



COMMUNITY  
HEALTH CARE  
ASSOCIATION  
of New York State



*CHCANYS NYS-HCCN presents*

## Data You Can Trust: Building Validation & Hygiene Practices in Health Centers

Session 1:

Data Gone Wrong: Causes and Validation Strategies

July 23, 2026

For more information, please email Anita Li at [ali@CHCANYS.org](mailto:ali@CHCANYS.org)



This resource is supported by the Health Resources and Services Administration (HRSA) of the U.S. Department of Health and Human Services (HHS) as part of an award to CHCANYS' New York State Health Center Controlled Network (NYS-HCCN) totaling \$1,705,000 with 0% financed with non-governmental sources. The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement, by HRSA, HHS, or the U.S. Government. For more information, please visit [HRSA.gov](https://www.hrsa.gov).

# Session Agenda

**I. Welcome & Session Context**

**II. Module A: Why Data Hygiene Matters in FQHCs**

**III. Module B: Validation Frameworks & Practices**

**V. Wrap-Up & Q&A**

# ➤ Meet Your Facilitators



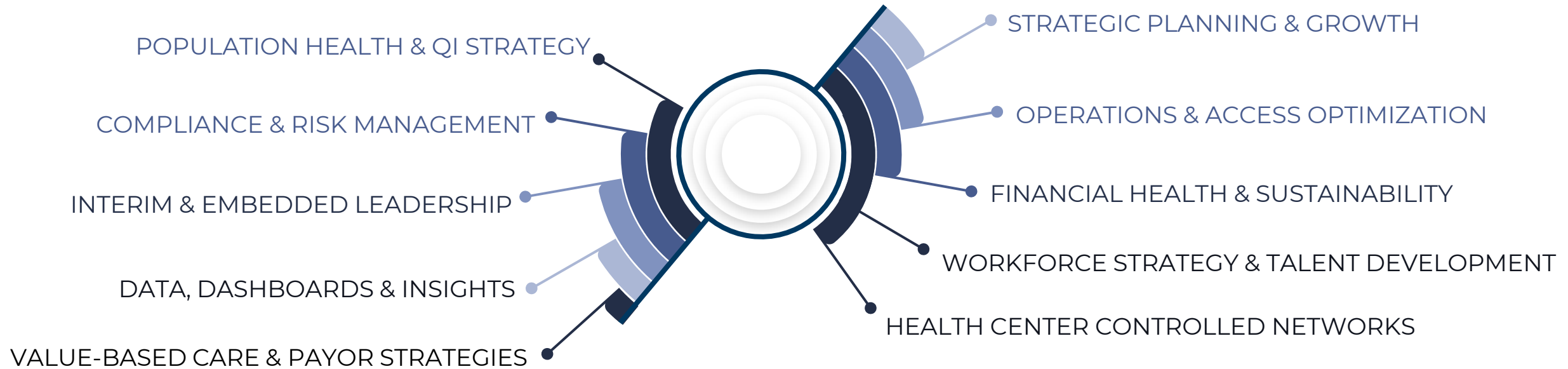
**Mike Wong**  
Managing Director  
Business Intelligence



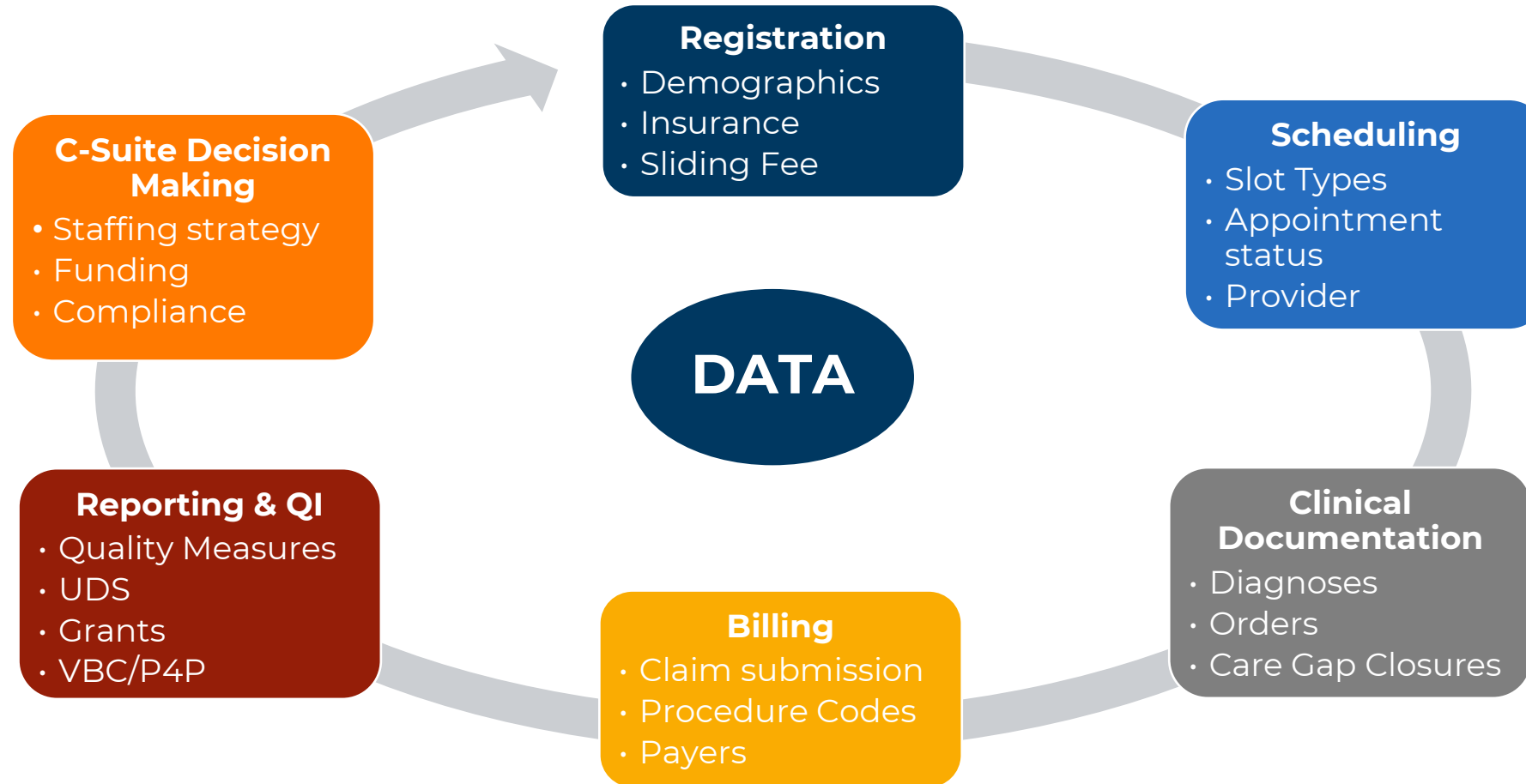
**Reyna Wong**  
Manager  
Business Intelligence

# ➤ Revolutionizing Community Health

Facktor partners with health centers, primary care associations and Health Center Controlled Networks to advance access, sustainability, and impact.



# ➤ Why Data Hygiene Matters in FQHC Operations



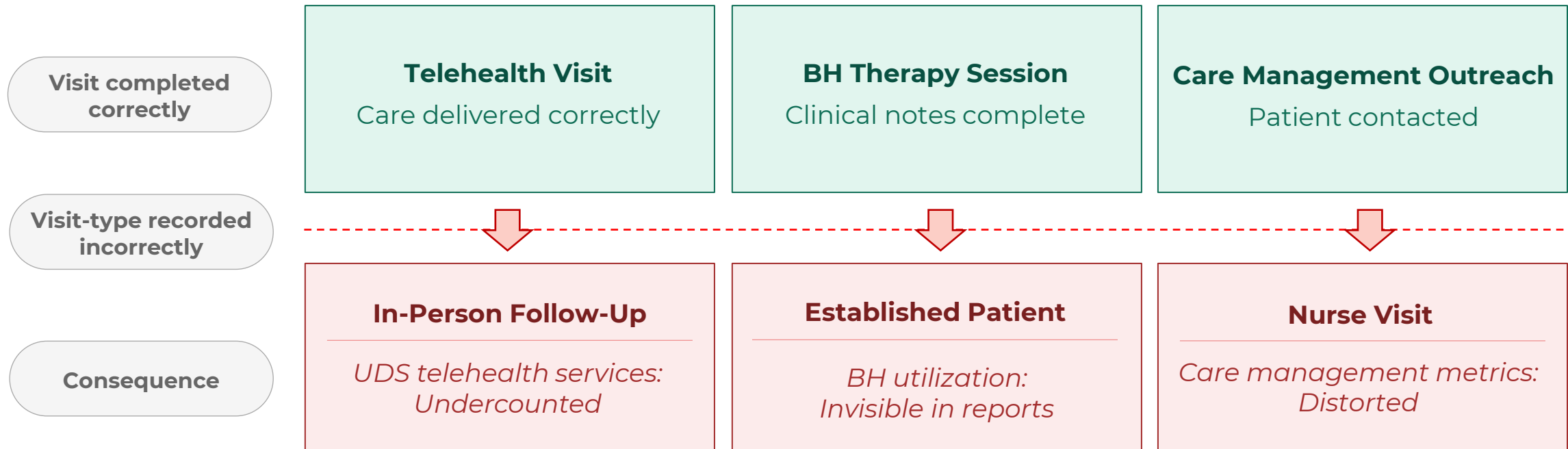
# ➤ Module A: Why Data Goes Bad

- 1 Common Sources of Dirty Data
- 2 Operational Origins of Reporting Problems
- 3 Importance of Aligning Workflows with Data Reports

# ➤ Inconsistent Visit-Type Classification



# ➤ Visit-Type Documentation Errors & Effects



**The patient may receive the correct care, but this won't be reflected in the data.**



# Visit-Type Clean-Up

## Keep visit types separate when...

- Visit types drive different billing codes or documentation requirements
- Distinct tracking is needed (program, grant, quality)
- Telehealth vs. in-person distinction affects reimbursement or reporting

### Examples

Telehealth Follow-Up

≠

In-Person Follow-Up

*Different UDS reporting lines*

BH Therapy Follow-Up

≠

BH Initial Assessment

*Different billing & documentation*

## Consolidate visit types when...

- No meaningful difference in billing, reporting, or care downstream
- Staff use the types interchangeably
- Distinction exists historically but serves no current purpose

### Examples

New Office Visit

=

New Patient Visit

Follow-up Visit

=

Established Patient Visit

# ➤ Duplicate Records: Before & After

*Same patient, same date of service: Two systems, Two records, Zero complete picture*

## BEFORE: Fragmented Records

| Field     | Record #1<br>(Scheduling)                                                                                                                                                   | Record #2<br>(Registration) |
|-----------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------|
| MRN       | 10024567                                                                                                                                                                    | 10089211                    |
| Name      | Michael Johnson                                                                                                                                                             | Mike Johnson                |
| Insurance | Sunshine Health                                                                                                                                                             | Sunshine Health<br>Medicaid |
| Source    | Scheduling<br>system                                                                                                                                                        | Registration<br>(new entry) |
| Impact    | <ul style="list-style-type: none"><li>• UDS count inflated</li><li>• Quality measure attribution split</li><li>• Payer mismatch</li><li>• Duplicate care outreach</li></ul> |                             |



## AFTER: Single Accurate Record

| Field     | Merged Record (Correct)                                                                                                                                                                |
|-----------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| MRN       | 10024567 (single canonical MRN)                                                                                                                                                        |
| Name      | Michael Johnson (verified at registration)                                                                                                                                             |
| Insurance | Sunshine Health Medicaid<br>(verified, structured field)                                                                                                                               |
| Source    | All systems linked to single record                                                                                                                                                    |
| Result    | <ul style="list-style-type: none"><li>• Accurate UDS count</li><li>• Correct payer attribution</li><li>• Quality measure properly attributed</li><li>• Complete care history</li></ul> |

# ➤ Coding Variation Across Sites & Teams

*Same workflows, different reporting outputs — and different decisions made from them*



## Procedure Code Variation

- Same chronic care visit:  
One provider bills 99213, another bills 99214 based on habit, not documentation level
- Result: cross-site productivity comparisons unreliable; audit exposure



## Diagnosis Coding Inconsistencies

- Active hypertension: Coded as I10, I10.9 (invalid), or Z87.39 (history) across three providers
- Specificity level varies:  
Some providers code to the highest levels
- Z-codes for housing instability:  
Documented at Site A, absent at Site B



## Visit Category Variation

- BH therapy follow-up coded as “Established Patient” at one site (not a BH visit type)
  - Telehealth coded as in-person at another site
  - Result: UDS BH line undercounted & grant deliverable unverifiable



## Service Location Miscoding

- Satellite clinic visits coded to main facility POS = UDS site data unreliable
- School-based visits coded as office visits = grant deliverables unverifiable
- Result: Access metrics skewed & grant reporting fails audit

# > Scheduling Template Irregularities



## Blocked Appointment Slots

- Slots held for same-day, admin, or provider preference reduce bookable time without being visible



## Inconsistent Provider Templates

- Provider templates reduce administrative burden & supports workflow flexibility
- Standardizing templates ensures consistency across sites and makes reporting reliable



## Overbook & Same-Day Scheduling

- Any ad hoc slots bypass template logic entirely, creating phantom capacity and skewing no-show and access metrics



## Visit Duration Inconsistencies

- Slot lengths that don't match actual visit time compress schedules and produce access metrics that don't reflect reality

### Metrics affected:

No-show  
rate

Provider  
utilization

Third Next Available  
Appointment (TNAA)

Slot fill  
percentage

Access  
reporting

# ➤ Scheduling Template Irregularities

| Provider A — Inconsistent template |                                     | Provider B — Consistent template |                                |
|------------------------------------|-------------------------------------|----------------------------------|--------------------------------|
| 8:00 AM                            | Office Visit (duration unclear)     | 8:00 AM                          | New Patient (60 min)           |
| 8:20 AM                            | <b>BLOCKED</b> — reason unspecified | 8:20 AM                          | Established — Office Visit     |
| 8:40 AM                            | Follow-Up (in-person or tele?)      | 8:40 AM                          | Established — Office Visit     |
| 9:00 AM                            | New Patient                         | 9:00 AM                          | <b>BLOCKED</b> — Admin / lunch |
| 9:20 AM                            | <b>BLOCKED</b> — reason unspecified | 9:20 AM                          | Telehealth Follow-Up           |
| 9:40 AM                            | Telehealth — added same day         | 9:40 AM                          | Established — Office Visit     |
| 10:00 AM                           | Office Visit (duration unclear)     | 10:00 AM                         | New Patient (60 min)           |
| 10:20 AM                           | <b>BLOCKED</b> — reason unspecified | 10:20 AM                         | Telehealth Follow-Up           |
| 10:40 AM                           | <b>Follow-Up (wrong visit type)</b> | 10:40 AM                         | Established — Office Visit     |

| Metric               | Provider A (inconsistent)   | Provider B (consistent)      |
|----------------------|-----------------------------|------------------------------|
| Slot fill %          | Unclear — mixed slot types  | Accurate — predictable slots |
| No-show rate         | Unreliable — ad hoc adds    | Reliable                     |
| Provider utilization | Inconsistent / unmeasurable | Accurate                     |
| Third next available | Appears open — misleading   | Reflects true access         |

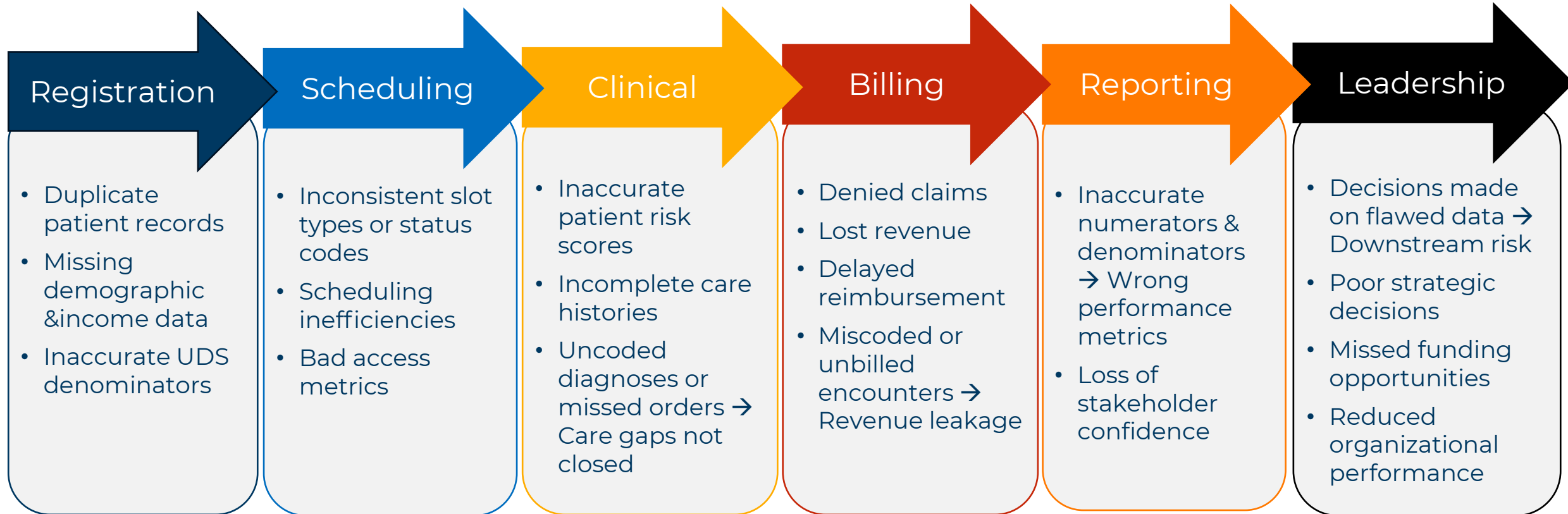
## Poll Question: Data Quality Problems Source

### ➤ Where do data quality problems most commonly originate at your health center?

- A. Front desk / registration workflows
- B. Scheduling templates and visit type setup
- C. Clinical documentation and coding
- D. We haven't formally identified the source
- E. Other: Please type in chat



# ➤ Operational Cost of Bad Data



# ➤ Key Takeaways from Module A

## The Scaling Effect

Dirty data is usually created long before it reaches the dashboard

WORKFLOW INCONSISTENCIES ➡ REPORTING, FINANCE, AND STRATEGY IMPACTS

### Bottom Line

- Root causes are operational, not reporting problems
- Better workflow design creates more reliable data without added burden

**Module B:** Frameworks and tools to detect, validate, and prevent dirty data

- including Excel skills and leave-behind checklist



# ➤ Module B: Validation Framework & Practices

Practical tools and frameworks to detect, validate, and prevent dirty data

- 1 Governance
- 2 Validation
- 3 Reconciliation
- 4 Practical Operational Tools

# ➤ Data Dictionaries Support Governance

## Why Data Dictionaries Matter

- Different teams define the same metric differently
- No shared standard = different numbers from same data
- Dictionaries create operational agreement, not just documentation
- Reduce reconciliation time before every report

## Core Components

- Field / Metric name and standardized definition
- Source system (EHR, PM, billing) and data owner
- Reporting usage (UDS, grants, productivity)
- *Known conflicts or common misinterpretations*

# ➤ Data Dictionary Example

| Visit Type                                | Definition                                                          | Billable?                                  | Required Fields                                   | Reporting Usage        | Common Pitfalls                                                                |
|-------------------------------------------|---------------------------------------------------------------------|--------------------------------------------|---------------------------------------------------|------------------------|--------------------------------------------------------------------------------|
| <b>Established Patient – Office Visit</b> | Routine primary care for established patient, in-person             | <b>Yes</b><br>CPT: 99213<br>Modifier: None | Visit type, POS (office), signed encounter        | UDS Table 5            | <b>Used as catch-all when staff is unsure of correct type</b>                  |
| <b>Telehealth Follow-Up</b>               | Synchronous audio-video follow-up for established patient           | <b>Yes</b><br>CPT: 99213<br>Modifier: 95   | Visit type, POS 02, telehealth modifier           | UDS – Telehealth Line  | <b>Coded as in-person follow-up when platform isn't confirmed</b>              |
| <b>BH Therapy Telehealth Follow-Up</b>    | Behavioral health individual therapy session with licensed provider | <b>Yes</b><br>CPT: 90834<br>Modifier: 95   | BH visit type, provider credential, session notes | UDS – BH Services Line | <b>Coded as established patient which makes it invisible in BH utilization</b> |

# ➤ Metric Definition Governance

The same metric can produce different results when definitions are not standardized.

- Who owns the metric? What's the definition for each metric?

| Metric              | Scheduling Definition<br>(<2-hour notice)                     | Billing Definition<br>(no encounter billed)                       | Gap                                                                                    |
|---------------------|---------------------------------------------------------------|-------------------------------------------------------------------|----------------------------------------------------------------------------------------|
| June No-Show Count  | 84<br>No-shows                                                | 61<br>No-shows                                                    | <b>23-visit discrepancy from<br/>the same dataset</b>                                  |
| No-Show Rate        | 18.4%                                                         | 13.4%                                                             | <b>5 percentage point<br/>difference of no-show rates</b>                              |
| Leadership Decision | May trigger patient<br>outreach program or<br>staffing review | May be considered within<br>acceptable range –<br>No action taken | <b>Both teams are right but<br/>neither knows if they're<br/>measuring differently</b> |

|                                                         |                  |                     |                          |                                            |
|---------------------------------------------------------|------------------|---------------------|--------------------------|--------------------------------------------|
| <b>Example Metrics<br/>Requiring Defined Ownership:</b> | No-show<br>Rates | Completed<br>Visits | Provider<br>Productivity | Third Next Available<br>Appointment (TNAA) |
|---------------------------------------------------------|------------------|---------------------|--------------------------|--------------------------------------------|

# ➤ Start with the Question, Not the Report

Before building a report...

|          |                                         |
|----------|-----------------------------------------|
| <b>1</b> | <b>Define the question</b>              |
| <b>2</b> | <b>Agree on the metric definition</b>   |
| <b>3</b> | <b>Identify the correct source data</b> |
| <b>4</b> | <b>Develop the report logic</b>         |

EX: How many patients did we see last month?

| <b>System</b>        | <b>What it Measures</b> | <b>Result</b> |
|----------------------|-------------------------|---------------|
| <b>Scheduling/PM</b> | Completed Encounters    | 10,500        |
| <b>EHR</b>           | Documented Encounters   | 10,100        |
| <b>Billing</b>       | Claims Submitted        | 9,800         |

# ➤ Source-to-Report Reconciliation

Identify discrepancies before they influence decisions



Routine reconciliation workflows regularly surface issues

| Validation Check Examples | What You're Looking For                                      |
|---------------------------|--------------------------------------------------------------|
| Visit Counts              | <b>Documented encounters that never make it to billing</b>   |
| Payer Mix                 | <b>Registration payer different than billed payer</b>        |
| Care Gap Reporting        | <b>Services documented but not in structured data fields</b> |

## Poll Question: Source-to-Report Reconciliation

➤ **How does your organization currently handle discrepancies between source systems and leadership-facing reports?**

- A. We have a formal reconciliation step before reports go out
- B. We catch issues informally, depending on who pulls the report
- C. Reports typically go out without a dedicated reconciliation step
- D. Other: Please type in chat



# ➤ Threshold-Based Anomaly Detection

*Monthly habits that can catch issues*

## Provider Visit Volume



- Compare visit volume month-over-month
- Flag any unexpected drops > 20-30%

## Telehealth vs In-Person Split



- Review the ratio monthly
- Flag unexpected spikes or reversals

## Payer Mix Changes



- Watch for sudden shifts
- Often signals registration or interface issues

## Unsigned/Incomplete Encounters



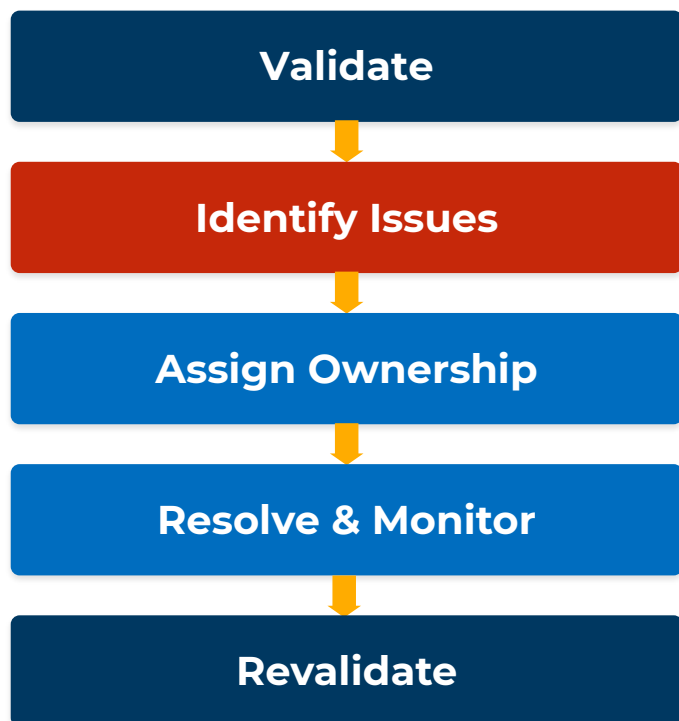
- Review weekly & follow-up on unsigned encounters
- Don't wait until month-end close



# ➤ Building a Sustainable Validation Workflow

Data quality improves when validation becomes operationalized and repeatable.

## Validation Cycle



## Core Components

| Component                         | Example                                                                 |
|-----------------------------------|-------------------------------------------------------------------------|
| <b>Validation Routine</b>         | Monthly review of key metrics and reports                               |
| <b>Defined Ownership</b>          | Assigned owner each report and/or collection of metrics                 |
| <b>Reconciliation Checkpoints</b> | EHR vs. Billing <i>OR</i> PM vs. Dashboard <i>OR</i> Quality vs. Claims |
| <b>Escalation Process</b>         | Defined path for unresolved discrepancies or system issues              |

# ➤ Practical Excel Skills for Data Validation

Recognize these tools and how they apply. Full instructions will be included in take-home deliverables.

| COUNTIF                                                                                                  | XLOOKUP / VLOOKUP                                                                                                                                      | Conditional Formatting                                                                                    |
|----------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------|
| <b>What It Does:</b><br>Counts cells matching a condition: Surfaces blanks and unexpected entries        | <b>What It Does:</b><br>Matches values across two datasets: finds records in one system missing in another                                             | <b>What It Does:</b><br>Color-flags cells meeting a rule: makes patterns and outliers immediately visible |
| <b>FQHC Example:</b><br><i>=COUNTIF(B:B,"") to flag registration records with blank insurance fields</i> | <b>FQHC Example:</b><br><i>=XLOOKUP("10123", A2:A10, B2:B10) to cross reference EHR encounters against billed encounters to find unmatched records</i> | <b>FQHC Example:</b><br><i>Highlight duplicate MRNs or visit counts that fall outside expected ranges</i> |
| Checklist Step 3:<br>Registration completeness check                                                     | Checklist Step 5:<br>Source-to-report reconciliation                                                                                                   | Checklist Step 4:<br>Anomaly flagging worksheet                                                           |

# ➤ Key Takeaways from Module B & Wrap Up

- 1 Small improvements in workflow consistency create significant reporting gains
- 2 Validation should be embedded into routine operations
- 3 Shared definitions are the foundation of trustworthy reporting

## Session Leave-Behinds

**Data Hygiene Validation Checklist**

**Data Dictionary Template**

Module C (Next Session)...

**Data Quality Communication Toolkit**

# ➤ Module C Preview

- 1 Closing the Feedback Loop
- 2 Reducing Friction at Documentation
- 3 Accountability Structures
- 4 Aligning Leadership Priorities with Frontline Workflows



# Please fill out our survey!

Please share your feedback using the survey link in the chat, the QR code, or the link in the follow up email!

Completing the survey helps us to provide relevant and helpful information. Thank you in advance!



# QUESTIONS?

 805.723.4107  
mike.wong@facktorhealth.com

 425.495.9938  
reyna.wong@facktorhealth.com