

Sampling, Metrics, and Denominators

Kunthel By
Mathematical Statistician
Office of Biostatistics
Center for Drug Evaluation and Research
FDA

Discussion Goals

- Measuring abuse-related outcomes in population
- Measuring change in abuse-related outcomes over time
- Impact of biased sampling

Abuse-Related Outcomes

- Abuse
- Misuse
- Addiction
- Overdose
- Death

Abuse-Related Outcomes: Route-specific

- Oral
- Chew
- Snort
- Inject
- Smoke

Quantifying Abuse-related Outcomes

- Precise quantification of abuse-related outcomes useful for:
 - Monitoring trends in population
 - Informing regulatory decisions affecting population
 - Assessing effectiveness of opioids with ADF in population

Sampling

Statements About Population: Ideal Approach

- Start with well-defined population
- Probability sample from population
- Ascertain outcomes
- Compute outcome metrics in sample
- Make statements about population:
 - Proportion of population abusing product X
 - Proportion of individuals snorting X among those who abuse X

Current Data: Less Than Ideal

- Non-probability sample
- Selection process not quantifiable
- Numerator-only
 - Case-only
 - Spontaneous

Numerator Data: Population?

Population = United States?

Population = High Risk of SUD?

Sample =
Treatment Center
Data

Current Data: Statements About Population

- Statements about underlying population difficult.
 - Proportion snorting X in sample estimates proportion snorting X in population?
- Problem: outcome-dependent sampling.

Current Data: Statements About Population

- Selection process may sample
 - Injectors of X at higher rate than snorters of X
 - Abusers of X at higher rate than abusers of Z

- Selection process not interesting:
 - Required for population inference.
 - Not available.

Current Data: Statements About Population

- Data indexed by time.
- Same problem at each time point.
- What about change over time?
- Change in proportion abusing X in sample valid for change in proportion abusing X in population?
- Depends on assumptions and metrics of change
 - Difference: Restrictive assumptions about selection process
 - Ratio: Assumptions less restrictive
 - Assumptions not verifiable

Metrics and Denominators

Current Data

- No denominator.
- Selection process unknown.
- How do we define abuse metrics (absolute/relative)?

Metrics & Denominators: Abuse Outcomes

- Abuse of product X as a *proportion* of
 - Surveyed (treatment center data)
 - Surveyed who indicate abuse of any opioid (treatment center data)
 - Poison center calls (poison center data)
 - Poison center calls with exposures to any opioid (poison center data)
 - Census population within catchment area of surveillance system

Metrics & Denominators: Route-specific Outcomes

- Route-specific abuse of X as *proportion* of
 - Surveyed
 - Surveyed who indicate abuse of X
 - Surveyed who indicate abuse of any opioid
 - Poison center calls
 - Poison center calls with exposures to any opioid
 - Poison center calls with exposures to X

Metrics & Denominators: Utilization-Adjusted

- Number of abuse of X depends on utilization of X :
 - *Prescriptions (Rx)*
 - *Dosage-units (Tx)*
 - *Unique individuals with Rx (Px)*
- Metrics based on utilization include *rate* of
 - Overall abuse of X per Rx/Tx/Px of X
 - Abuse of X via route r per Rx/Tx/Px of X
- Numerator captured by surveillance systems
- Denominator measured within catchment area.

Metrics & Denominators: Useful?

- Two broad types of abuse metrics
 - Proportions
 - Rates

- Not entirely clear which metric is most useful for
 - Monitoring trends in population
 - Regulatory decision making affecting population
 - Assessing impact of ADF in population

Metrics & Denominators: Hybrid Metric

- $$\frac{\text{no. of abuse events for product } X}{\text{no. surveyed} \times \text{utilization of } X}$$
- “Proportion adjusting for utilization”?
- “Rate adjusting for number surveyed”?

Metrics & Denominators: Change

- Compute abuse-related measures in two periods
- Relevant in context of ADF
 - Pre period: μ_0 (prevalence or rate)
 - Post period: μ_1 (prevalence or rate)
- Change
 - Difference: $\Delta = \mu_1 - \mu_0$
 - Ratio: $\Delta = \mu_1 / \mu_0$
- For some product, no pre period

Metrics & Denominators: Change

- Ideal length of periods unclear
- Long pre period
 - More info on “pre-existing” trends
 - Pre period population structure potentially different than post
- Long post period
 - More info on long-term impact of ADF
 - Post period population structure potentially different than pre
- Long pre and post periods
 - More susceptible to changes in biased sampling



Discussion

Question 1: Analyses Under Biased Sampling

- Discuss analytical approaches that enable inference about the underlying population without
 - requiring knowledge of the selection process, or
 - without making assumptions about it.
- Discuss the utility of making assumptions about the selection process. Discuss assumptions that may be considered reasonable.

Question 2: Temporal Sample Heterogeneity

- Discuss methodological approaches that address changes in the studied population over time (for example, changes in individual geography, demographics, ...)

Question 3: Numerator Data & Abuse Metrics

- Discuss the usefulness of these metrics for measuring and assessing the impact of ADFs on abuse-related outcomes in the population:
 - Number abusing X as a proportion of 1) surveyed; 2) surveyed and who indicate abuse of any opioid; 3) calls to poison centers; 4) calls to poison centers with exposures to any opioids; 5) census population within catchment area of surveillance system
 - Number abusing X via route R as a proportion of 1) surveyed; 2) surveyed and who indicate abuse of any opioid; 3) surveyed who indicate abuse of X; 4) calls to poison centers; 4) calls to poison centers with exposures to any opioids; 5) calls to poison centers with exposures to X;
 - Number abusing X relative to the 1) number of prescriptions; 2) number of dosage-units; 3) number of unique individuals with prescriptions
- Discuss metrics we have not considered that are potentially useful.
- Discuss interpretations when different metrics imply different conclusions.

Question 4: Catchment Area

- Outcome metrics based on utilization use prescriptions dispensed or dosage-units dispensed. For treatment center data, it's not entirely clear how to define the catchment area within which prescriptions or dosage-units are counted. Possibilities we have considered include:
 - a) States in which treatment centers reside
 - b) 3-digit ZIP codes of residence of individuals who interact with treatment centers
 - c) Different catchment area for each time period based on b) versus a single catchment area for both time periods using **all** 3-digit ZIP codes of individuals assessed during the study period
- Discuss advantages and disadvantages of these catchment area definitions.
- Discuss other catchment area definitions that you think are useful.

Question 5: Length of Pre/Post Periods

- Pre-post comparisons have been considered extensively in the context of measuring change between a pre-ADF period and a post-ADF period. Discuss criteria that may be useful for determining the length of the pre and post period. Discuss the balance between the ability to observe trends and changing population characteristics.

Question 6: Hybrid Metric

- For treatment center data, we mentioned the metric

$$\frac{\text{no. of abuse events for product } X}{\text{no. surveyed} \times \text{utilization of product } X}$$

- Discuss the merits of such a metric.

Back-up

Utilization-Adjusted Rates Under Biased Sampling

Period	Sample	Population
0 (pre)	$r_0 = \frac{\pi_0 X_0}{U_0}$	$R_0 = \frac{X_0}{U_0}$
1 (post)	$r_1 = \frac{\pi_1 X_1}{U_1}$	$R_1 = \frac{X_1}{U_1}$
Rate Difference	$r_1 - r_0 \neq R_1 - R_0$ unless $\pi_j = 1; j = 0,1$	
Rate Ratio	$\frac{r_1}{r_0} \neq \frac{R_1}{R_0}$ unless $\pi_1 = \pi_0$	

Proportion/Prevalence Under Biased Sampling

- Y_x denotes abuse of product X (0 = no; 1 = yes)
- T denotes time period (0 = pre; 1 = post)
- S denote selection (0 = not sampled; 1 = sampled)
- $\mu(t) = \Pr(Y_x = 1|T = t)$ denotes prevalence of abuse of X in period t in the **population**
- $\nu(t) = \Pr(Y_x = 1|S = 1, T = t)$ denotes prevalence of abuse of X in period t in the **sample**

$$\nu(t) = \left[\frac{\Pr(S = 1|Y_x = 1, T = t)}{\Pr(S = 1|T = t)} \right] \Pr(Y_x = 1|T = t) = \frac{\pi_x(t)}{\pi(t)} \mu(t)$$

Proportion/Prevalence Under Biased Sampling

Period	Sample	Population
0 (Pre)	$v(0)$	$\mu(0)$
1 (Post)	$v(1)$	$\mu(1)$
Prevalence Difference	$v(1) - v(0) \neq \mu(1) - \mu(0)$ unless $S \perp Y_x$ or equivalently $\pi_x(t) = \pi(t); t = 0,1$	
Prevalence Ratio	$\frac{v(1)}{v(0)} \neq \frac{\mu(1)}{\mu(0)}$ unless (1) $S \perp Y_x$, or (2) $\pi_x(1)/\pi(1) = \pi_x(0)/\pi(0)$	

