RADARS® SYSTEM 2020

14TH ANNUAL SCIENTIFIC MEETING

How Low Can You Go: Solving the Challenge of Low Volume Endorsements in the General Population

Joshua C. Black, Ph.D.

Rocky Mountain Poison & Drug Safety

Outline

- Background & Hypothesis: Detection of rare events
- Analytic Strategy
- Defining Two Epidemiological Paradigms with General Population Surveys
- Test-Retest Reliability as a Function of Dispensing
- Useful Information from the Signal Detection Paradigm

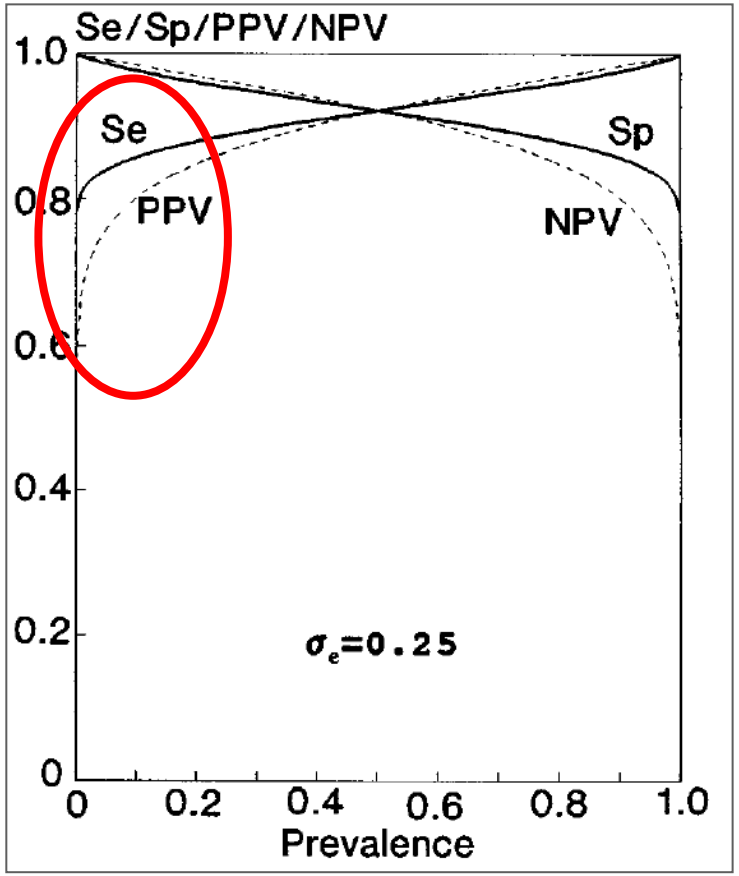
Objectives: Investigating Rare Endorsements

- Detecting new and low volume products in the general population can be framed as a rare outcome problem
- RADARS System General Population Survey:
  - Survey of Non-Medical Use of Prescription Drugs (NMURx) Program
- Assume:
  - Dispensing can be used as a measure of "rarity"
- Hypothesis:
  1. A relationship between dispensing and estimates can be used to validate surveillance of a product in the general population

Background

- Detection as function of prevalence modelled previously[1]

Smaller prevalence leads to lower sensitivity and precision.
False positives starts to outweigh false negative.



[1]Brenner H, Gefeller O. "Variation of Sensitivity, Specificity, likelihood ratios, and predictive values with disease prevalence." Statistics in Medicine 1997; 16(9): 981-991

Analytic Strategy

- Truth of individual responses cannot be directly confirmed
  - Dispensing is a proxy for actual use
- Goals:
  - Quantify threshold in association of dispensing and use estimates (Part 1)
  - Demonstrate reliable estimates and association with dispensing (Part 2)
  - Explore qualitative analysis of rare behavior (Part 3)
- Joinpoint Regression[2]
  - Past experience indicated a threshold relationship was likely
  - Disjointed linear regression; estimates "transition" points with CIs

[2]Joinpoint Regression Program, Version 4.8.0.1 - April 2020; Statistical Methodology and Applications Branch, Surveillance Research Program, National Cancer Institute

Study Design Summary of the NMURx Program

- Setting: Online survey panel of general population
  - Digital distribution through commercial company
- Sample Size: 120,000 respondents
  - 4 launches to date: 3rd quarter 2018 to 1st quarter 2020
  - Average Completion Rate: 75.6%
- Key measure: Drug use (any medical or non-medical)
- Key adjustments[3]:
  - Calibration weighting to address non-probability sampling
  - Careless response exclusion

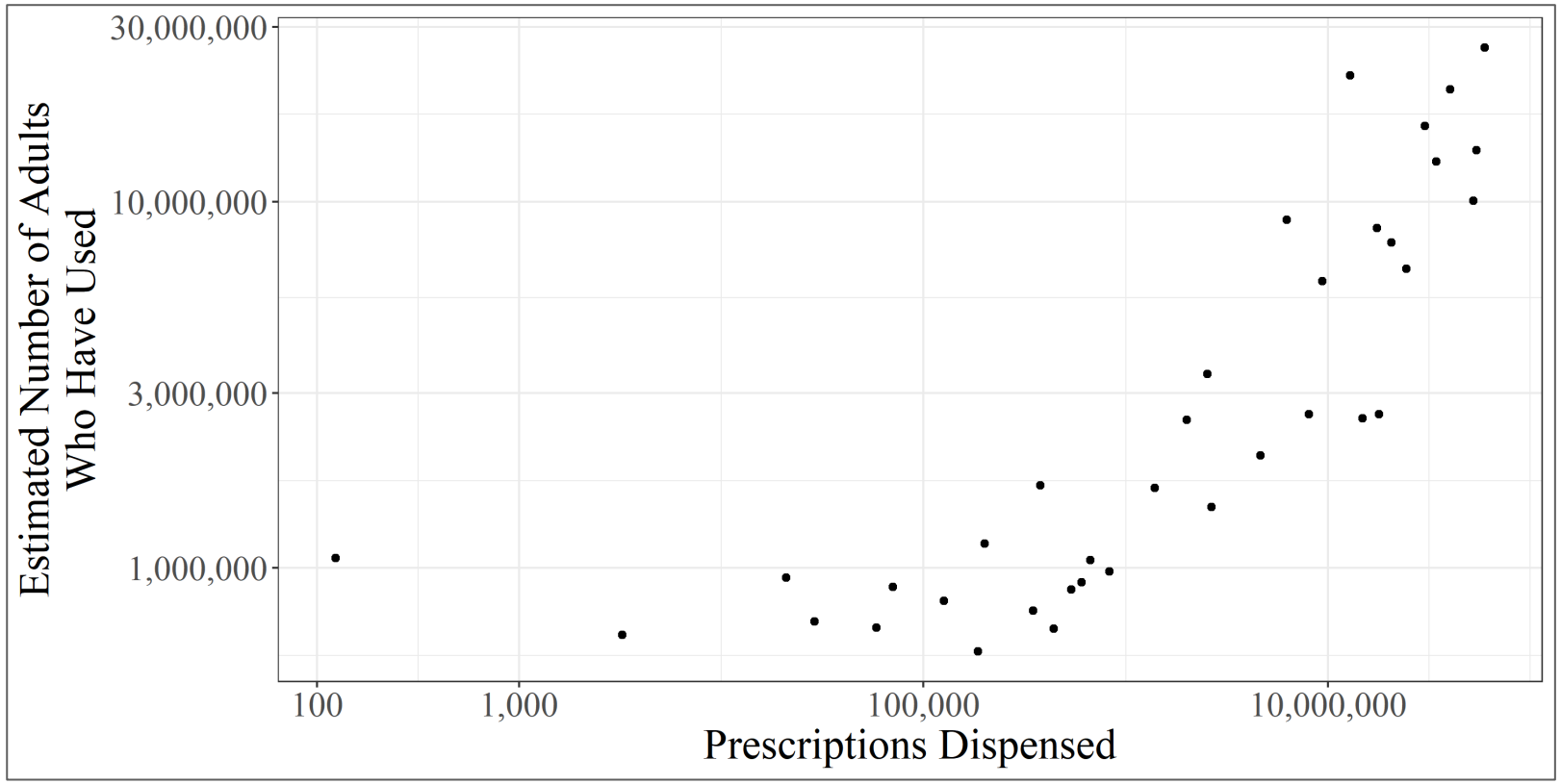
[3]Black JC, Rockhill K, Forber A, Amioka E, May KP, Haynes CM, Dasgupta N, Dart RC. "An Online Survey for Pharmacoepidemiological Investigation (Survey of Non-Medical Use of Prescription Drugs Program): Validation Study." J Med Internet Res 2019; 21(10):e15830

Part 1: Threshold Function in Dispensing and Use Relationship

Threshold Identification Methods

- Past experience indicated that API endorsements have lower limit
  - Lead us to using joinpoint method to quantify threshold
- Average past year prescriptions dispensed (IQVIA™) for active ingredients (independent x-axis)
- Estimated of number of adults who have used (dependent y-axis)
- Joinpoint
  - Identified best fit model between 0, 1, 2, 3, & 4 thresholds
  - Resampling to estimate threshold and CIs
  - Model on log-log scale

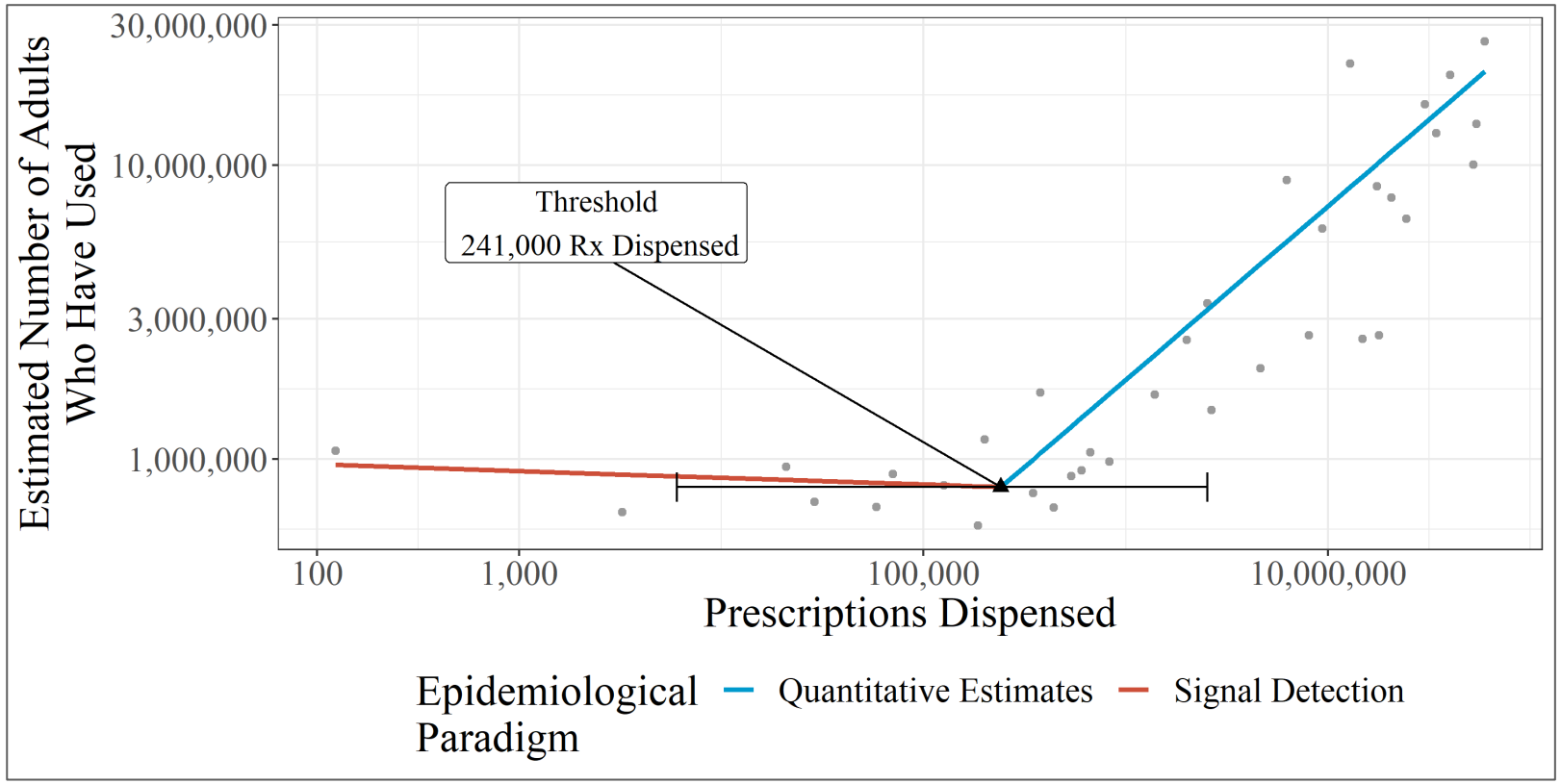
Threshold Identification Results: Joinpoint Analysis



Results Summary:

- Visually, there appears to be a change in the association around 100,000 Rx dispensed

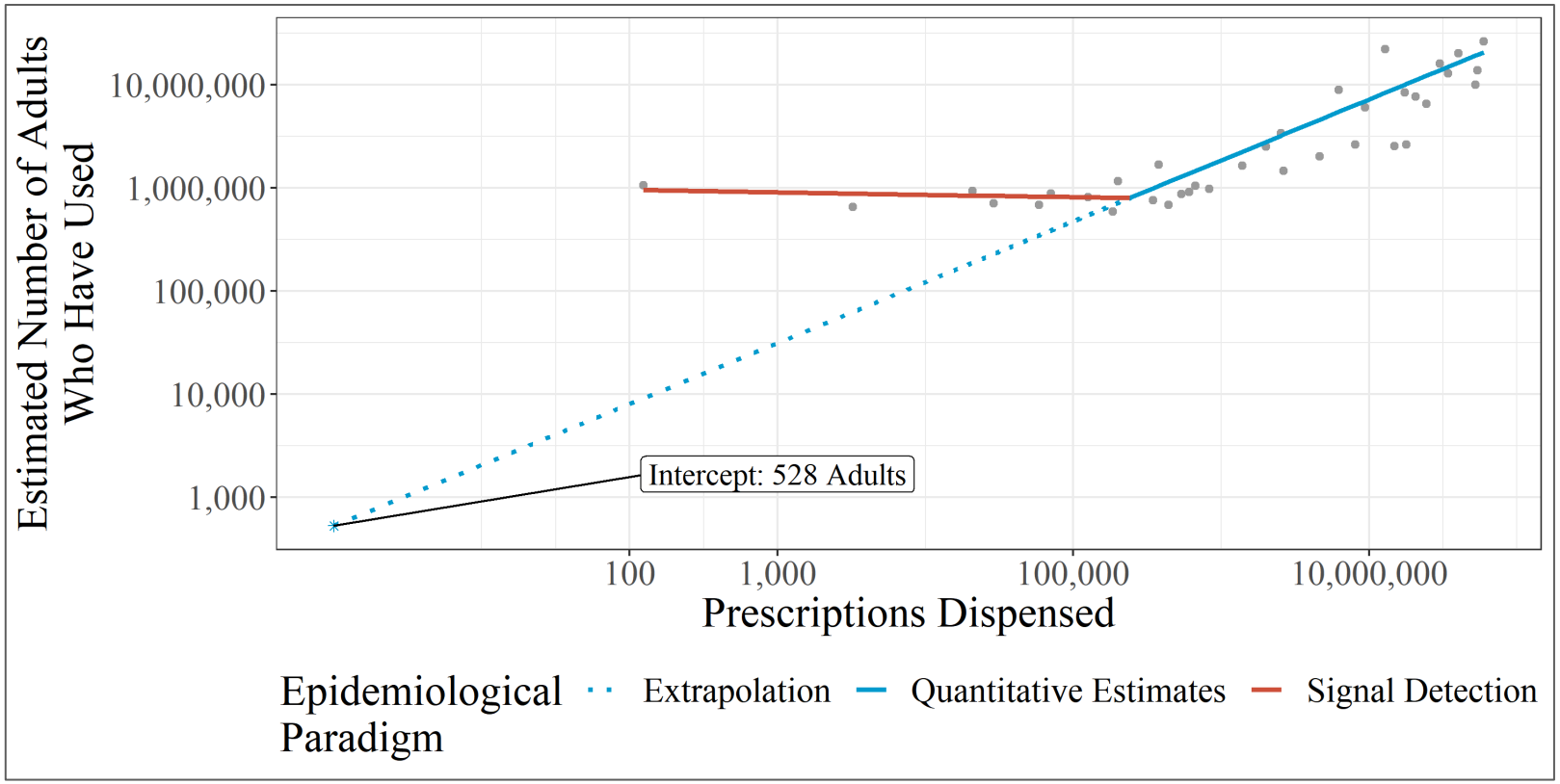
Threshold Identification Results: Joinpoint Analysis



Results Summary:

- Best model
  - Single threshold
  - Better fit than 0, 2, 3 or 4 thresholds
- Above threshold
  - Significant linear association (log scale)
- Below threshold
  - No association (not significant)
- Suggests two different paradigms for investigation

Threshold Identification Results: Extrapolation



Results Summary:

- Extrapolation
  - Intercepts near 1
  - Shows internal validity of model
- Implies estimates above the threshold are valid with respect to dispensing

Threshold Identification Conclusions

- A single threshold demarcates two distinct epidemiological paradigms
  - Quantitative estimates: Statistically valid, generalizable estimates of use, within the context of the sampling frame
  - Signal detection: Not generalizable, but true positives still informative
- Dispensing can be used to guide which surveillance is best suited
- Limitations:
  - Uncertainty in dispensing not accounted for
  - Nonprobability sampling – Mitigated by calibration weights

Part 2: Reliability in the NMURx Program

Reliability in the NMURx Program
Study Design

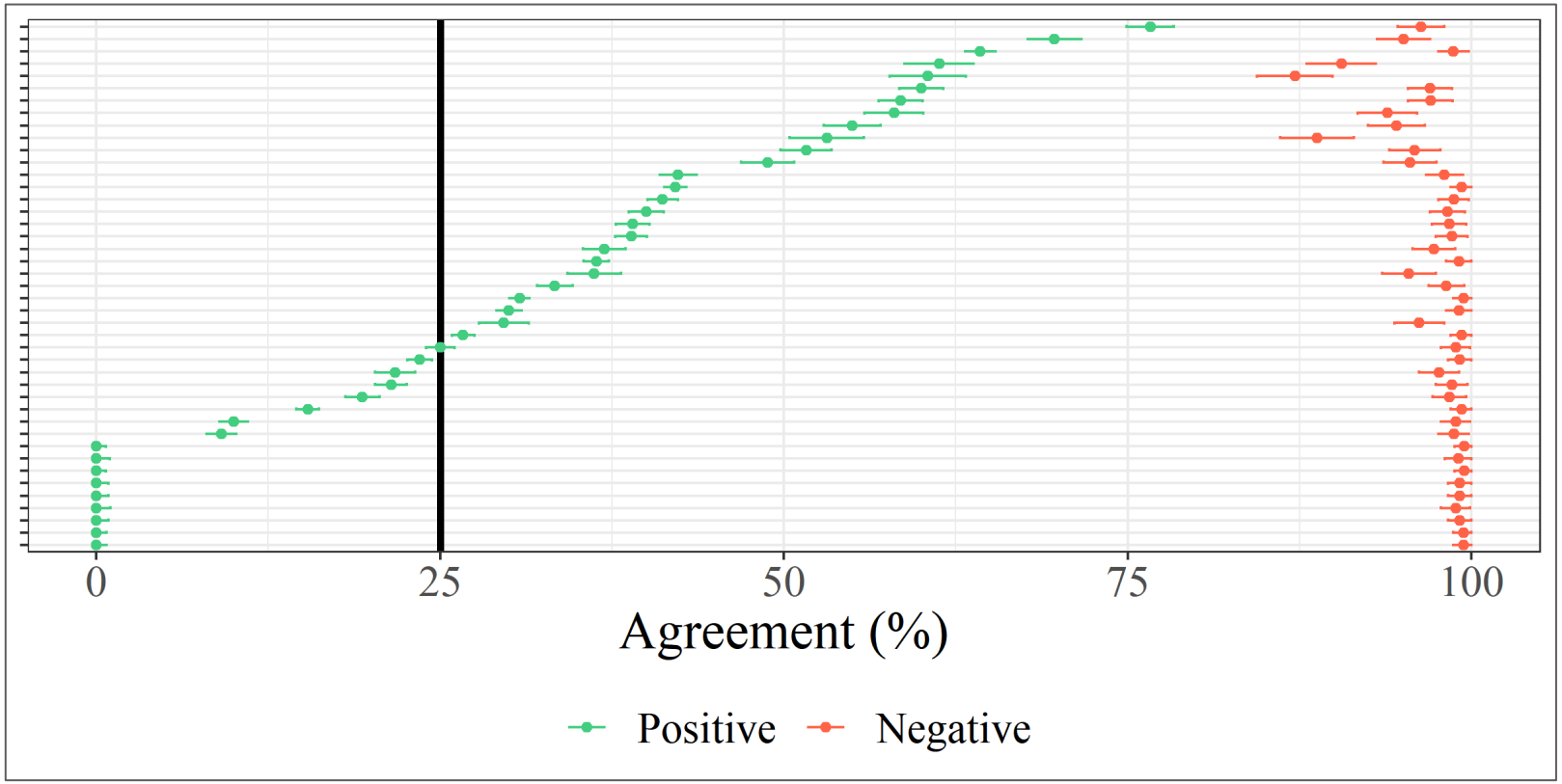
- Re-contacted 1,844 respondents from 3rd quarter 2019
  - 789 retook survey (42.8% re-contact rate)
- Same questionnaire (Past year use question)
- 1-2 months after initial contact
- Drug class and individual drug reliability measures
  - Kappa: Modelled with joinpoint regression
  - Prevalence-adjusted, bias adjusted kappa (PABAK)
  - Overall, positive, and negative agreement

Reliability in the NMURx Program
Test-Retest Results: Drug Class

Drug Class (Past Year Use)	Overall Agreement	Positive Agreement	Negative Agreement	Kappa	Prevalence Adjusted Bias Adjusted Kappa
Pain Relievers	81.6%	83.9%	78.6%	0.63	0.63
Sedatives	81.7%	72.2%	86.4%	0.59	0.63
Stimulants	88.3%	59.6%	93.2%	0.53	0.77
Rx Cannabinoids	92.0%	40.0%	95.7%	0.36	0.84

- Overall agreement at the drug class level is good
- Adjustments to kappa needed to account for prevalence imbalance in the observational data

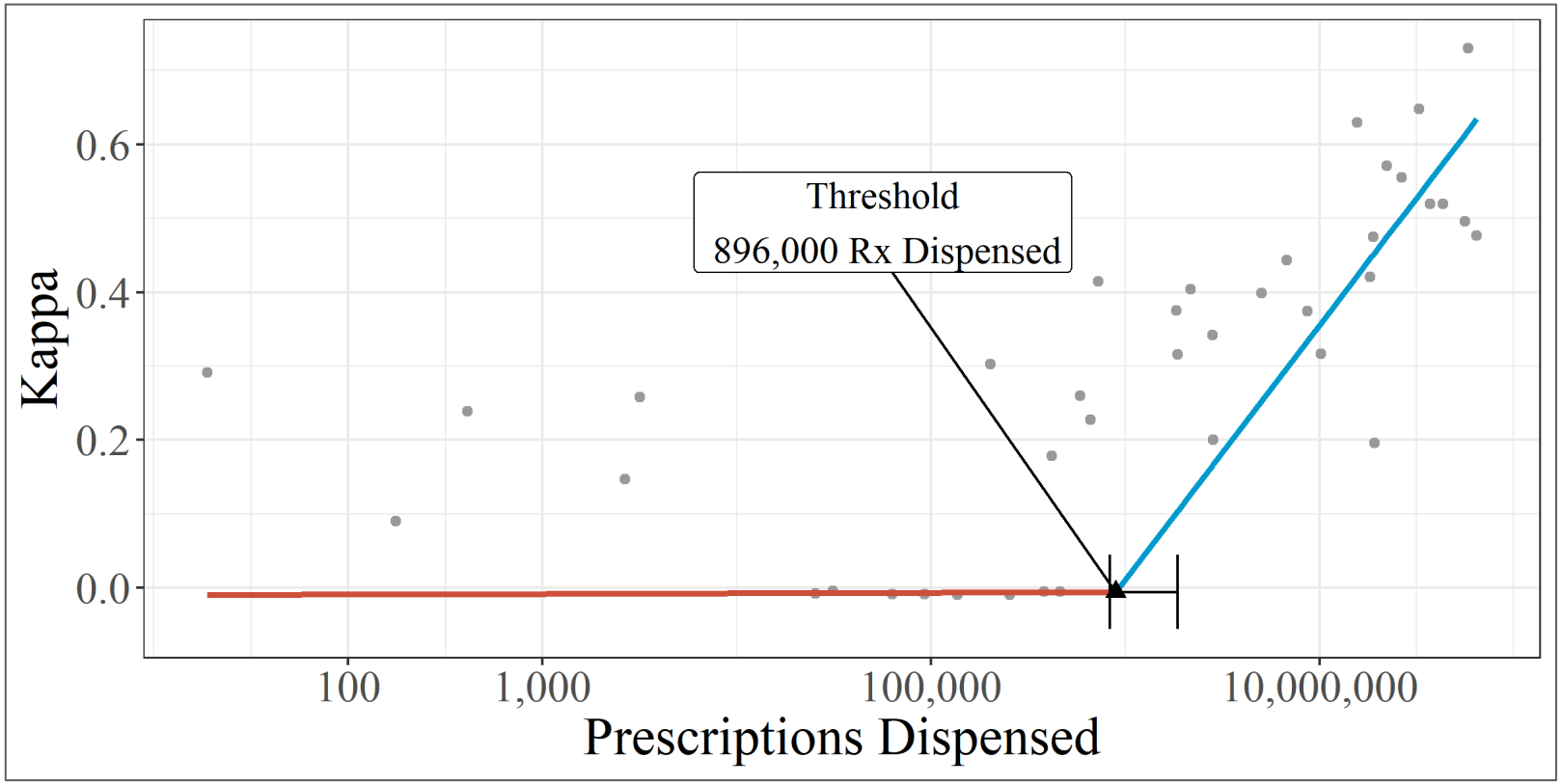
Reliability in the NMURx Program
Test-Retest Results: Individual Drugs



Results Summary:

- Results more diverse by API
  - Each horizontal line is 1 API
  - Ordered by positive agreement
- Vertical black line represents chance agreement
- Negative agreements mostly >80%
- Positive agreements lower, but more reliable than random chance

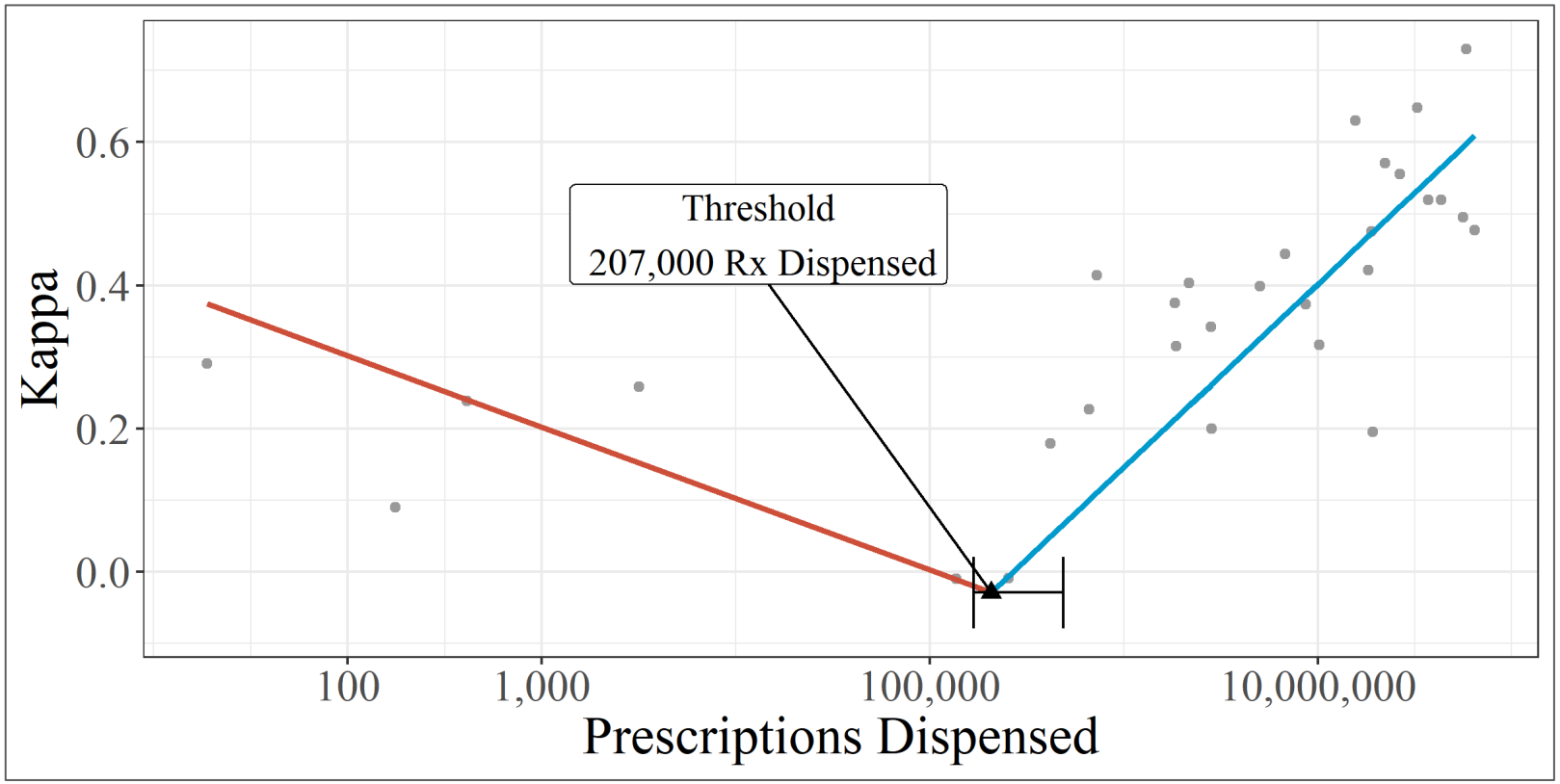
Reliability in the NMURx Program
Test-Retest Results: Threshold Analysis



Results Summary:

- Each point is the kappa for one API
- Functional relationship similar to prevalence estimates
- Single threshold
- Inconclusive results due to several APIs with low sample size
  - Low sample size drives kappas toward zero
  - Biases the threshold estimate and confidence interval

Reliability in the NMURx Program
Sensitivity: Removing low sample sizes



Sensitivity Summary:

- Removed APIs in lower quartile of sample size
- Model performing better relative to actual data
  - Suggests low sample sizes perturbing model
- Threshold similar to prevalence vs dispensing threshold
  - 207,000 vs 241,000

Reliability in the NMURx Program Conclusions

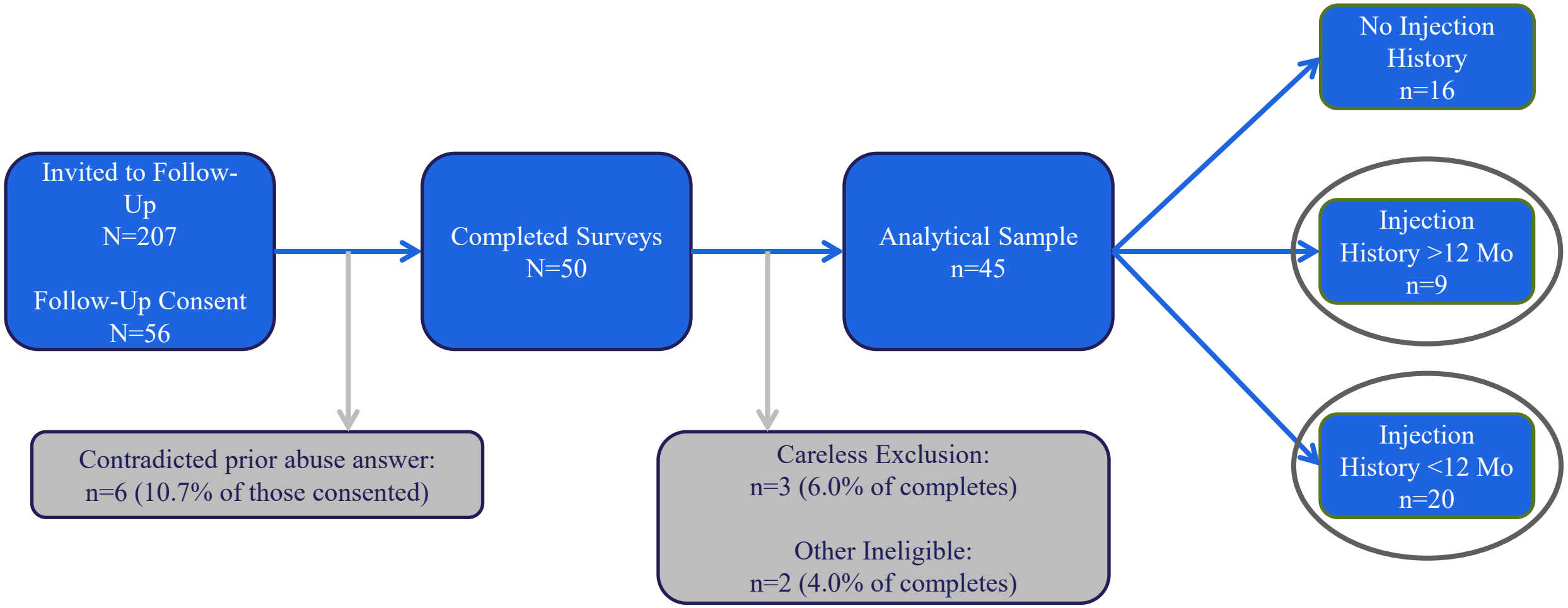
- Drug use questions demonstrate good reliability
  - Kappa estimates influenced by prevalence
- Single threshold also observed in reliability data
  - Also suggests two epidemiological paradigms
- Limitations:
  - Results only suggestive due to sample size

Part 3: Useful Information under the Signal Detection Paradigm

Opioid Injection Follow-up Survey: Pilot Study Design

- Signal detection paradigm can be used to study rare behaviors qualitatively
- Current Example: Investigate injection of opioid pills
  - Setting: Respondent re-contact via online portal
  - Sample: Respondents who reported abuse of an opioid pill in the past year
  - Unique questionnaire focused on relevant injection behavior
  - Delivered after each of 1st and 3rd quarter launches in 2019
- Respondents must confirm abuse behavior

Opioid Injection Follow-up Survey: Pilot Respondent Flow Chart



Opioid Injection Follow-up Survey: Pilot Results: Motivations for initiation injection

We asked: "What lead you to inject the [opioid] pill? Tell us about your **first** experience"

It was exhilarating experience and I enjoyed it so. It was absolutely amazing.

Feels good honestly

To get high

It was related to pains and other issues

Much pain

Chronic pain

Knee replace for pain

To feel less pain

Out of 29 responses:

- 8 referred to pain
- 3 referred to a high experience
- 1 referred to drug switching
- 4 were uninformative

Opioid Injection Follow-up Survey: Pilot Results: Injection behavior in past year

Most Frequently Used Ingredient	Frequency, n (%) (N=20)
Oxycodone	5 (25%)
Oxymorphone	5 (25%)
Hydrocodone	3 (15%)
Codeine	3 (15%)
Morphine	2 (10%)
Hydromorphone	1 (5%)
Tramadol	1 (5%)

Narrative accounts can be combined with targeted questions to understand why and how people use products

Injection Regularity	Frequency, n (%) (N=20)
Once a month or less often	8 (40%)
Once a week	8 (40%)
Once a day	2 (10%)
Multiple times a day	2 (10%)

Typical Time Spent Preparing	Frequency, n (%) (N=20)
<5 Minutes	6 (30%)
6 to 15 Minutes	9 (45%)
16 to 30 Minutes	4 (20%)
>6 hours	1 (5%)

Needle Sharing	Frequency, n (%) (N=20)
No sharing	5 (25%)
1 person	9 (45%)
2+ people	6 (30%)

Results Summary:

- Diverse APIs
- Most injected once a week or less often
- Most spent 15 minutes or less preparing
- Most shared needles with at least one person

Opioid Injection Follow-up Survey: Pilot Conclusions

- Follow-up verifies endorsement
  - Limits false positive bias
- Allows narrative and close-ended question development
- Tailored questions can address emergent concerns (e.g., needle sharing)
- Limitations:
  - Rapid follow-up required
  - Might require waves of follow-up

Overall Conclusions

Overall Conclusions

- Identified a threshold to demarcate paradigms and inference frameworks using a general population survey
  - Above the threshold, estimates are valid, reliable, and representative of the population
  - Below the threshold, tailored questionnaires and qualitative analysis are informative of emergent behavior
- Use of an online panel can work within both paradigms using a single participant resource

ROCKY MOUNTAIN POISON & DRUG SAFETY
Saving lives with answers.™

DENVER HEALTH
est. 1860
FOR LIFE'S JOURNEY

Questions?

Joshua C. Black, Ph.D.

Joshua.black@rmpds.org

RADARS® SYSTEM 2020
14TH ANNUAL SCIENTIFIC MEETING