Transportability Index

A Scalar Summary of Transportation Robustness

Xinran Miao

University of Wisconsin-Madison

Joint work with Jiwei Zhao and Hyunseung Kang

ENAR 2024

Background

> "Science is about generalization, and generalization requires that conclusions obtained in the laboratory be transported and applied elsewhere" (Pearl and Bareinboim 2014)

Background

> "Science is about generalization, and generalization requires that conclusions obtained in the laboratory be transported and applied elsewhere" (Pearl and Bareinboim 2014)

	R	Covariate $\mathbf{X}$	Outcome Y
Source (n)	1	$\mathbf{x}_1$	y_1
	$\vdots$	$\vdots$	$\vdots$
	1	$\mathbf{x}_n$	y_n
Target (m)	0	$\mathbf{x}_{n+1}$	
	$\vdots$	$\vdots$	
	0	$\mathbf{x}_{n+m}$	

Background

> "Science is about generalization, and generalization requires that conclusions obtained in the laboratory be transported and applied elsewhere" (Pearl and Bareinboim 2014)

	R	Covariate $\mathbf{X}$	Outcome Y
Source (n)	1	$\mathbf{x}_1$	y_1
	$\vdots$	$\vdots$	$\vdots$
	1	$\mathbf{x}_n$	y_n
Target (m)	0	$\mathbf{x}_{n+1}$	
	$\vdots$	$\vdots$	
	0	$\mathbf{x}_{n+m}$	

Goal

Estimate the mean outcome on target: $\beta = \mathrm{E}(Y \mid R = 0)$.

Popular Assumption: Transportability

Transportability assumption between source ($R = 1$) and target ($R = 0$)

To identify $\beta = \mathrm{E}(Y \mid R = 0)$, it's common to assume "transportability",

$$p(y \mid \mathbf{x}, R = 0) = p(y \mid \mathbf{x}, R = 1), \tag{1}$$

which allows for covariate shift

$$p(\mathbf{x} \mid R = 0) \not\equiv p(\mathbf{x} \mid R = 1).$$

Popular Assumption: Transportability

Transportability assumption between source ($R = 1$) and target ($R = 0$)

To identify $\beta = \mathrm{E}(Y \mid R = 0)$, it's common to assume "transportability",

$$p(y \mid \mathbf{x}, R = 0) = p(y \mid \mathbf{x}, R = 1), \tag{1}$$

which allows for covariate shift

$$p(\mathbf{x} \mid R = 0) \not\equiv p(\mathbf{x} \mid R = 1).$$

Under (1),

$$\begin{aligned} \beta &= \mathrm{E}(Y \mid R = 0) \\ &= \mathrm{E}\left\{ \boxed{\mathrm{E}(Y \mid \mathbf{X}, R = 0)} \mid R = 0 \right\} \\ &= \mathrm{E}\left\{ \boxed{\mathrm{E}(Y \mid \mathbf{X}, R = 1)} \mid R = 0 \right\} \end{aligned}$$

can be identified.

Popular Assumption: Transportability

Transportability assumption between source ($R = 1$) and target ($R = 0$)

To identify $\beta = \mathrm{E}(Y \mid R = 0)$, it's common to assume "transportability",

$$p(y \mid \mathbf{x}, R = 0) = p(y \mid \mathbf{x}, R = 1), \tag{1}$$

which allows for covariate shift

$$p(\mathbf{x} \mid R = 0) \neq p(\mathbf{x} \mid R = 1).$$

Under (1),

$$\begin{aligned}
\beta &= \mathrm{E}(Y \mid R = 0) \\
&= \mathrm{E}\left\{ \boxed{\mathrm{E}(Y \mid \mathbf{X}, R = 0)} \mid R = 0 \right\} \\
&= \mathrm{E}\left\{ \boxed{\mathrm{E}(Y \mid \mathbf{X}, R = 1)} \mid R = 0 \right\}
\end{aligned}$$

can be identified.

> Unfortunately, assumption (1) cannot be verified from data.

Outline

1 **Transportability Index and Its Estimation**

2 Data Application on Transportations between ICUs

Sensitivity Model

To quantify the violations of the transportability assumption, we adopt the selection odds model with an exponential tilting shift (Robins 2000; AlexanderM. Franks and Feller 2020; Scharfstein et al. 2021; Dahabreh et al. 2022),

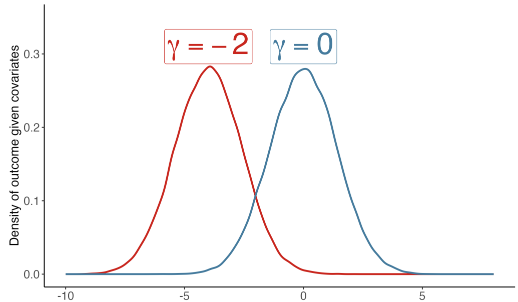
$$\underbrace{\boxed{p(y \mid \mathbf{x}, R = 0)}}_{\text{target}} \propto \exp(\gamma y) \underbrace{\boxed{p(y \mid \mathbf{x}, R = 1)}}_{\text{source}}. \tag{2}$$

Sensitivity Model

To quantify the violations of the transportability assumption, we adopt the selection odds model with an exponential tilting shift (Robins 2000; AlexanderM. Franks and Feller 2020; Scharfstein et al. 2021; Dahabreh et al. 2022),

$$\underbrace{\boxed{p(y \mid \mathbf{x}, R=0)}}_{\text{target}} \propto \exp(\gamma y) \underbrace{\boxed{p(y \mid \mathbf{x}, R=1)}}_{\text{source}}. \tag{2}$$

- When $\gamma = 0$, (2) reduces to the transportability assumption (1).
- When $\gamma \neq 0$, $p(y \mid \mathbf{x}, R=0)$ deviates from $p(y \mid \mathbf{x}, R=1)$. The magnitude of deviation is calibrated by γ.
- For example, suppose $p(y \mid \mathbf{x}, R=1) = \text{Normal}(0, \sigma^2)$, then $p(y \mid \mathbf{x}, R=0) = \text{Normal}(\gamma\sigma^2, \sigma^2)$.



Sensitivity Model

To quantify the violations of the transportability assumption, we adopt the selection odds model with an exponential tilting shift (Robins 2000; AlexanderM. Franks and Feller 2020; Scharfstein et al. 2021; Dahabreh et al. 2022),

$$\underbrace{\boxed{p(y \mid \mathbf{x}, R = 0)}}_{\text{target}} \propto \exp(\gamma y) \underbrace{\boxed{p(y \mid \mathbf{x}, R = 1)}}_{\text{source}}. \tag{2}$$

- Under (2), the target outcome mean $\beta = \mathrm{E}(Y \mid R = 0)$ becomes a function of γ,

$$\beta(\gamma) = \mathrm{E}\left[\frac{\mathrm{E}\{Y \exp(\gamma Y) \mid \mathbf{X}, R = 1\}}{\mathrm{E}\{\exp(\gamma Y) \mid \mathbf{X}, R = 1\}} \middle| R = 0 \right].$$

- When $\gamma = 0$,

$$\beta(0) = \mathrm{E}\{\mathrm{E}(Y \mid \mathbf{X}, R = 1) \mid R = 0\},$$

which reduces to the case when the transportability assumption (1) holds.

Transportatility Index

Definition 1 (Transportability Index)

We define transportability index as the derivative of $\beta(\gamma)$ *with respect to* γ *evaluated at* $\gamma = 0$*:*

$$\lambda = \left. \frac{\partial \beta(\gamma)}{\partial \gamma} \right|_{\gamma=0}.$$

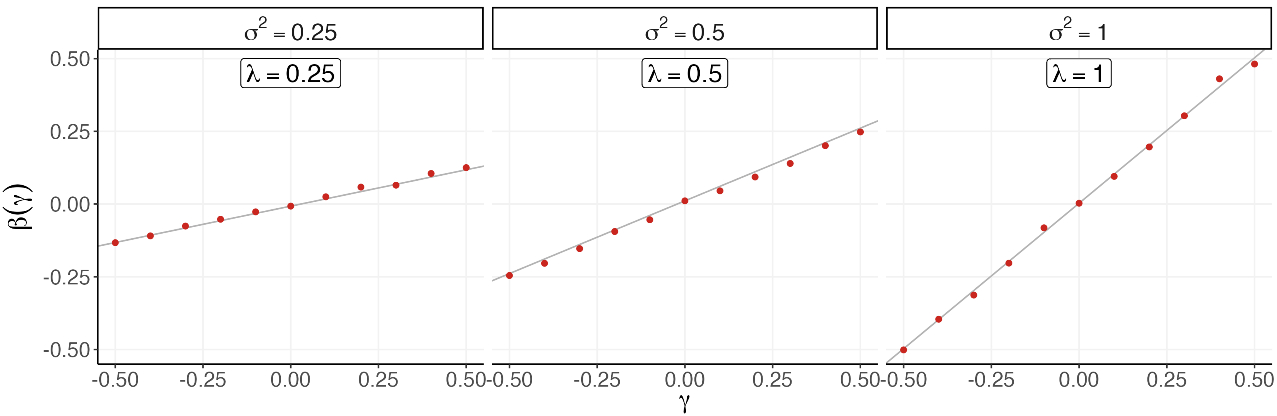
- It quantifies how a small change in γ (near zero) influences the target estimand.
- When $|\lambda| \approx 0$, a slight deviation from the transportability assumption does not change the estimand.
- When $|\lambda|$ is far from zero, a slight deviation from the transportability assumption leads to a dramatic change in the estimand.

Definition 1 (Transportability Index)

We define transportability index as the derivative of $\beta(\gamma)$ with respect to γ evaluated at $\gamma = 0$:

$$\lambda = \left.\frac{\partial \beta(\gamma)}{\partial \gamma}\right|_{\gamma=0}.$$

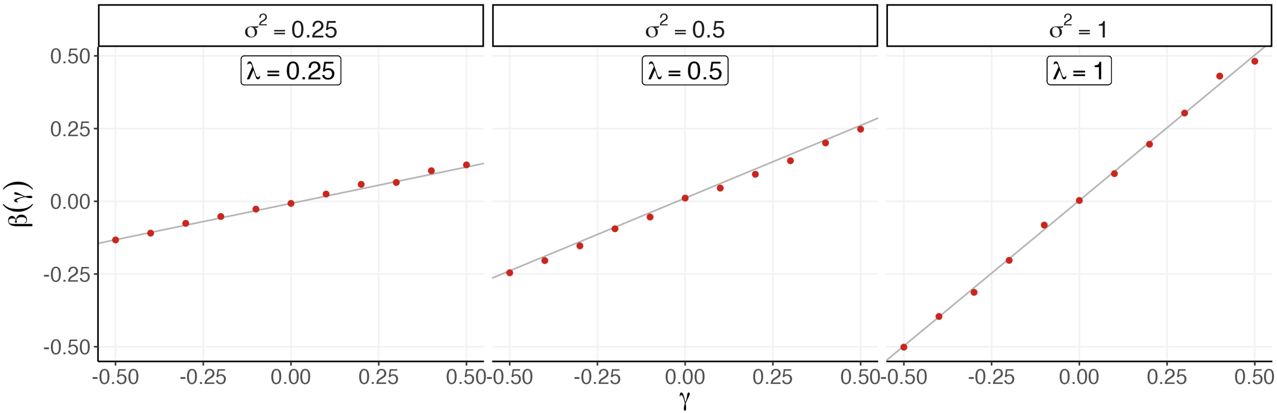
Revisiting the normal example, $p(y \mid \mathbf{x}, R = 1) = \text{Normal}(0, \sigma^2)$.



Transportatility Index

- More heterogeneity in the source population $\Rightarrow$ more sensitive to violations of the transportability assumption.
- It matches existing intuition about ease of generalizability when the treatment effect is homogeneous (Tipton and Olsen 2018).

Revisiting the normal example, $p(y \mid \mathbf{x}, R = 1) = \text{Normal}(0, \sigma^2)$.



Transportability Index

$$\lambda = \left.\frac{\partial \beta(\gamma)}{\partial \gamma}\right|_{\gamma=0}$$

Relations to existing literature

- Relates to the influence curve in robust statistics (Huber 1964; Hampel 1974), but at an estimand level with respect to the sensitivity parameter γ.
- Relates to the sensitivity analysis (Robins 2000; AlexanderM. Franks and Feller 2020), providing a scalar summary of the sensitivity to violations of the transportability assumption.

Universiality of the transportability index

- Extends beyond the exponential tilting function $\exp(\gamma y) \Rightarrow$ any function $\rho(y, \mathbf{x}; \gamma)$ with $\rho(y, \mathbf{x}; 0) = 1$.
- Extends beyond the outcome mean $\Rightarrow$ any parameter β defined through estimating equations (e.g., median, GLM, GMM, GEE).

For simplicity, this talk focuses on the transportability index for the outcome mean.

Transportability Index for Outcome Mean

For the mean of the outcome in the target population, the transportability index simplifies to

$$\lambda = \mathrm{E}\{w(\mathbf{X})\mathrm{var}(Y \mid \mathbf{X}, R = 1) \mid R = 1\},$$

where $w(\mathbf{x}) = p(\mathbf{x} \mid R = 0)/p(\mathbf{x} \mid R = 1)$.

The transportability index depends on

- $\mathrm{var}(Y \mid \mathbf{x}, R = 1)$, the heteroskedasticitiy of the outcome variance in the source population, and
- $w(\mathbf{x})$, the covariate shift between the source and the target populations.

Estimation

Motivated by $\lambda = \mathrm{E}\{w(\mathbf{X})\mathrm{var}(Y \mid \mathbf{X}, R = 1) \mid R = 1\}$, we propose to estimate λ by

$$\widehat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} \widehat{w}(\mathbf{x}_i)\{y_i - \widehat{\mu}(\mathbf{x}_i)\}^2, \text{ where} \tag{3}$$

- $i = 1, \cdots, n$ indexes source samples,
- $\widehat{w}(\mathbf{x}) = n \cdot \widehat{\mathrm{pr}}(R = 0 \mid \mathbf{x})/\{m \cdot \widehat{\mathrm{pr}}(R = 1 \mid \mathbf{x})\}$, and
- $\widehat{\mu}(\mathbf{x}) = \widehat{\mathrm{E}}(Y \mid \mathbf{x}, R = 1)$ is the outcome regression fitted from the source sample.

Outline

Application on Robustness Diagnostics

MIMIC-III database contains hospital admission information of adult patients admitted to five types of critical care units between 2001 and 2012 (Johnson et al. 2016).

Source indicator R: initial care unit

- Source: Medical Intensive Care Unit ($n = 14,824$).
- Target:
  - Cardiac Surgery Recovery Unit ($m = 7,865$),
  - Trauma Surgical Intensive Care Unit ($m = 4,727$).

Outcome Y

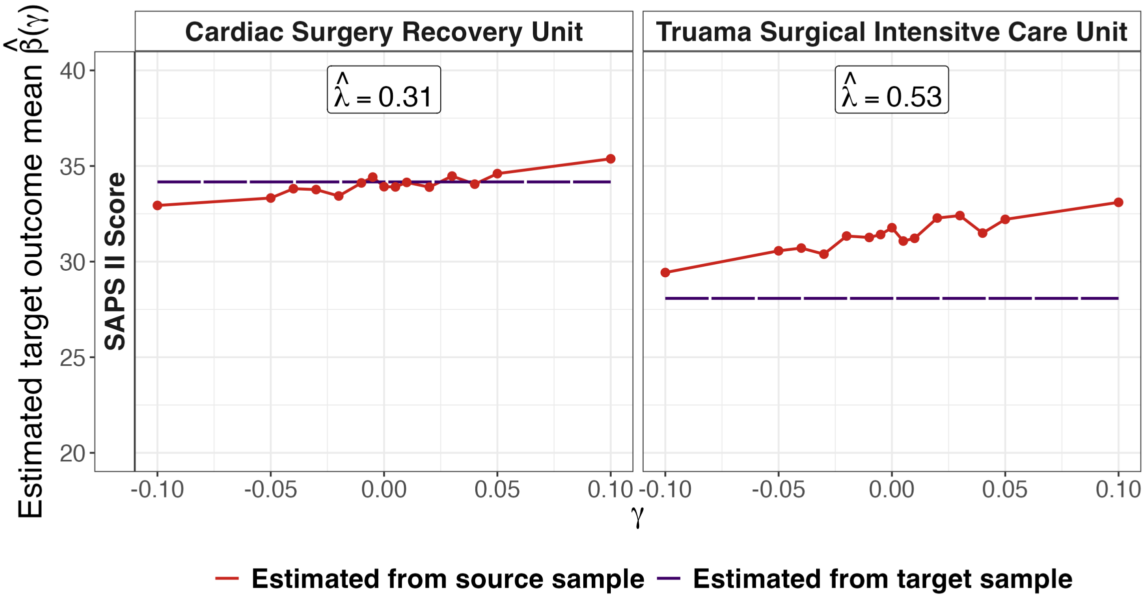
- SAPS II score: Simplified Acute Physiology Score
- It ranges from 0 to 163. The higher, the worse.

Goal: estimate the average SAPS II score in a target ICU

Covariate X

- 14 variables including demographics, chart events (e.g., body temperature) and laboratory tests (e.g., red blood cell count).

Transportability Across Critical Care Units



Summary

- This talk addresses the question of how sensitive an estimand is to violations of the transportability assumption.
- We propose a scalar summary, the transportability index, measuring the change in the estimand with respect to a small perturbation to the transportatibility assumption,

$$\lambda = \left. \frac{\partial \beta(\gamma)}{\partial \gamma} \right|_{\gamma=0}.$$

- Investigators can use this tool to diagnose whether their estimation problem is robust to violations of the transportability assumption.
- Future work: quantifying robustness of different estimands/policies with respect to the transportability assumption.

THANK YOU!

xinran.miao@wisc.edu

https://xinranmiao.github.io

References I

- AlexanderM. Franks, Alexander D'Amour and Avi Feller (2020). **"Flexible Sensitivity Analysis for Observational Studies Without Observable Implications".** In: *Journal of the American Statistical Association* 115.532, pp. 1730–1746.
- Carroll, Raymond J et al. (1997). **"Generalized partially linear single-index models".** In: *Journal of the American Statistical Association* 92.438, pp. 477–489.
- Dahabreh, Issa J et al. (2022). **"Global sensitivity analysis for studies extending inferences from a randomized trial to a target population".** In: *arXiv preprint arXiv:2207.09982.*
- Hampel, Frank R (1974). **"The influence curve and its role in robust estimation".** In: *Journal of the american statistical association* 69.346, pp. 383–393.
- Huber, Peter J. (1964). **"Robust Estimation of a Location Parameter".** In: *The Annals of Mathematical Statistics* 35.1, pp. 73–101.
- Johnson, Alistair EW et al. (2016). **"MIMIC-III, a freely accessible critical care database".** In: *Scientific Data* 3.1, pp. 1–9.
- Pearl, Judea and Elias Bareinboim (2014). **"External Validity: From Do-Calculus to Transportability Across Populations".** In: *Statistical Science* 29.4, pp. 579–595.
- Robins, James M (2000). **"Robust estimation in sequentially ignorable missing data and causal inference models".** In: *Proceedings of the American Statistical Association.* Vol. 1999. Indianapolis, IN, pp. 6–10.
- Scharfstein, Daniel O et al. (2021). **"Semiparametric sensitivity analysis: Unmeasured confounding in observational studies".** In: *arXiv preprint arXiv:2104.08300.*

References II

Tipton, Elizabeth and Robert B Olsen (2018). **"A review of statistical methods for generalizing from evaluations of educational interventions".** In: *Educational Researcher* 47.8, pp. 516–524.

Exponential Tilting Function

- Under the exponential tilting shift,

$$\underbrace{\boxed{p(y \mid \mathbf{x}, R = 0)}}_{\text{target}} \propto \exp(\gamma y) \underbrace{\boxed{p(y \mid \mathbf{x}, R = 1)}}_{\text{source}}, \tag{2}$$

  one can write

$$\mathrm{pr}(R = 1 \mid y, \mathbf{x}) = \frac{1}{1 + \exp\{\gamma y + h(\mathbf{x})\}}, \tag{4}$$

  where $h(\mathbf{x}) = \log \left(\frac{p(\mathbf{x} \mid R = 0)}{p(\mathbf{x} \mid R = 1)\mathrm{E}\{\exp(\gamma Y) Y \mid \mathbf{x}, R = 1\}} \right)$ is a function of $\mathbf{x}$ only.

- From (4), the selection probability of being in the source population is related to $(y, \mathbf{x})$ via a partially linear logistic regression model (Carroll et al. 1997).

Transportability Index for Parameters in Estimating Equations

- Consider a q-dimensional parameter of interest $\boldsymbol{\beta}$ defined through

$$\mathrm{E}\{\boldsymbol{\xi}(Y, \mathbf{X}; \boldsymbol{\beta}) \mid R = 0\} = 0.$$

- To study the sensitivity of the estimand to the transportability assumption (1), we assume,

$$p(y \mid \mathbf{x}, R = 0) \propto \rho(y, \mathbf{x}; \gamma) p(y \mid \mathbf{x}, R = 1), \text{ where}$$

$\rho(y, \mathbf{x}; \gamma)$ is a user-defined sensitivity function with $\rho(y, \mathbf{x}; \gamma) = 0$, e.g., $\rho(y, \mathbf{x}; \gamma) = \exp(\gamma y)$.

- The estimand on the target is now coded as $\beta(\gamma)$, which is the solution to

$$\mathrm{E}\left[\frac{\mathrm{E}\{\boldsymbol{\xi}(Y, \mathbf{X}; \boldsymbol{\beta}(\gamma)) \rho(Y, \mathbf{X}; \gamma) \mid \mathbf{X}, R = 1\}}{\mathrm{E}\{\rho(Y, \mathbf{X}; \gamma) \mid \mathbf{X}, R = 1\}} \middle| R = 0 \right] = \mathbf{0}.$$

- The transportability index is defined as the derivative of $\boldsymbol{\beta}(\gamma)$ with respect to γ evaluated at $\gamma = 0$:

$$\boldsymbol{\lambda} = \left. \frac{\partial \boldsymbol{\beta}(\gamma)}{\partial \gamma} \right|_{\gamma=0} = -\mathbf{M}^{-1} \mathrm{E}\left\{ \mathrm{Cov}\left(\boldsymbol{\xi}, \frac{\partial \rho}{\partial \gamma}|_{\gamma=0} \mid \mathbf{X}, R = 1 \right) \mid R = 0 \right\},$$

where $\mathbf{M} = \mathrm{E}\left\{ \mathrm{E}\left(\frac{\partial \boldsymbol{\xi}}{\partial \boldsymbol{\beta}} \mid \mathbf{X}, R = 1 \right) \mid R = 0 \right\}$ is assumed invertible.

Estimating $\beta(\gamma)$

Noticing that

$$\beta(\gamma) = \mathrm{E}\left[\mathbf{w}(\mathbf{X}) \frac{\mathrm{E}\{Y \exp(\gamma Y) \mid \mathbf{X}, R = 1\}}{\mathrm{E}\{\exp(\gamma Y) \mid \mathbf{X}, R = 1\}} \middle| R = 1\right],$$

we propose to estimate $\beta(\gamma)$ with

$$\widehat{\beta}(\gamma) = \frac{1}{n} \sum_{i=1}^{n} \widehat{w}(\mathbf{x}_i) \frac{\exp(\gamma y_i) y_i}{\widehat{\mathrm{E}}\{\exp(\gamma Y) \mid \mathbf{x}_i, R = 1\}},$$

where we recall $w(\mathbf{x}) = p(\mathbf{x} \mid R = 0)/p(\mathbf{x} \mid R = 1)$ can be estimated by $n \cdot \widehat{\mathrm{pr}}(R = 0 \mid \mathbf{x})/\{m \cdot \widehat{\mathrm{pr}}(R = 1 \mid \mathbf{x})\}$.

MIMIC III Data Application

Table. Covariate description in MIMIC III data

Type	Name	Description
Demographics	`age`	Age of a patient
	`gender`	Gender of a patient
Chart events	`diasbp_mean`	Diastolic blood pressure (on average)
	`glucose_mean`	Blood glucose (on average)
	`resprate_mean`	Respiratory rate per minute (on average)
	`sysbp_mean`	Systolic blood pressure (on average)
	`temp_mean`	Body temperature (on average)
	`hr_mean`	Heart rate per minute (on average)
Laboratory Tests	`hemotocrit_mean`	Hematocrit level (on average)
	`platelets_mean`	Platelets count (on average)
	`redbloodcell_mean`	Red blood cell count (on average)
	`whitebloodcell_mean`	White blood cell count (on average)
	`urea_n_mean`	Blood urea nitrogen (on average)
	`calcium_mean`	Calcium level in blood (on average)