	a year ago
PICU First Admin Abx Within 60 m...	ad hoc analysis results and base datasets	3 years ago
README.md	Update README.md	5 months ago
cohort differences.xlsx	update to sepsis data mart -- ED portion	3 years ago
sepsis abx data.xlsx	Updating Outcomes data for complete Q1 2016	3 years ago
update_datamart.bat	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 text "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	-	-	

To the right of the pipeline list, there's a "Jenkinsfile Generator" panel for the "OCQI Jenkins R Jobs" pipeline. It includes sections for "R Script" (containing "main.R"), "Schedule" (set to "Every week around 09:00 AM on Monday"), "Slack Notifications" (ADD), "Network Drives" (Q_DRIVE, ETL_ARCHIVE, QLVAAPDW1), "DEPARTMENT_SHARE" (selected), "DEPARTMENT" (DEPARTMENT), and "Vault Secrets" (ADD, CDWPRD, CDWUAT). A "COPY TO CLIPBOARD" button is also present in the generator panel.

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 issues. It contains a single branch named 'master'. The commits listed are:

- mirizioj Fix bug with .Rprofile (4 days ago)
- Dockerfile Fix bug with .Rprofile (4 days ago)
- README.md Add openxlsx (a month ago)
- r-packages.txt Add external file for r dependencies (4 days ago)

The 'r-base' section describes it as "The base R image for OCQI". The 'Features' section lists additional functionality:

- unixODBC
- Netezza Drivers
- CDW ODBC DSNs (CDWPRD , OCQI_PRD , CDWUAT , and OCQI_UAT)
- pandoc 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. At the top, there's a preview of an R Markdown file titled "impact_fy18_report.Rmd". Below it, a "Publish to Server" dialog is open, showing the file has been published to the account "minichc:rstudio-connect". The main area displays a dashboard titled "CCU Discharge QI Metrics" with two line charts: "AVERAGE CCU LENGTH OF STAY (HOURS)" and "MEDIAN CCU LENGTH OF STAY (HOURS)". The charts show data from December 2017 to June 2018, with various metrics and control limits labeled. On the right side, there are sections for managing document access, changing, and running.

- 📦 Package management
 - via packrat

- 👮 User-authenticated content
 - Works with LDAP

AUTOMARTS



Automated data mart creation



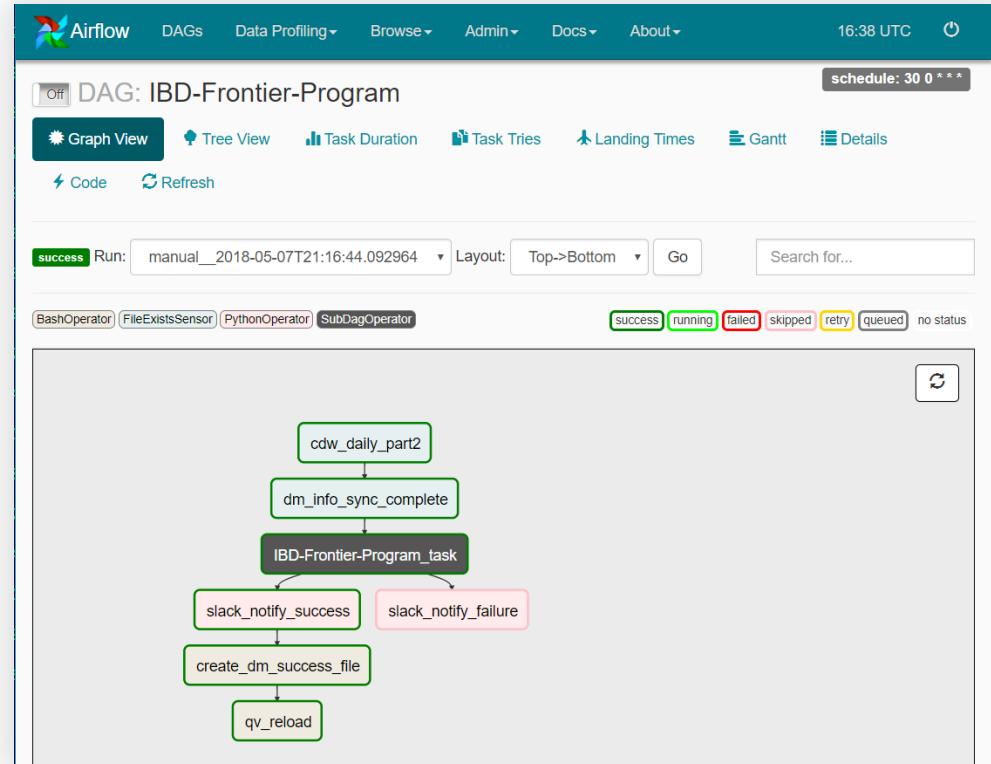
- Airflow jobs create data marts
 - DAGs scheduled daily



- Reads SQL from GitHub
 - Auto-discovered dependencies



- Feedback via Slack



DATA SOLUTIONS

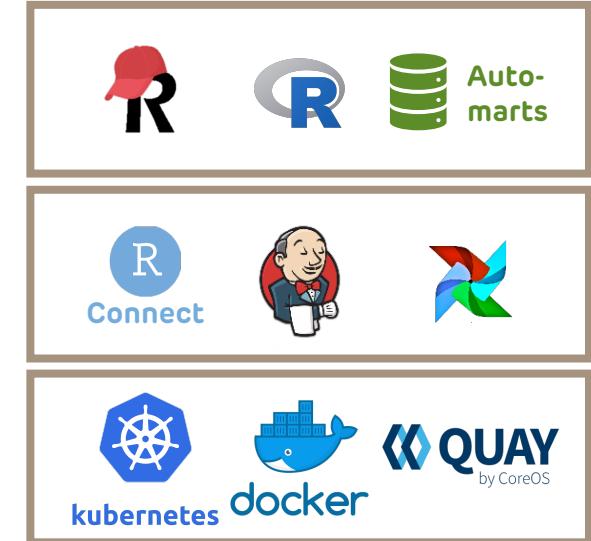
TECHNOLOGY RECIPES

📊 **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 went from rags to riches

- Saves clinical team 10+ 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!