A Semi-Automated Method to Analyze and Report Data from *Neisseria gonorrheae* Antimicrobial Susceptibility Testing

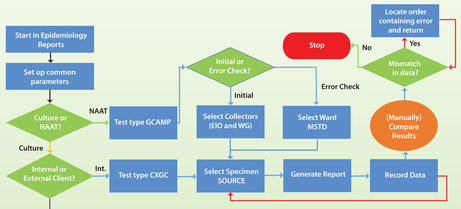
J. Weiner, M. Khubbar, and S. Bhattacharyya, City of Milwaukee Health Department Laboratory, Milwaukee, WI

Introduction:

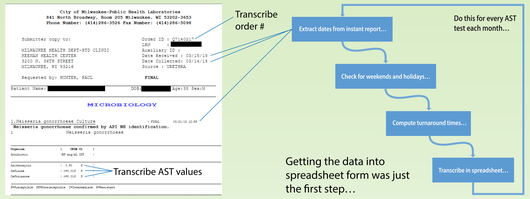
The City of Milwaukee Health Department Laboratory (MHDL) uses the SoftLab/SoftMic platform from Soft Computer (SCC)[1] for Laboratory Information System (LIS). The platform includes a front-end Graphical User Interface (GUI) for daily lab operations and a back-end database (DB) that can be queried via Microsoft Access[2]. We designed a system of two Access queries coupled to three programs in the Python[3] language that accomplish tasks of compiling, analyzing, and validating laboratory results for NAATs, cultures and susceptibility testing for *N. gonorrheae.*

MHDL is one of the participants in the CDC-funded Strengthening U.S. Response to Resistant Gonorrhea (SURRG)[4] grant study and follows specific reporting guidelines for *N. gonorrheae* culture and susceptibility testing. The processes described herein are generalizable to any disease surveillance given suitable adaptations.

Before process improvements implemented: LIS-only method – Convoluted

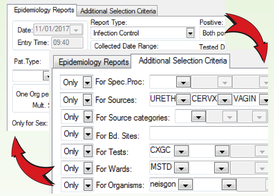


Manual accession of turnaround time and AST values – Error prone



Goals for data reporting process improvement

Previous method – labor intensive



- Requires a separate report for each test code (3), each specimen type (9), each clinic (3, soon 4), and each result type (2) to compile data
- 3*9*3*2 = 162 LIS reports to generate needed data each month

New method – streamlined

Drawbacks of manual data methods

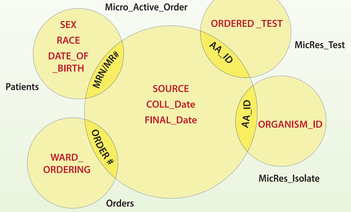
- Accessing data source by source, test by test, and clinic by clinic is highly labor-intensive
- Manual calculations are error-prone
- Despite adaptations, not all data are readily viewable from the LIS front-end so reports may be incomplete

Improving data management through automation 1: Standardizing data mining with SQL queries

- The previous method of data mining within the LIS could lead to omission errors if procedures aren't followed precisely
- Using standardized SQL queries (below) prevents this

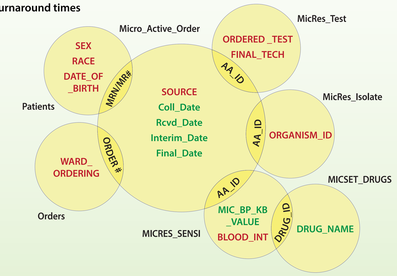
Query design: Total tests

- Venn diagram representation of SURRG table 1 query
- Each circle represents a table in DB
- Overlaps are linkages (keys) between tables
- Data elements in query are shown in red



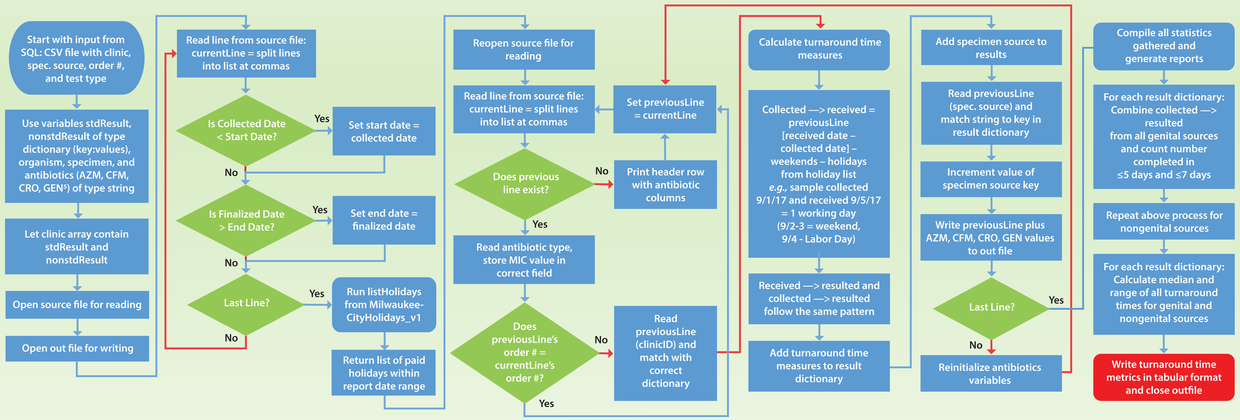
Query design: Etest results with turnaround times

- Venn diagram representation of SURRG table 2-3 query
- Each circle represents a table in DB
- Overlaps are linkages (keys) between tables
- Data elements distinct to Etest/TAT metrics are shown in green



Improving data management through automation 2: Algorithms for data analysis

Sample algorithm: *N. gonorrhoeae* Etest AST results



- The previous method allowed for calculation and entry errors due to manual data entry and analysis
- Data is now analyzed by a Python program which was coded with accurate formulae and has been validated against years of data
- Fast- Programs run in O(n) (linear) time
- Flowchart is somewhat simplified in this presentation

Data viewed Pre- and post-processing in Python: basic

- **SQL output**
  - Includes unwanted duplicate line and irrelevant test result

ORDER.	MR.	DATE_OF_ BIRTH	SEX	SOURCE	COLL_DATE	FINAL_DATE	ORDERED_ TEST	ORGAN- ISM_ID	WARD_ORDERING
P8050054	115611	10/9/86 0:00	M	URETH	6/5/17	6/8/17	CXGC	neismen	MSTD
P8050055	266204	8/6/96 0:00	M	URINE	6/5/17	6/6/17	GCAMP		MSTD
P8050056	266204	8/6/96 0:00	M	THR	6/5/17	6/6/17	GCAMP		MSTD
P8050057	266204	8/6/96 0:00	M	RECTM	6/5/17	6/6/17	GCAMP	CTRACH	MSTD
P8050057	266204	8/6/96 0:00	M	RECTM	6/5/17	6/6/17	GCAMP	neisgon	MSTD
P8050058	266204	8/6/96 0:00	M	RECTM	6/5/17	6/14/17	CXGC	neisgon	MSTD
P8050059	266204	8/6/96 0:00	M	THR	6/5/17	6/7/17	CXGC		MSTD

- **Python output**
  - Removes duplicate line with irrelevant result

ORDER.	MR.	DATE_OF_ BIRTH	SEX	SOURCE	COLL_DATE	FINAL_DATE	ORDERED_ TEST	ORGAN- ISM_ID	WARD_ORDERING
P8050054	115611	10/9/86 0:00	M	URETH	6/5/17	6/8/17	CXGC	neismen	MSTD
P8050055	266204	8/6/96 0:00	M	URINE	6/5/17	6/6/17	GCAMP		MSTD
P8050056	266204	8/6/96 0:00	M	THR	6/5/17	6/6/17	GCAMP		MSTD
P8050057	266204	8/6/96 0:00	M	RECTM	6/5/17	6/6/17	GCAMP	neisgon	MSTD
P8050057	266204	8/6/96 0:00	M	RECTM	6/5/17	6/14/17	CXGC	neisgon	MSTD
P8050058	266204	8/6/96 0:00	M	RECTM	6/5/17	6/14/17	CXGC	neisgon	MSTD
P8050059	266204	8/6/96 0:00	M	THR	6/5/17	6/7/17	CXGC		MSTD

Data viewed Pre- and post-processing in Python: advanced

- **SQL output**
  - Outputs four lines per test, one for each Etest MIC value

	SURRG_SpecimenID	PATIENTID	Site	Specimen_ Source	DOB	Sex	Collected Date	Received Date	Confirmed Date	Final Result	Scientist	DRUG_NAME	Interpret	MIC_BP_KB_VALUE
Test 1	MILQ7010048	272290	MSTD	THR	5/3/00 0:00	M	3/1/18 0:00	3/2/18 0:00	3/6/18 0:00	3/7/18 0:00	RG	Azithromycin	S	0.25
Test 1	MILQ7010048	272290	MSTD	THR	5/3/00 0:00	M	3/1/18 0:00	3/2/18 0:00	3/6/18 0:00	3/7/18 0:00	RG	Cefixime	S	<=0.016
Test 1	MILQ7010048	272290	MSTD	THR	5/3/00 0:00	M	3/1/18 0:00	3/2/18 0:00	3/6/18 0:00	3/7/18 0:00	RG	Ceftriaxone	S	<=0.016
Test 1	MILQ7010048	272290	MSTD	THR	5/3/00 0:00	M	3/1/18 0:00	3/2/18 0:00	3/6/18 0:00	3/7/18 0:00	RG	Gentamicin		8
Test 2	MILQ7010077	276218	PPWA	CERVX	6/6/01 0:00	F	3/1/18 0:00	3/1/18 0:00	3/6/18 0:00	3/7/18 0:00	RG	Azithromycin	S	0.125
Test 2	MILQ7010077	276218	PPWA	CERVX	6/6/01 0:00	F	3/1/18 0:00	3/1/18 0:00	3/6/18 0:00	3/7/18 0:00	RG	Cefixime	S	<=0.016
Test 2	MILQ7010077	276218	PPWA	CERVX	6/6/01 0:00	F	3/1/18 0:00	3/1/18 0:00	3/6/18 0:00	3/7/18 0:00	RG	Ceftriaxone	S	<=0.016
Test 2	MILQ7010077	276218	PPWA	CERVX	6/6/01 0:00	F	3/1/18 0:00	3/1/18 0:00	3/6/18 0:00	3/7/18 0:00	RG	Gentamicin		4

- **Python output**
  - Condenses each test to a single line with a column for each antibiotic MIC
  - Calculates turnaround times

SURRG_ SPEC_ID	PATIENTID	Site	SPEC_ SOURCE	DOB	Sex	Collected	Received	Confirmed	Final Result	COLL_REC	REC_RES	COLL_RES	Scientist	Azithromycin	Cefixime	Ceftriaxone	Gentamicin
MILQ7010048	0000272290	MSTD	THR	5/3/00	M	3/1/18	3/2/18	3/6/18	3/7/18	1	3	4	RG	0.25	<=0.016	<=0.016	8
MILQ7010077	0000276218	PPWA	CERVX	6/6/01	F	3/1/18	3/1/18	3/6/18	3/7/18	0	4	4	RG	0.125	<=0.016	<=0.016	4

How do we generate reports?

- **Most fields simply flow through to output, but several are captured:**

ORDER.	MR.	DATE_OF_ BIRTH	SEX	SOURCE	COLL_DATE	FINAL_DATE	ORDERED_ TEST	ORGAN- ISM_ID	WARD_ORDERING
P8050054	115611	10/9/86 0:00	M	URETH	6/5/17	6/8/17	CXGC	neismen	MSTD
P8050055	266204	8/6/96 0:00	M	URINE	6/5/17	6/6/17	GCAMP		MSTD
P8050056	266204	8/6/96 0:00	M	THR	6/5/17	6/6/17	GCAMP		MSTD
P8050057	266204	8/6/96 0:00	M	RECTM	6/5/17	6/6/17	GCAMP		MSTD
P8050057	266204	8/6/96 0:00	M	RECTM	6/5/17	6/8/17	GCAMP	neisgon	MSTD
P8050058	266204	8/6/96 0:00	M	RECTM	6/5/17	6/14/17	CXGC	neisgon	MSTD
P8050059	266204	8/6/96 0:00	M	THR	6/5/17	6/7/17	CXGC		MSTD

- **Order#** used to separate tests
- **Ordered test** and **Ward Ordering** are used to group metrics categories
- **Source** and **Organism ID** are used to categorize # of pos/neg by anatomic site

- **Etest results reporting:**

SURRG_SpecimenID	PATIENTID	Site	Specimen_ Source	DOB	Sex	Collected Date	Received Date	Confirmed Date	Final Result	Scientist	DRUG_NAME	Interpret	MIC_BP_KB_VALUE
MILQ7010048	272290	MSTD	THR	5/3/00 0:00	M	3/1/18 0:00	3/2/18 0:00	3/6/18 0:00	3/7/18 0:00	RG	Azithromycin	S	0.25
MILQ7010048	272290	MSTD	THR	5/3/00 0:00	M	3/1/18 0:00	3/2/18 0:00	3/6/18 0:00	3/7/18 0:00	RG	Cefixime	S	<=0.016
MILQ7010048	272290	MSTD	THR	5/3/00 0:00	M	3/1/18 0:00	3/2/18 0:00	3/6/18 0:00	3/7/18 0:00	RG	Ceftriaxone	S	<=0.016
MILQ7010048	272290	MSTD	THR	5/3/00 0:00	M	3/1/18 0:00	3/2/18 0:00	3/6/18 0:00	3/7/18 0:00	RG	Gentamicin		8
MILQ7010077	276218	PPWA	CERVX	6/6/01 0:00	F	3/1/18 0:00	3/1/18 0:00	3/6/18 0:00	3/7/18 0:00	RG	Azithromycin	S	0.125
MILQ7010077	276218	PPWA	CERVX	6/6/01 0:00	F	3/1/18 0:00	3/1/18 0:00	3/6/18 0:00	3/7/18 0:00	RG	Cefixime	S	<=0.016
MILQ7010077	276218	PPWA	CERVX	6/6/01 0:00	F	3/1/18 0:00	3/1/18 0:00	3/6/18 0:00	3/7/18 0:00	RG	Ceftriaxone	S	<=0.016
MILQ7010077	276218	PPWA	CERVX	6/6/01 0:00	F	3/1/18 0:00	3/1/18 0:00	3/6/18 0:00	3/7/18 0:00	RG	Gentamicin		4

- **SpecimenID** is used to group lines into tests
- **Site and specimen source** are used to group results
- **Dates** are used in turnaround time calculations
- **Drug name** and **breakpoint value** are sorted and alert values are flagged

Python report generation

STD CLINIC NAAT METRICS

GENITAL SAMPLES:								NONGENITAL SAMPLES:													
Urethral		Vaginal		Endocervical		Genital (unspecified)		Urine		Rectal		Pharyngeal		Other		Unknown		SUM	
Total	Pos	Total	Pos	Total	Pos	Total	Pos	Total	Pos	Total	Pos	Total	Pos	Total	Pos	Total	Pos	Total	Pos
0	0	0	0	0	0	0	0	389	37	33	4	253	10	0	0	0	0	675	51

STD CLINIC CULTURE METRICS

GENITAL SAMPLES:								NONGENITAL SAMPLES:											
Urethral		Vaginal		Endocervical		Genital (unspecified)		Urine		Rectal		Pharyngeal		Other		Unknown		SUM	
Total	Pos	Total	Pos	Total	Pos	Total	Pos	Total	Pos	Total	Pos	Total	Pos	Total	Pos	Total	Pos	Total	Pos
71	28	0	0	12	5	0	0	0	0	22	1	52	3	0	0	0	0	157	37

- Total GC tests algorithm tallies all samples and generates reports as above
- Reports are in the same format as CDC SURRG metrics form, allowing results to be directly transcribed after verification

- Etest results algorithm tallies all GC-AST and generate a report as on the right
- Average and range of turnaround times are reported for genital and non-genital samples based on clinic type

STD CLINIC METRICS

	GENITAL SAMPLES:			NONGENITAL SAMPLES:		
	Col2Rec‡	Rec2Res	Col2Res	Col2Rec‡	Rec2Res	Col2Res
Median:	1	3	4	1	4	5.5
Min:	1	2	3	1	3	4
Max:	5	8	13	2	6	7
All steps completed in 1-5 days:	24					2
in 1-7 days:	32					4

Non-STD CLINIC METRICS

	GENITAL SAMPLES:			NONGENITAL SAMPLES:		
	Col2Rec‡	Rec2Res	Col2Res	Col2Rec‡	Rec2Res	Col2Res
Median:	N/a	N/a	N/a	N/a	N/a	N/a
Min:	N/a	N/a	N/a	N/a	N/a	N/a
Max:	N/a	N/a	N/a	N/a	N/a	N/a
All steps completed in 1-5 days:	0					0
in 1-7 days:	0					0

The overall number of positives were:

	STD	Non-STD
Genital:	33	0
Nongenital:	4	0
Total:	37	0

Future directions:

- Automated entry error checking:
  - A third Python script will detect mismatches between ordering ward, collector, and test type so they can be inspected and corrected by Data Specialist/LIS Specialist
  - Other error checking features coming soon (✔: implemented): City misspellings ✔, Zip code mismatch, Isolate/test mismatch
- Automated PDF filling using PYPFDTK (Python PDF toolkit)

References:

1. "SCC Soft Computer: Laboratory Information Systems Suite (LIS)." SCC Soft Computer Consultants, Inc. Accessed May 15, 2018. http://www.softcomputer.com/products-services/laboratory/
2. "Access: Elevate Data." Database Software and Applications | Microsoft Access. Accessed May 15, 2018. https://products.office.com/en-us/access.
3. "Python 3.6.5 Documentation." Python Software Foundation. Accessed May 15, 2018. https://docs.python.org/3/.
4. "Combating the Threat of Antibiotic-Resistant Gonorrhea." Centers for Disease Control and Prevention. February 15, 2018. Accessed May 15, 2018. https://www.cdc.gov/std/gonorrhea/arg/carb.htm.
5. "Antimicrobial Agents and Chemotherapy, Abbreviations." American Society for Microbiology. Accessed May 15, 2018. http://aac.asm.org/site/misc/journal-ita_abb.xhtml.

Acknowledgements: The authors gratefully acknowledge the Centers for Disease Control and Prevention for funding grant #G3802517000, Noah Leigh and Raquel Gomez from MHD Lab for technical assistance, Lori Amsterdam and John Pfister from WI Division of Public Health for project coordination and data analytic support.

For technical questions on Access or Python programs, contact Joshua Weiner @ jweine@milwaukee.gov.
For all other questions, contact Sanjib Bhattacharyya @ sbhatt@milwaukee.gov

Think Health Act Now!
CITY OF MILWAUKEE HEALTH DEPARTMENT