RMPDS 19TH ANNUAL SCIENTIFIC MEETING

Frankenstein Surveys

Data Fusion to Estimate Behaviors for Rare Populations

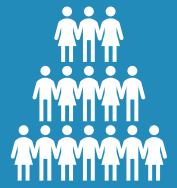
Joshua C. Black, PhD
Senior Statistical Scientist

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

DENVER HEALTH
est. 1860
FOR LIFE'S JOURNEY

Validating Opt-In, Online, Panel Surveys

Survey validation proactively uses best practices to minimize bias, then tests accuracy using a variety of methods.



Sample Validation

Goal is to ensure the sample collected is a faithful representation of the target population



Psychometric Validation

Goal is to ensure that variable measurement is accurate and represents what the researchers intend it to represent

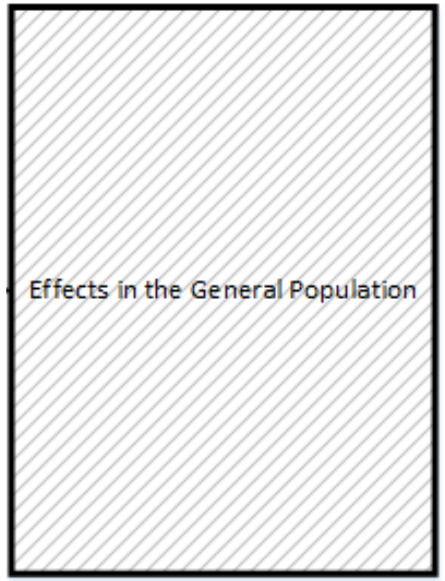
Making Estimates Generalizable

- Health professionals use population health data to guide intervention development and rollout

- Large surveys measure drug use prevalence very well
    - Fail to measure detailed behaviors
- Small surveys measure detailed behavior very well
    - Challenge is generalizing estimates

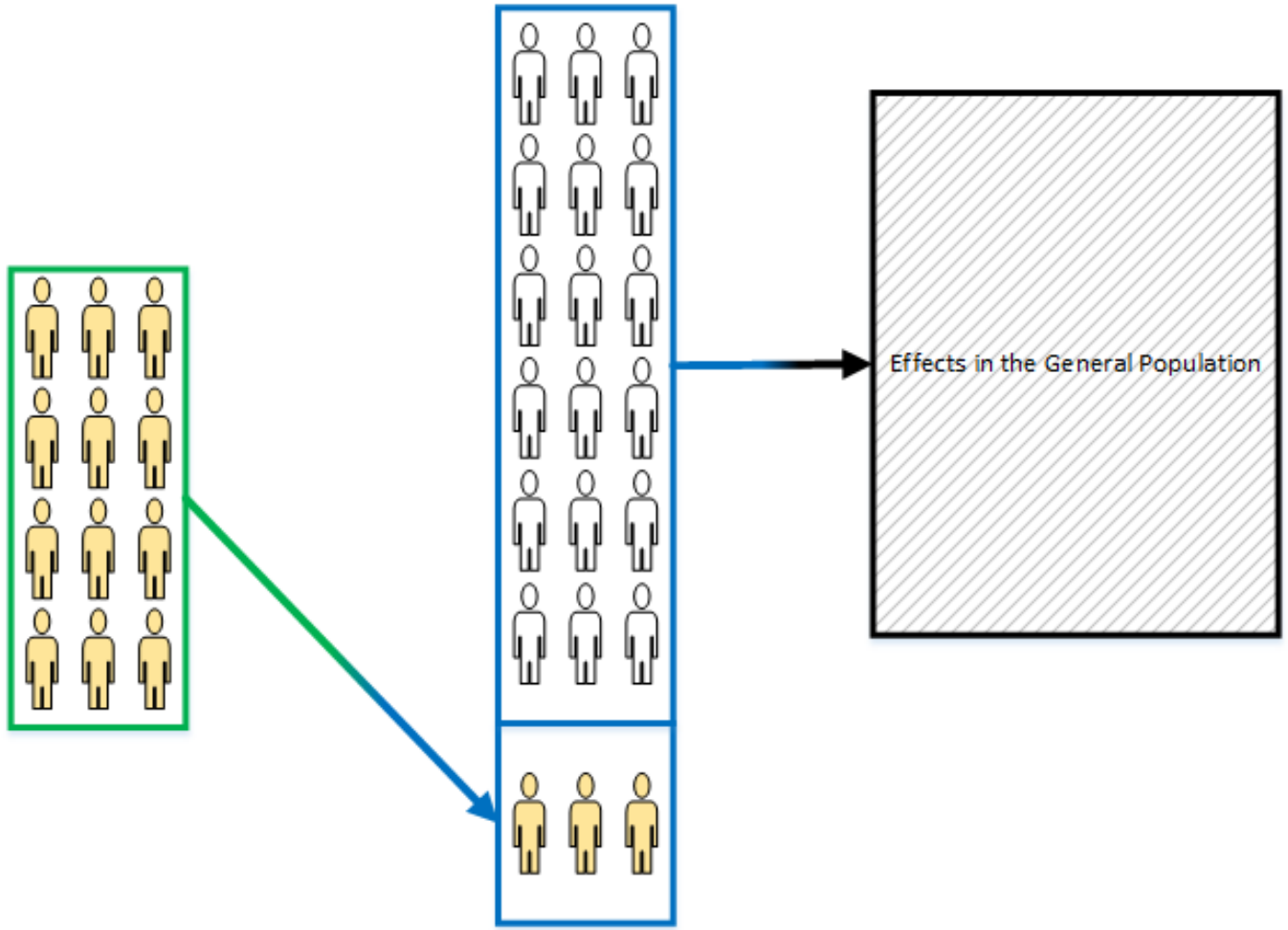
- But the smaller surveyed group is just a subset of the larger survey's group…there should be a way to combine the data for stronger inference…

- <u>Presentation Goal: Demonstrate a novel method to fuse surveys together, enriching information content for national estimates</u>



Data Fusion for Survey Surveillance

- "Data fusion" is a statistical process whereby one data source is combined or integrated with another data source
- Fusion occurs by adjusting for all the factors that would cause differences in selection between them
- Once the "green-to-blue" transition is corrected, then inference to the "black" is valid

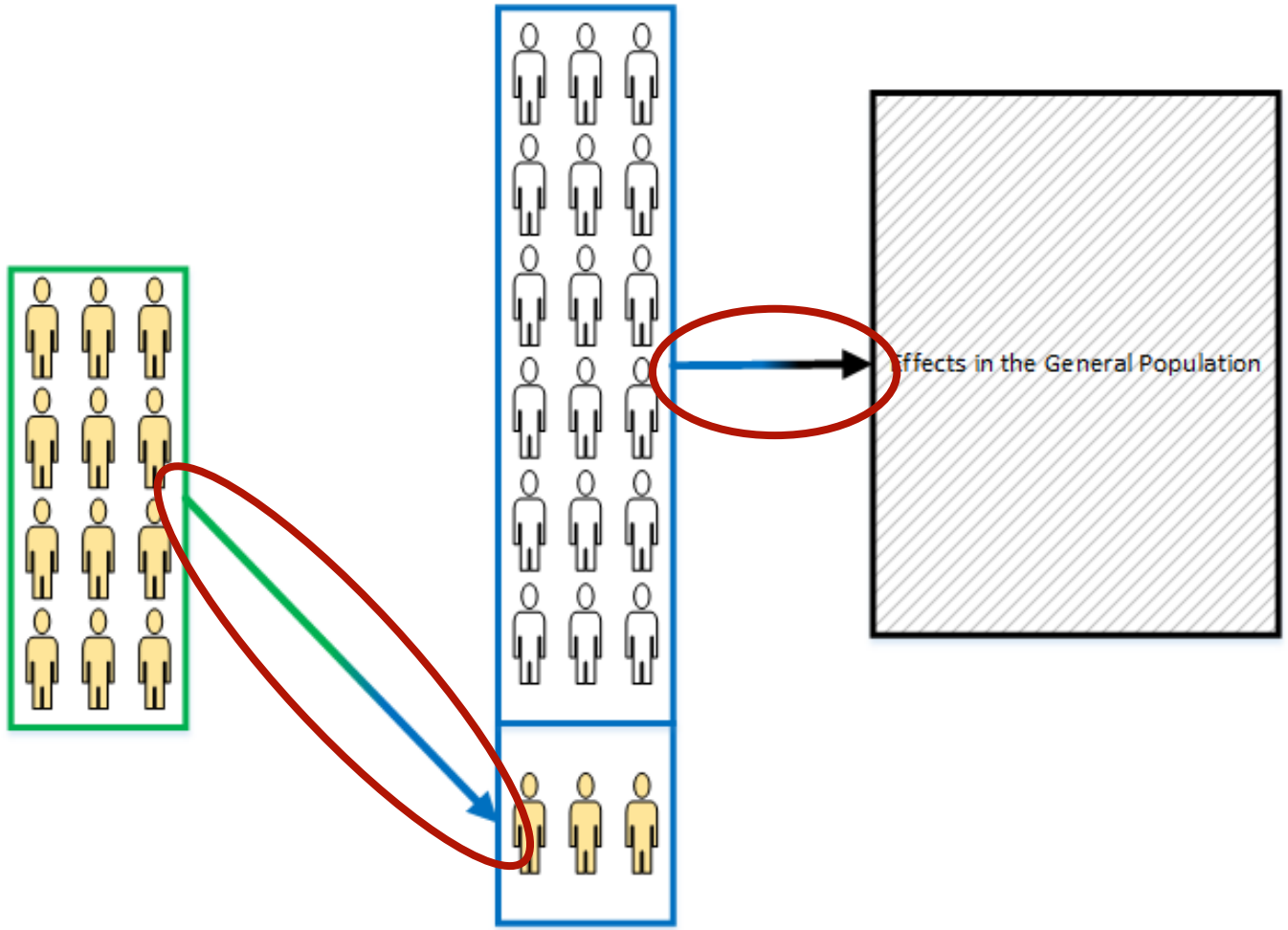


How does it work?

- We make 2 serial adjustments to the data
  - 1st: Fusion, or "transport weighting" corrects for between sample selection differences
  - 2nd: Benchmarking adjustment corrects for overarching selection bias
- Combined adjustments are represented as analytical weights, correcting for selection along the pathway

- Full statistical details and diagramming won Best Poster Award at the Society for Epidemiologic Research (SER) 2025
  - **Dr. Rockhill**, Bemis, Schow, Olsen, Beekman, Monte
  - Poster available upon request
  - Full publication forthcoming

RMPDS 19TH ANNUAL SCIENTIFIC MEETING



What Surveys are We Fusing?

RMPDS fielded a psychedelic-focused survey, which will fuse to the NMURx survey.



NSIHT

Contains dozens of detailed behavioral information not measured on anchor survey

Reason, setting, adverse events, etc



NMURx

Statistical design provides valid national estimates which are generalizable

Approach to selecting weights:

1. Find analogous sample population
2. Iterate 2,048 weighting schemes; pick one
3. Check internal consistency between samples
4. Check external validity against national benchmarks
5. Estimate novel results

1: Find analogous sample population

The small sample is biased!!

Characteristic	NSIHT Unadjusted sample… People Using Psychedelics	NMURx, weighted… People Using Psychedelics
Sex (Female)	40.4%	35.0%
Income (>$75,000)	52.4%	47.0%
Self-assessed health (Poor/Fair)	16.5%	10.5%
Tobacco use (Smoking cigarettes)	46.4%	38.6%

If we don't apply corrections, estimates would be from people who are…

- Disproportionately female
- Earn higher amounts of money
- Use more drugs
- Are less healthy (among other factors)

2 & 3: Choosing a Weighting Scheme and Internal Consistency

2,048 different transport weighting schema were tested vs anchor sample. Diamond was selected.

Residual Transport Bias is Small

Transport Bias: Average differences between anchor and weighted data

Transport weighting barely changed demographics…but they didn't need much correction anyway

- Residual transport bias: 0.09 percentage points (pps)

Health metrics shifted the most

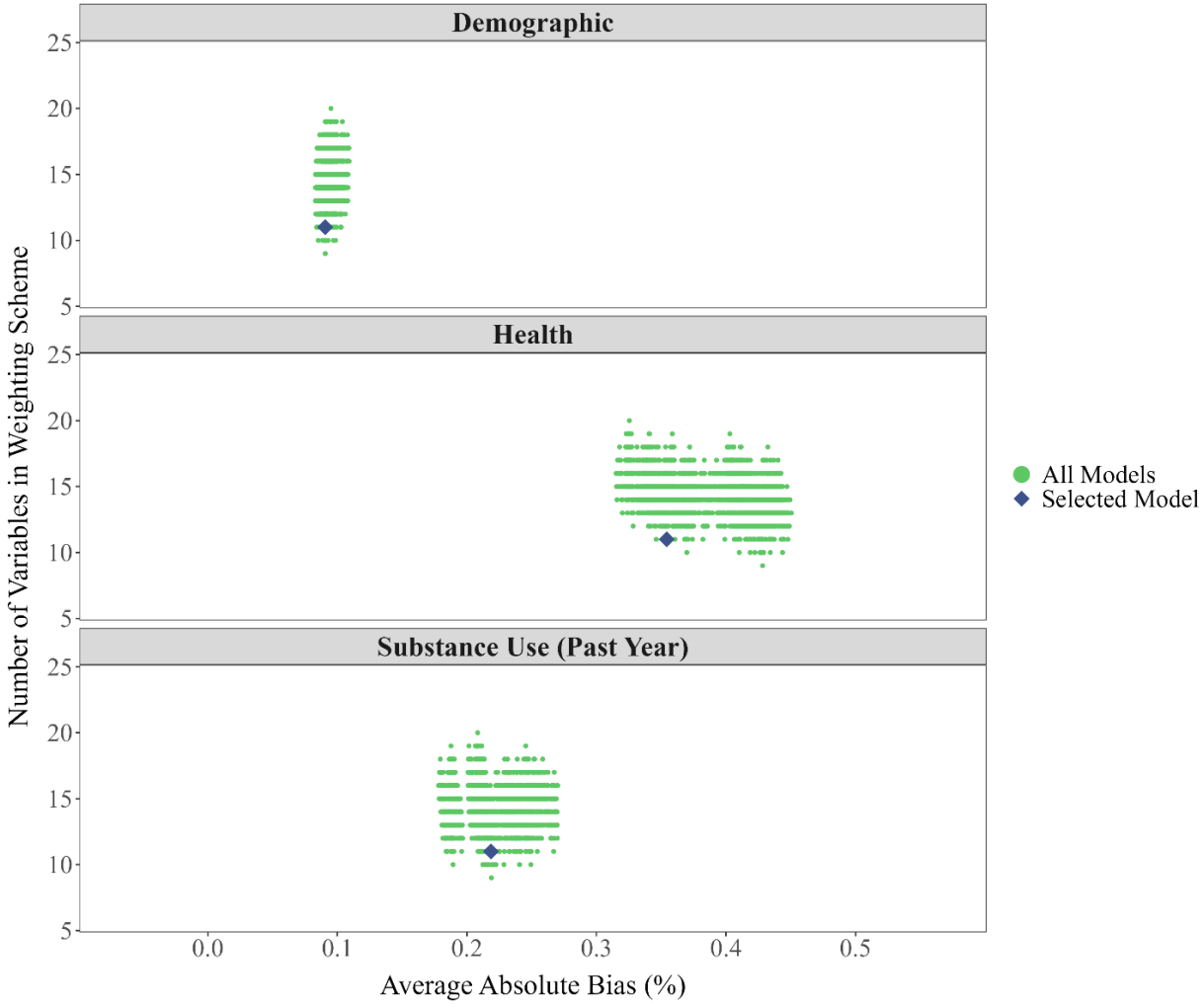
- Residual transport bias: 0.35 pps

Substance use shifted somewhat

- Residual transport bias: 0.22 pps

Multidomain weighting was needed!

- Demographics
- Health
- Substance use
- Trust in people metric



RMPDS 19TH ANNUAL SCIENTIFIC MEETING

Fused sample recreated metrics from anchor sample (<5pps on average)

4: Check External Validity

After fusion, final error got substantially better

Error in demographics mostly unchanged

- Unweighted error: 1.82 pps
- Fused error: 1.88 pps
- 3.3% increase

Error in health metrics substantially improved

- Unweighted error: 7.30 pps
- Fused error: 3.38 pps
- 53.7% decrease

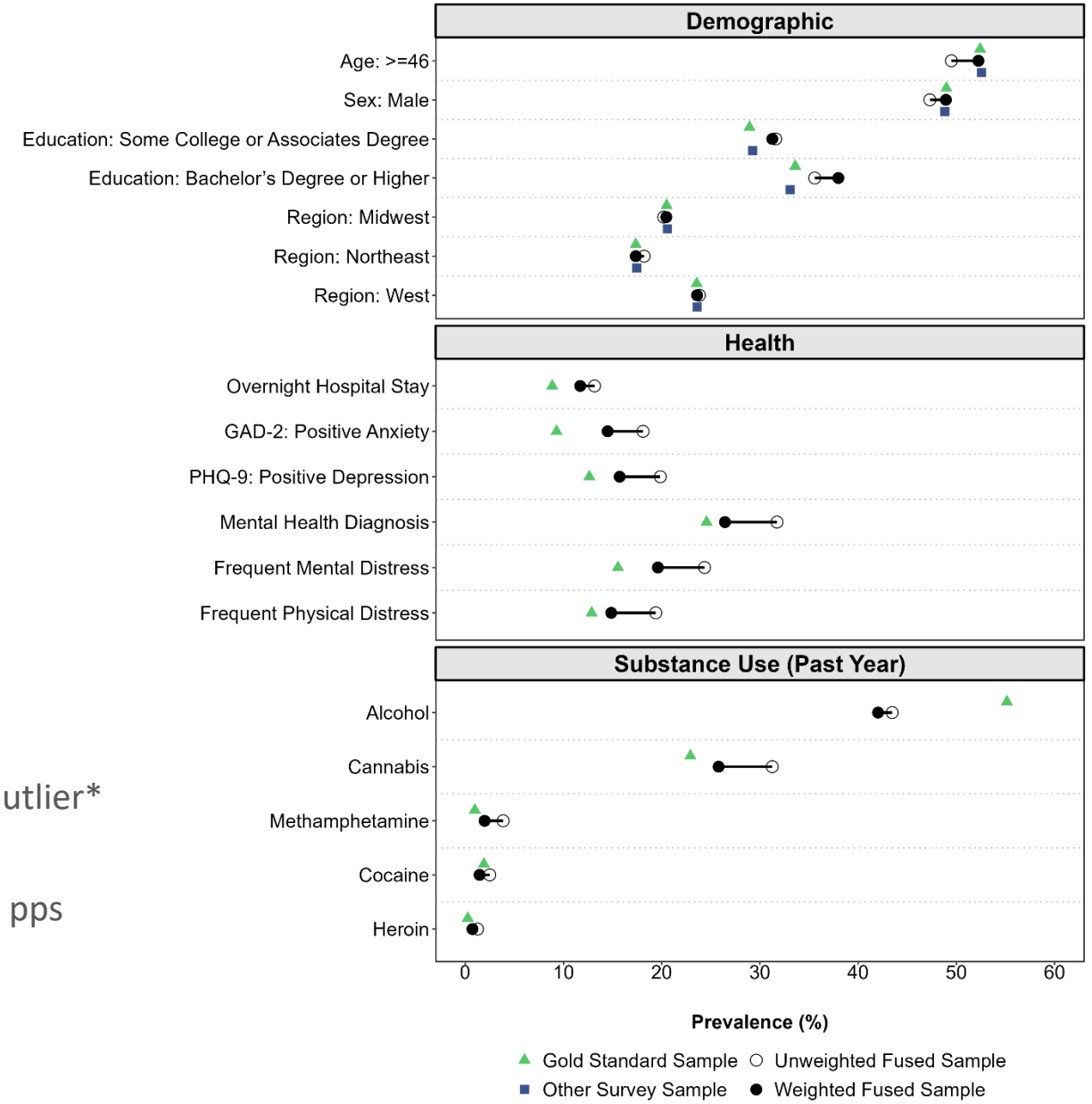
Error in substance use substantially improved, but alcohol is an outlier*

With alcohol

- Unweighted error: 6.56 pps
- Fused error: 6.03 pps
- 8.1% decrease

Without alcohol

- Unweighted error: 6.03 pps
- Fused error: 1.55 pps
- 68.3% decrease



*Why is alcohol an outlier? It is measured differently between RMPDS surveys and external benchmarks.

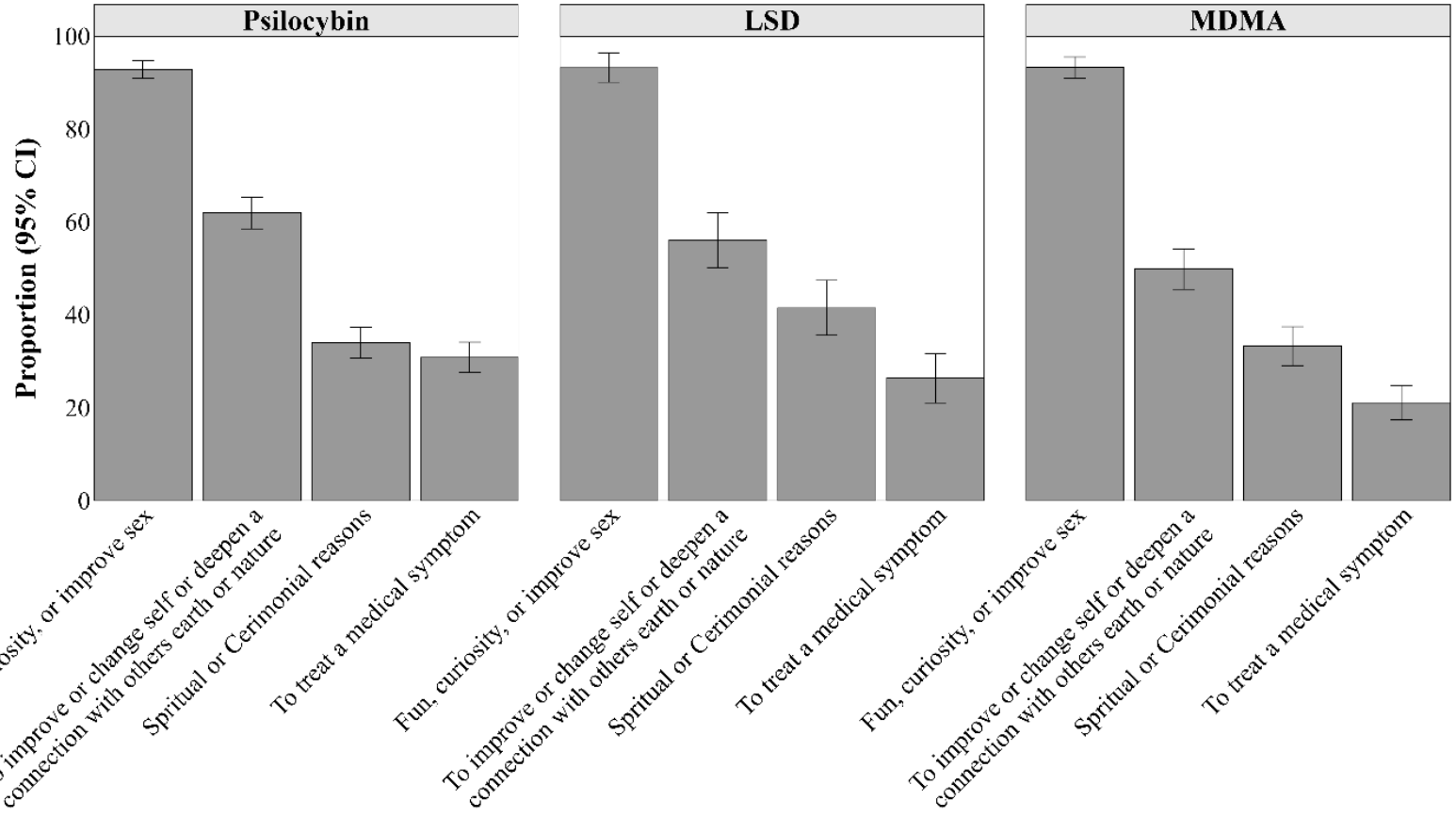
5: Estimates of Reason for Use

National Estimates for Reasons of Use

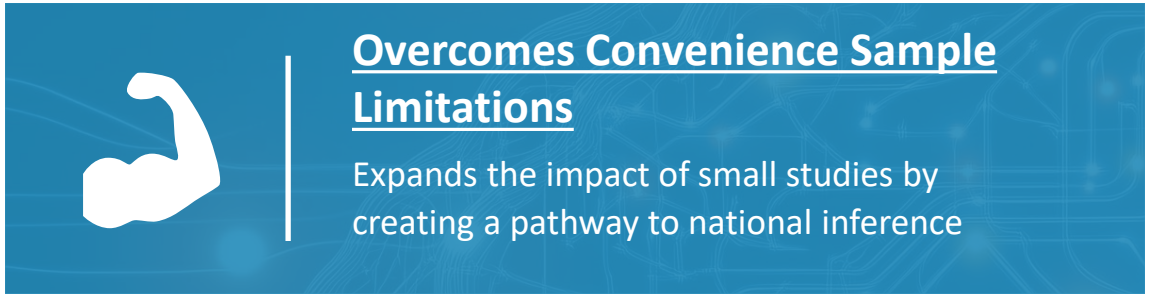
Between 50% and 60% of people who use psychedelics use to improve themselves

Only about a third see use as treating a medical symptoms

Most still consider their use as recreational

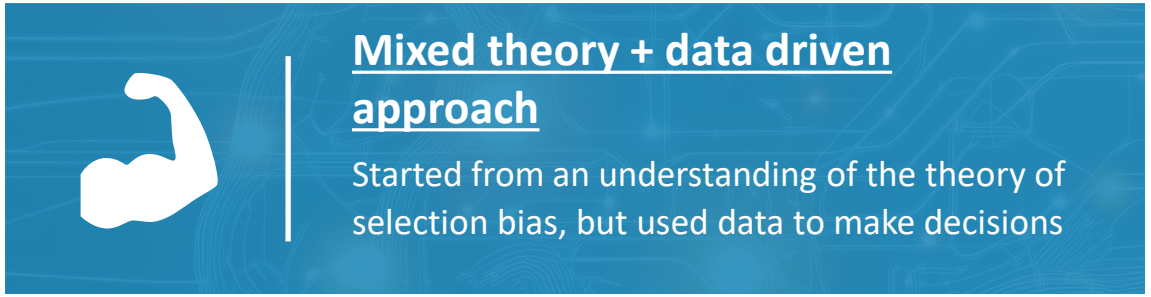


Strengths and Limitations



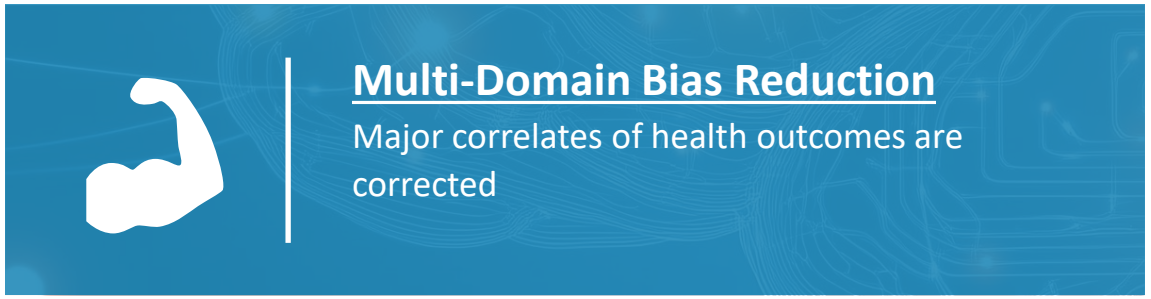
Overcomes Convenience Sample Limitations

Expands the impact of small studies by creating a pathway to national inference



Mixed theory + data driven approach

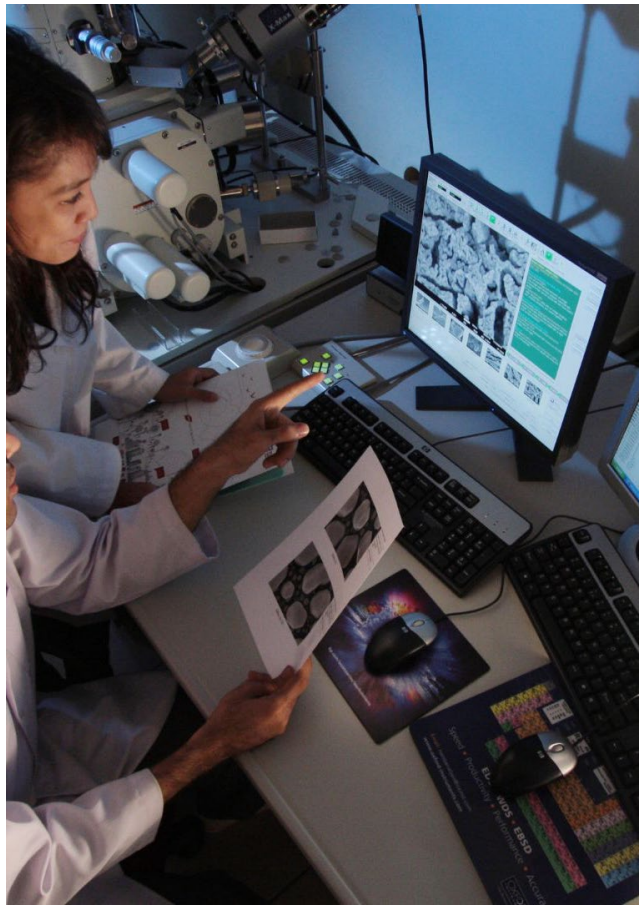
Started from an understanding of the theory of selection bias, but used data to make decisions



Multi-Domain Bias Reduction

Major correlates of health outcomes are corrected

Conclusions: The Fused Survey Study Design



Here we have demonstrated a novel statistical method to draw valid national-level inference, iterating on classic study designs.

Enriching information content

- Collecting many detailed behaviors may be infeasible
- Approach creates a path for inference

Convenience samples are limited

- Naïve samples are highly biased
- Even rigorously collected samples must contend with selection bias

Residual selection bias was small

- Compared to 4 national benchmarks
- Required multi-domain adjustments

Treating Symptoms

- A minority of adults currently using see their use as treating symptoms
- A recreational paradigm of use may predominate for years to come

RMPDS 19TH ANNUAL SCIENTIFIC MEETING

Thank you

Joshua C. Black, PhD
Senior Statistical Scientist
Rocky Mountain Poison & Drug Safety
Denver Health and Hospital Authority

Disclosure

This work was conducted by the staff at Rocky Mountain Poison and Drug Safety (RMPDS). RMPDS is a division of nonprofit Denver Health and Hospital Authority, a political subdivision of the State of Colorado, USA. Outside of this work, RMPDS provides independent surveillance, research, and reporting services to pharmaceutical manufacturers, government, and non-government agencies. No manufacturer or agency participated in the conception, analysis, drafting, or review of this presentation.

Funding

The methods work for this presentation was funded through initiatives #SAM318573 and #75S20124P00003 to investigator Andrew Monte from the Substance Abuse and Mental Health Services Administration (SAMHSA).

The views and opinions expressed are the presenters' own and do not necessarily represent the views, opinions, or policies of SAMHSA, the Department of Health and Human Services, or the United States government.

DENVER HEALTH est. 1860 FOR LIFE'S JOURNEY | ROCKY MOUNTAIN POISON & DRUG SAFETY Saving lives with answers.™ | RADARS® SYSTEM

3: Check Internal Consistency

Transport did not perturb the characteristics of the sample reporting use of psychedelics

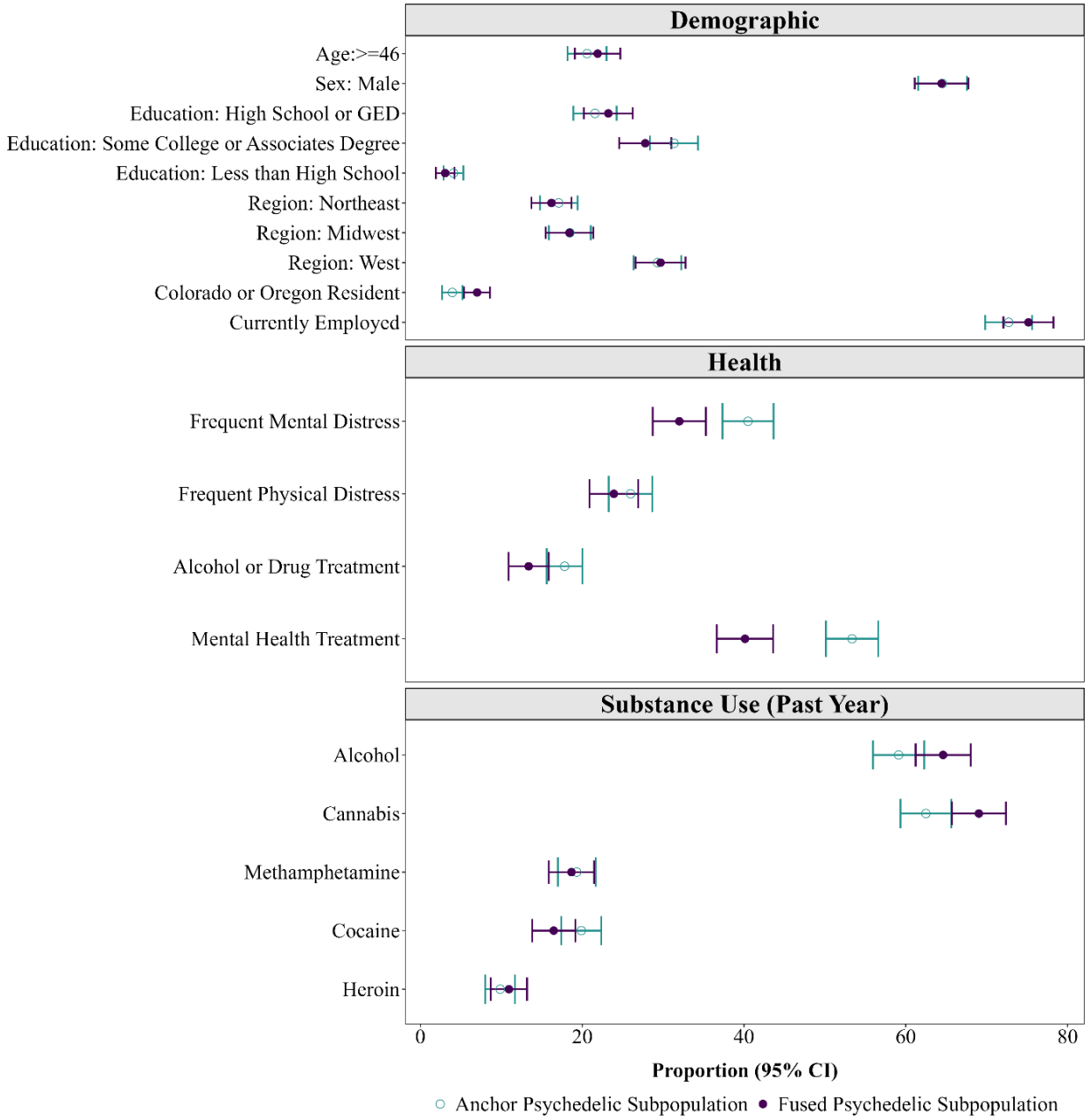
Demographics are nearly overlapping

Substance use varied slightly, but well within uncertainty

Physical distress and alcohol/drug treatment similar

Current mental stress/treatment differed modestly

- Current theories on why this difference stands out:
  - Seasonal differences (14-day mental health distress)
  - Positivity bias potentially influential (15% misaligned by drug)



Surely, this isn’t the whole story of validity…

While the naïve NSIHT sample was affected by selection bias, this data fusion methodology corrects for that bias to generate valid estimates



But what about questions (measurements)? Are the measurements themselves valid?