MRC | Biostatistics Unit

Demystifying causal inference in randomised trials

Lecture 2: Non-compliance using Inverse Probability Weighting

ISCB 2016

Ian White
MRC Biostatistics Unit, Cambridge, UK

MRC | Medical Research Council

Plan

1. IPW for missing data
2. IPW for artificially missing data
3. Introduction to IPCW (= IPW for longitudinal data)
4. Assumptions in IPCW
5. Modelling switching
6. Estimating and using the weights
7. Illustration using SANAD trial
8. Complexities - for future reference
9. Summary

MRC | Medical Research Council

2

Learning objectives

By the end of this session, you should be able to

- implement IPW in a cross-sectional setting
- understand the assumptions of IPW and IPCW
- understand the steps in implementing IPCW

MRC | Medical Research Council

3

Plan

1. **IPW for missing data**
   - picking up missing data idea from lecture 1
2. IPW for artificially missing data
3. Introduction to IPCW (= IPW for longitudinal data)
4. Assumptions in IPCW
5. Modelling switching
6. Estimating and using the weights
7. Illustration using SANAD trial
8. Complexities - for future reference
9. Summary

MRC | Medical Research Council

4

Introduction to IPW

- You are given a sample of 16-year-old children
- 12 boys have heights in cm:
  - 163 166 171 174 178 182 183 184 185 188 191 195
  - mean height of 180cm
- 8 girls have heights in cm:
  - 155 158 166 168 174 175 179 185
  - mean height of 170cm
- How do you estimate the mean height of 16-year-old children?

MRC | Medical Research Council

5

Introduction to IPW

Sample of 16-year-olds:
- 12 boys have a mean height of 180cm
- 8 girls have a mean height of 170cm

What's the mean height of 16-year-olds?

- Best estimate = (180+170)/2 = 175cm
  - because we know that the population of 16-year-olds contains (roughly) equal numbers of boys and girls
- This is an unweighted average of the gender-specific means
- It is a weighted average of the individual data
  - boys get weight 10/12 or any weight $\propto 1/12$
  - girls get weight 10/8 or any weight $\propto 1/8$
- These weights $\propto$ the inverse probabilities of selection, because (from n boys and n girls) it appears that
  - the probability of selecting each boy was $12/n$
  - the probability of selecting each girl was $8/n$

Doesn't matter if these are estimated or true probabilities of selection!

MRC | Medical Research Council

6

How IPW works: simple example

Trial centre	Rand-omised	Observed outcome		
		n	%	mean
St George's	135	76	56%	16.63
Manchester	107	89	83%	15.69
St Mary's	143	94	66%	17.79
King's	115	90	78%	18.18
All	500	349	70%	17.10
p across centres			<0.001	<0.001

- "UK500 trial" recruited patients with severe mental illness from 4 London centres
- Centres have different mean outcome *and* different % observed
- How can we estimate the mean outcome across the original 500 patients?

MRC | Medical Research Council

7

How IPW works: simple example

Trial centre	Rand-omised	Observed outcome		
		n	%	mean
St George's	135	76	56%	**16.63**
Manchester	107	89	83%	**15.69**
St Mary's	143	94	66%	**17.79**
King's	115	90	78%	**18.18**
All	500	349	70%	17.10
p across centres			<0.001	<0.001

- The mean of the observed data is the average of the **centre-specific means** with weights = #observed =17.10
- We can get the mean across the 500 by averaging the **centre-specific means** with weights = #randomised =17.12
- Equivalently, we can average the observed data with weights = #randomised / #observed =17.12

MRC | Medical Research Council

8

How IPW works: simple example

Trial centre	Rand-omised	Observed outcome		
		n	%	mean
St George's	135	76	56%	16.63
Manchester	107	89	83%	15.69
St Mary's	143	94	66%	17.79
King's	115	90	78%	18.18
All	500	349	70%	17.10
p across centres			<0.001	<0.001

- These weights could be obtained from a logistic regression of {outcome observed?} on centre
  - fitted values are p(observed)
  - weights are 1/p(observed)
- The logistic regression idea generalises nicely to more complex problems

MRC | Medical Research Council

9

Mathematical theory of IPW

With no missing data, suppose that $\sum U(Y_i, \theta) = 0$ is an unbiased estimating equation for a parameter θ

- e.g. θ is the mean of a random variable Y and $U(Y_i, \theta) = Y_i - \theta$: then solution is $\hat{\theta} = \frac{1}{n}\sum Y_i$

Then we know that the solution $\hat{\theta}$ is consistent for θ

With missing data, let R_i indicate whether $U(Y_i, \theta)$ is observed.

- $\sum R_i U(Y_i, \theta) = 0$ is NOT an unbiased estimating equation
  - "complete cases"
- But $\sum \frac{R_i U(Y_i, \theta)}{p(R_i=1|Y_i)} = 0$ IS an unbiased estimating equation
  - proof by taking expectation conditional on Y_i

weighting by the inverse of the probability (of being observed)

MRC | Medical Research Council

10

Exercise

- In a randomised trial of the effect of health advice on exercise, there is an unexpected failure to follow up men in the intervention arm. Can you use IPW to correct for this?

	Intervention	No intervention
Randomised, women / men	100 / 50	100 / 50
Followed up, women / men	100 / 20	100 / 50
Mean exercise in those followed up (minutes/day)	27	21

MRC | Medical Research Council

11

Exercise

- In a randomised trial of the effect of health advice on exercise, there is an unexpected failure to follow up men in the intervention arm. Can you use IPW to correct for this?

	Intervention			No intervention		
	Women	Men	All	Women	Men	All
Randomised	100	50	150	100	50	150
Followed up	100	20	120	100	50	150
Mean exercise in those followed up (minutes/day)	30	12	27	27	9	21

MRC | Medical Research Council

12

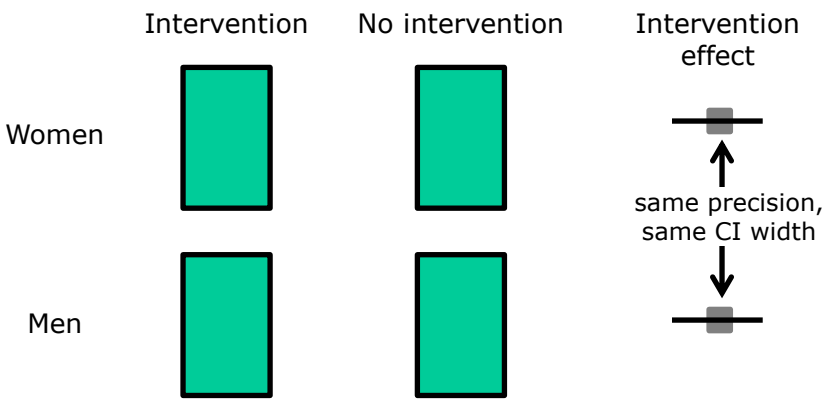
What have we assumed?

- We have assumed that the missing men in the intervention arm were comparable to the observed men in the intervention arm
  - or: missing participants were comparable to observed participants, conditional on gender and randomised group
  - "missing at random"
- We could also have estimated the intervention effect on this assumption by regression adjustment
  - see next slides

MRC | Medical Research Council

13

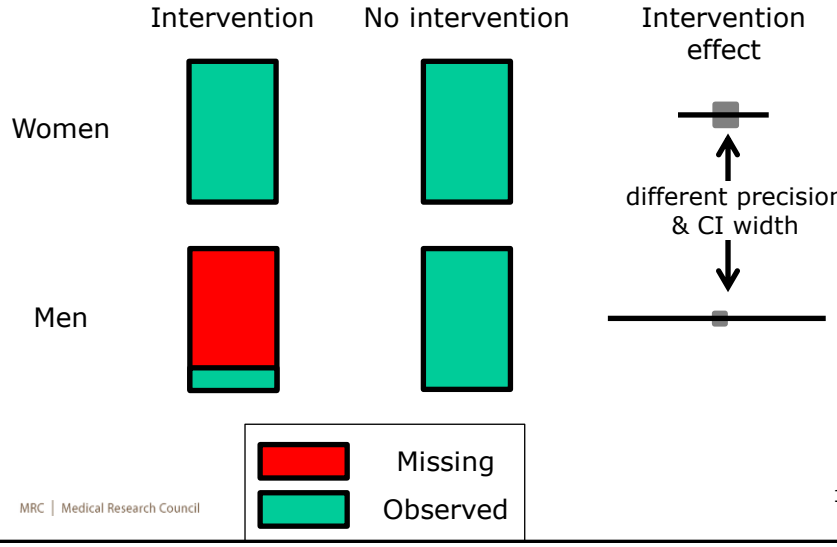
IPW vs regression adjustment



MRC | Medical Research Council

14

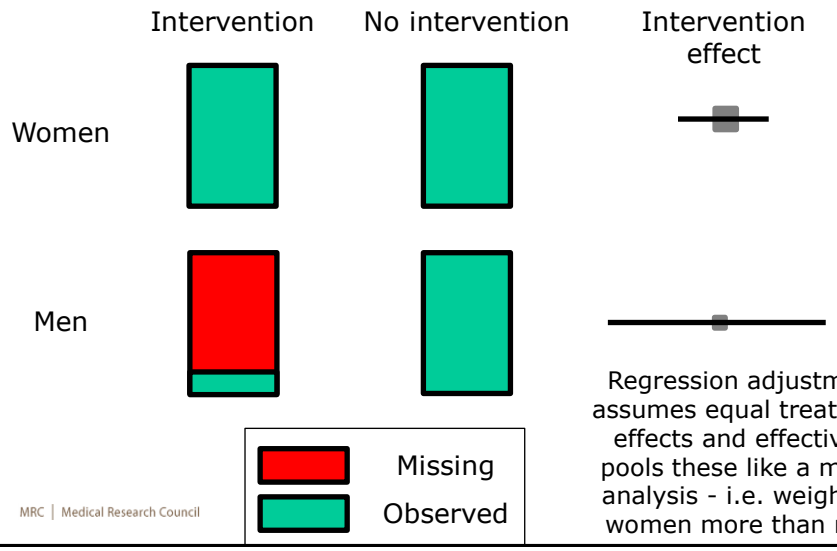
IPW vs regression adjustment



MRC | Medical Research Council

15

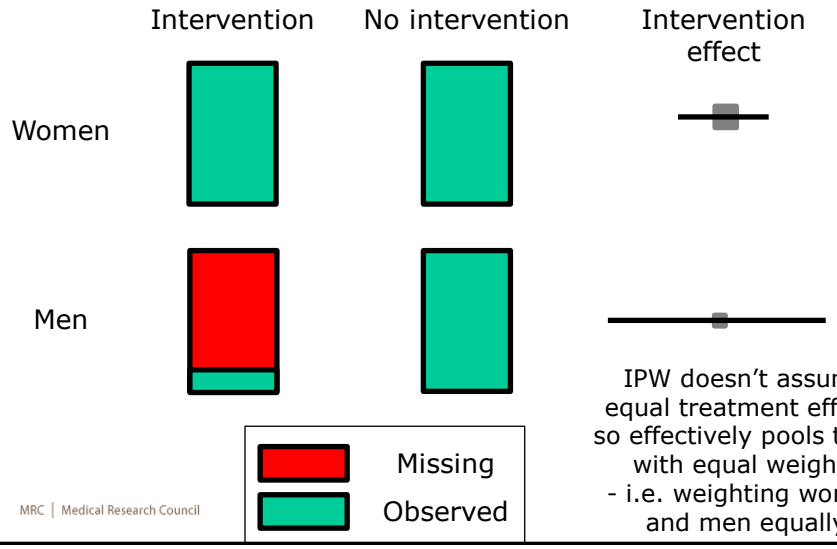
IPW vs regression adjustment



Regression adjustment assumes equal treatment effects and effectively pools these like a meta-analysis - i.e. weighting women more than men

MRC | Medical Research Council

IPW vs regression adjustment



IPW doesn't assume equal treatment effects so effectively pools these with equal weights - i.e. weighting women and men equally

MRC | Medical Research Council

IPW vs regression adjustment

- IPW and standard regression adjustment do similar jobs
  - regression adjustment is more precise
  - IPW is more robust to interactions
- [With non-linear effect measures like the odds ratio, they estimate (slightly) different parameters]

MRC | Medical Research Council

18

Exercise (revisited)

- Now suppose no men were followed up in the intervention arm.
- Can you use IPW to correct for this?
- How else can you estimate the intervention effect?

	Intervention			No intervention		
	Women	Men	All	Women	Men	All
Randomised	100	50	150	100	50	150
Followed up	100	0	120	100	50	150
Mean exercise in those followed up (minutes/day)	30	-	30	27	9	21

MRC | Medical Research Council

19

IPW: second assumption

- IPW requires "positivity" - for each individual, there is a positive probability of the data being non-missing
- Without positivity, there are some missing individuals who you can't allow for because you don't have comparable observed individuals to weight up
- [Strictly, IPW is consistent if the probability of data being non-missing is greater than some ϵ as sample size goes to ∞]

MRC | Medical Research Council

20

Plan

1. IPW for missing data
2. **IPW for artificially missing data**
3. Introduction to IPCW (= IPW for longitudinal data)
4. Assumptions in IPCW
5. Modelling switching
6. Estimating and using the weights
7. Illustration using SANAD trial
8. Complexities - for future reference
9. Summary

When else can we use IPW? (1)

We can use IPW to allow for confounding in observational studies.

- Why? Because we observe
  - $Y = Y(1)$ in the $D = 1$ group
  - $Y = Y(0)$ in the $D = 0$ group
- But we want the average treatment effect,

$$ATE = E[Y(1) - Y(0)]$$

  (or ATT / ATU)
- So we can regard
  - $Y(1)$ as missing data in the $D = 0$ group
  - $Y(0)$ as missing data in the $D = 1$ group
- (points to propensity score methods)

When else can we use IPW? (2)

We can use IPW to allow for non-compliance in randomised trials under the "Exclude" approach of Lecture 1 (=per-protocol)

- Why? Because we observe
  - $Y(0)$ when $D = 0$
  - $Y(1)$ when $D = 1$
- We want to estimate
  - the mean of $Y(0)$ in the $Z = 0$ arm
  - the mean of $Y(1)$ in the $Z = 1$ arm
- So we can regard
  - $Y(0)$ as missing whenever $Z = 0$ and $D = 1$
  - $Y(1)$ as missing whenever $Z = 1$ and $D = 0$
- However this is usually undesirable (see Lecture 1)

MRC | Medical Research Council

23

Plan

1. IPW for missing data
2. IPW for artificially missing data
3. **Introduction to IPCW (= IPW for longitudinal data)**
4. Assumptions in IPCW
5. Modelling switching
6. Estimating and using the weights
7. Illustration using SANAD trial
8. Complexities - for future reference
9. Summary

MRC | Medical Research Council

24

IPCW is like IPW, but

- IPCW is a longitudinal version of IPW
- Simple data are fully observed or fully missing
- Longitudinal data are observed up to a certain time and then missing
  - i.e. we have censoring
- As an individual advances through time, their probability of being censored increases
  - probability of remaining uncensored decreases
  - weight = 1/{probability of remaining uncensored} increases
- So we have time-dependent weights

What's in a name?

- IPCW:
  "Inverse probability of censoring weighting"
- But it is really:
  "Inverse probability of not-having-been-censored weighting"
- IPNHBCW???

Steps in IPCW for missing data

1. Identify when each observation is censored
2. Model the probability of being censored over time - "censoring model"
3. For each individual at each time, compute the inverse probability of remaining uncensored
4. Use these inverse probabilities as weights in a weighted analysis - "outcome model"

Estimand in IPCW

- We are going to artificially censor the data at the point of a treatment switch
  - user decides what sort of treatment switch implies censoring
  - could also censor at any other protocol violation
- We will use IPCW to up-weight the observed data to represent the censored data
  - and hence we will estimate the ATE (average treatment effect)

Steps in IPCW for treatment switches

1. Censor observations at treatment switches
   - hence I'll use "censor" and "switch" interchangeably
2. Model the probability of being censored over time - "switching model"
3. For each individual at each time, compute the inverse probability of remaining unswitched
4. Use these inverse probabilities as weights in a weighted analysis - "outcome model"

I will focus on switching in the control arm
- but IPCW can be applied in both arms

Can use IPCW for missing data **and** treatment switches
- multiply the weights together

MRC | Medical Research Council

29

Simple example from a hypothetical trial (observed counts)

Randomise: Exp or Pbo

1st follow-up: detect disease progression; no deaths at this stage

2nd follow-up: look at mortality

Arm	Time 1		Time 2 status Dead/Total
	Progression	Switch	
Exp	No (800)		10 / 800
	Yes (200)		90 / 200
Pbo	No (600)	No (600)	10 / 600
	Yes (400)	No (100)	40 / 100
		Yes (300)	90 / 300

Pbo arm progressors may switch to Exp

Question: what would be the difference between the two arms if no switching occurred in the Pbo arm (ATE)?

MRC | Medical Research Council

30

Simple example: per-protocol analysis

Arm	Time 1		Time 2 status Dead/Total
	Progression	Switch	
Exp	No (800)		10 / 800
	Yes (200)		90 / 200
Pbo	No (600)	No (600)	10 / 600
	Yes (400)	No (100)	40 / 100
		~~Yes (300)~~	~~90 / 300~~

Result: 100/1000 vs. 50/700

MRC | Medical Research Council

31

Simple example: IPCW analysis

step 1 step 2 step 3 step 4

Arm	Time 1		Time 2 status Dead/Total	weights
	Progression	Switch		
Exp	No (800)		10 / 800	1
	Yes (200)		90 / 200	1
Pbo	No (600)	No (600)	10 / 600	1
	Yes (400)	No (100)	40 / 100	4
		~~Yes (300)~~	~~90 / 300~~	0

P(no switch | Pbo, progressed) = 0.25 → non-switchers get weight 4

Assumes switchers & non-switchers have comparable counterfactual switch-free outcomes

Result: 100/1000 vs. 170/1000

32

Simple example: causal assumption of IPCW analysis

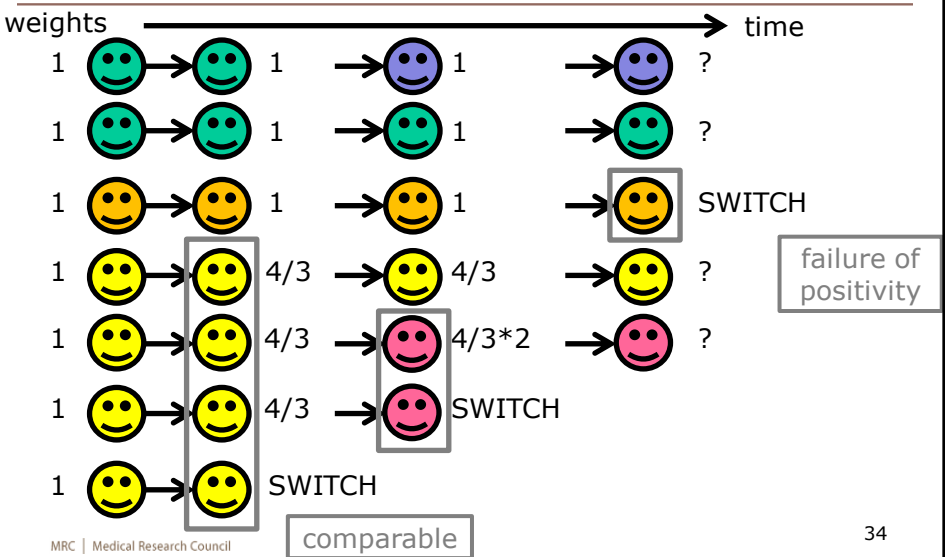
Arm	Time 1		Time 2 status Dead/Total
	Progression	Switch	
Exp	No (800)		$Y(1)$
	Yes (200)		$Y(1)$
Pbo	No (600)	No (600)	$Y(0)$
	Yes (400)	No (100)	$Y(0)$
		Yes (300)	$Y(0)$ then

IPCW assumes these (missing) values of Y(0)

are like these (observed) values of Y(0)

MRC | Medical Research Council

More complex example of IPCW (one trial arm)



MRC | Medical Research Council

34

Plan

1. IPW for missing data
2. IPW for artificially missing data
3. Introduction to IPCW (= IPW for longitudinal data)
4. **Assumptions in IPCW**
5. Modelling switching
6. Estimating and using the weights
7. Illustration using SANAD trial
8. Complexities - for future reference
9. Summary

No unmeasured confounders (1)

- We are assuming that, at each time, there are no systematic differences between those who are censored and those who remain uncensored
  - conditional on variables in the model
- This is usually described as "no unmeasured confounders", where a confounder is a variable which
  - predicts switching, and
  - predicts *counterfactual* time to event (i.e. time to event if no switch)
- Any confounders must be measured and included in the model

No unmeasured confounders (2)

- Time-dependent confounders can be used
  - this is the key feature that makes IPCW typically more plausible than ordinary IPW
- What are these time-dependent confounders?
  - first example: progression
  - second example: history of face colour
  - in practice: can involve full clinical history …

Arm	Time 1		Time 2 status Dead/Total
	Progression	Switch	
Exp	No (800)		10 / 800
	Yes (200)		90 / 200
Pbo	No (600)	No (600)	10 / 600
	Yes (400)	No (100)	40 / 100
		Yes (300)	90 / 300

MRC | Medical Research Council

37

No unmeasured confounders (3)

- Typically, we use many baseline confounders and time-dependent confounders
- Validity of NUC cannot be assessed statistically
  - as in observational epidemiology, we can never be sure we have no unmeasured confounders
  - have to argue what are the likely confounders and measure them
  - but we can explore adding more and more potential confounders into the model: little impact on the estimated treatment effect → we may be OK

MRC | Medical Research Council

38

Positivity

- IPCW also assumes that for each individual there is a positive probability that they will remain unswitched (uncensored) at any given time

… pause for questions

MRC | Medical Research Council

39

Plan

1. IPW for missing data
2. IPW for artificially missing data
3. Introduction to IPCW (= IPW for longitudinal data)
4. Assumptions in IPCW
5. **Modelling switching**
   - **IPCW step 2**
6. Estimating and using the weights
7. Illustration using SANAD trial
8. Complexities - for future reference
9. Summary

MRC | Medical Research Council

40

Discrete or continuous time?

- We will work in discrete time as this is generally found to be simpler and adequate.
  - this means we will divide follow-up into intervals
  - the number of intervals shouldn't be too important
  - we can vary the number of intervals in a sensitivity analysis
- Later I'll mention an extension to continuous time.
  - see "complexities"
- Our "pooled logistic models" in discrete time are analogous to Cox models in continuous time

Notation (omitting subscript i)

- X = an array of an individual's baseline covariates [V in Hernan et al (2001)]
- Z = randomised group
- $A(k)$ = switching in interval k (regarded as happening at the start of the interval)
  - i.e. $A(k) = 1$ if $D(k-1) = Z$ and $D(k) \neq Z$
- $L(k)$ = time-dependent covariates measured at or prior to the beginning of interval k
- T = interval in which event occurs

Also use overbars to denote history:

- $\bar{A}(k)$ = individual's treatment history up until the end of interval k
- $\bar{L}(k)$ = history of an individual's time-dependent covariates up to the beginning of interval k

Switching model

X = baseline covariates
Z = randomised group
$A(k)$ = treatment in interval k
$L(k)$ = time-dependent covariates
Overbar = history

- We need a model for switching in the control arm $Z = 0$
  - "switching model"
- For interval 1, propose:
$$logit\ P(A(1) = 1 \mid Z = 0, X) = \alpha_1' X$$
- For interval 2, propose:
$$logit\ P(A(2) = 1 \mid Z = 0, X, A(1) = 0, L(1), T \geq 2) = \alpha_2' X + \gamma_2' L(1)$$
- In general, propose a model for
$$logit\ P(A(k) = 1 \mid Z = 0, X, A(k-1) = 0, \bar{L}(k), T \geq k) \ldots$$
- Typically needs us to specify a suitable function of $\bar{L}(k)$: e.g. set $L^*(k)$ = (progressed at k?, progressed at $k-1$?, severity at k, severity at $k-1$) and model
$$\ldots = \alpha_k' X + \gamma_k' L^*(k)$$
  - how do we specify this model?

MRC | Medical Research Council

43

Selecting the switching model

X = baseline covariates
Z = randomised group
$A(k)$ = treatment in interval k
$L(k)$ = time-dependent covariates
Overbar = history

- Selecting covariates: how to define $L^*(k)$ from the very rich history $\bar{L}(k)$?
- Selecting functional forms: how to enter $L^*(k)$ in the model?
- How to combine models over different intervals?
  - in practice we don't have enough data to fit one model at each interval
  - instead choose a model across intervals, e.g.
$$logit\ P(A(k) = 1 | Z = 0, X, A(k-1) = 0, \bar{L}(k), T \geq k) = \alpha' X + \gamma' L^*(k) + \delta' f(k)$$

where $f(k)$ is some sort of spline.

See Dodd (2014).

MRC | Medical Research Council

44

Plan

1. IPW for missing data
2. IPW for artificially missing data
3. Introduction to IPCW (= IPW for longitudinal data)
4. Assumptions in IPCW
5. Modelling switching
6. **Estimating and using the weights**
   - **IPCW steps 3 & 4**
7. Illustration using SANAD trial
8. Complexities - for future reference
9. Summary

MRC | Medical Research Council

45

Constructing the weights (IPCW step 3)

X = baseline covariates
Z = randomised group
$A(k)$ = treatment in interval k
$L(k)$ = time-dependent covariates
Overbar = history

- First use the fitted switching model to predict the probability of switching for individual i in interval k:

$$p_{ik} = Pr[A(k) = 1 \mid Z = 0, X, A(k-1) = 0, \bar{L}(k), T \geq k]$$

- Then construct the weight for individual i in interval t by multiplying together all their probabilities of remaining unswitched up to interval t:

$$w_{it} = \prod_{k=0}^{t} \frac{1}{1 - p_{ik}}$$

- An improvement is stabilised weights
  - see "complexities"

MRC | Medical Research Council

46

Using the weights (IPCW step 4)

- The last step should be the easiest
- We just have to fit our outcome model using the weights
- If we're working in discrete time intervals, this is another pooled logistic regression
  - but now fitted to both arms

Plan

1. IPW for missing data
2. IPW for artificially missing data
3. Introduction to IPCW (= IPW for longitudinal data)
4. Assumptions in IPCW
5. Modelling switching
6. Estimating and using the weights
7. **Illustration using SANAD trial**
8. Complexities - for future reference
9. Summary

SANAD B trial

- SANAD = "Standard And New Antiepileptic Drugs"
- Marson AG et al. (2007) The SANAD study of effectiveness of valproate, lamotrigine, or topiramate for generalised and unclassifiable epilepsy: an unblinded randomised controlled trial. *The Lancet* 369: 1016–26.
- Here:
  - compare lamotrigine (LTG) with valproate (VPS)
  - ignore the topiramate arm
  - outcome is time to 52-week remission (i.e. 52 weeks without seizures)
  - based on Susie Dodd's PhD thesis (Liverpool, 2014), but with a simplified analysis strategy
- Thank you to the SANAD investigators for permission to use the data and to Susie for supplying the data

MRC | Medical Research Council

49

SANAD variables (1): Baseline covariates X

Variable	Range	Description
hist2	2 levels	treatment history
feb	2 levels	febrile seizure history
rels	2 levels	first degree relative with epilepsy
type	2 levels	epilepsy type (generalised (G) / unclassified (U))
sex	2 levels	sex
ni	2 levels	neurological insult
eeg	4 levels	EEG
ln_age_t95	1.7-4.0	age, log-transformed
ln_fi_t95	2.6-9.1	interval from first ever seizure to randomisation, log-transformed
ln_tt95	0-3.0	total number of tonic-clonic seizures, log-transformed

SANAD variables (2)

Variable	Range	Description
Randomisation		
trialno	1-387	Participant ID
treat	LTG/VPS	Randomised group (Z)
Time-varying (L)		
bw	1-130	bi-week (time)
ln_sez_t95	0-5.6	cumulative seizure count since randomisation, log-transformed
cum_AE_t95	0-4	Cumulative AE count
dose_t95	0-2000	Dose, log-transformed
Outcomes		
wcens	0/1	switching indicator ($A(k)$)
rcens	0/1	remission indicator (whether $T = k$)

SANAD: what's the question?

- Participants in both arms change treatment
  - for various reasons: inadequate seizure control, adverse events, patient choice
  - to various treatments: other trial treatment, other anti-epileptic drug, drug combination, …
- Here, we ask what would be the effect of randomisation to LTG compared to VPS if there were no treatment changes
  - thus we censor at all the above treatment changes
  - and use IPCW to (try to) adjust for selection bias due to this censoring
  - estimating ATE
- We could also adjust only for a subset of these treatment changes

MRC | Medical Research Council

52

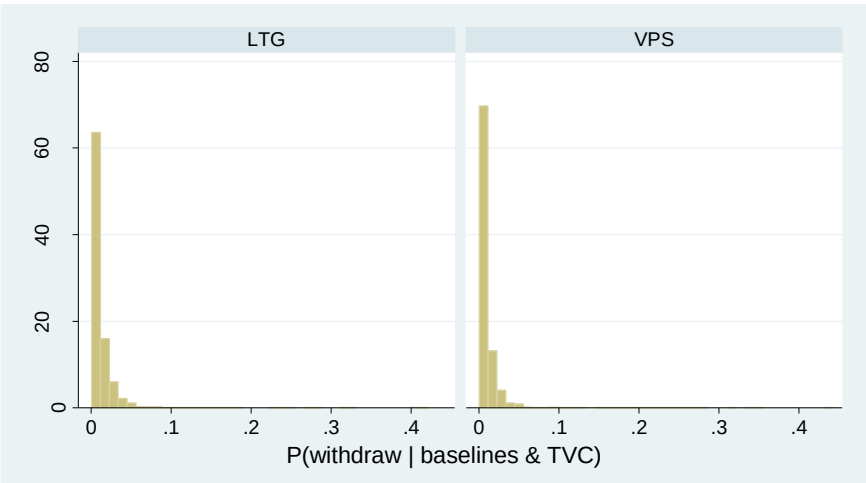
Modelling strategy: switching model

- Outcome = switching
- Covariates:
  - baseline covariates
  - time-varying covariates (TVCs)
  - time
- Model separately by randomised group
- Use multivariable fractional polynomials to select functional form of quantitative variables
  - Royston P, Sauerbrei W (2008). *Multivariable Model-building: a Pragmatic Approach to Regression Analysis Based on Fractional Polynomials for Modelling Continuous Variables*. Wiley, Chichester.
- Check fitted probabilities and weights

MRC | Medical Research Council

53

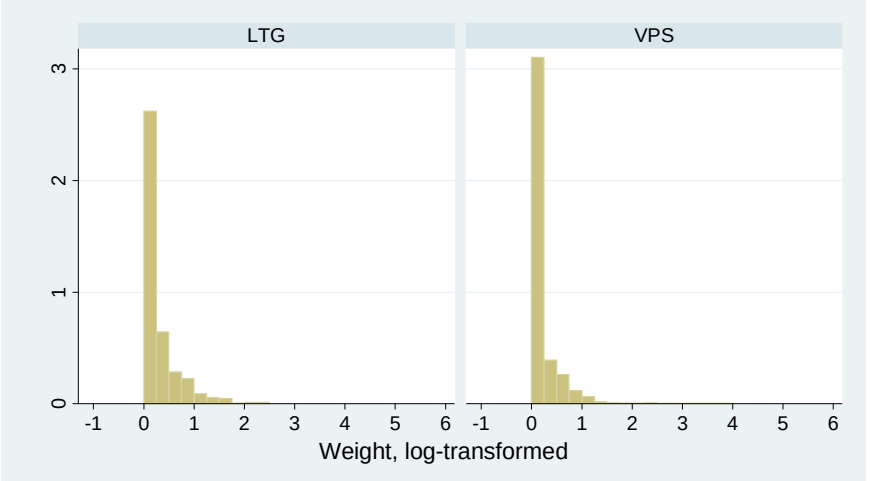
SANAD: fitted withdrawal probabilities



MRC | Medical Research Council

54

SANAD: weights

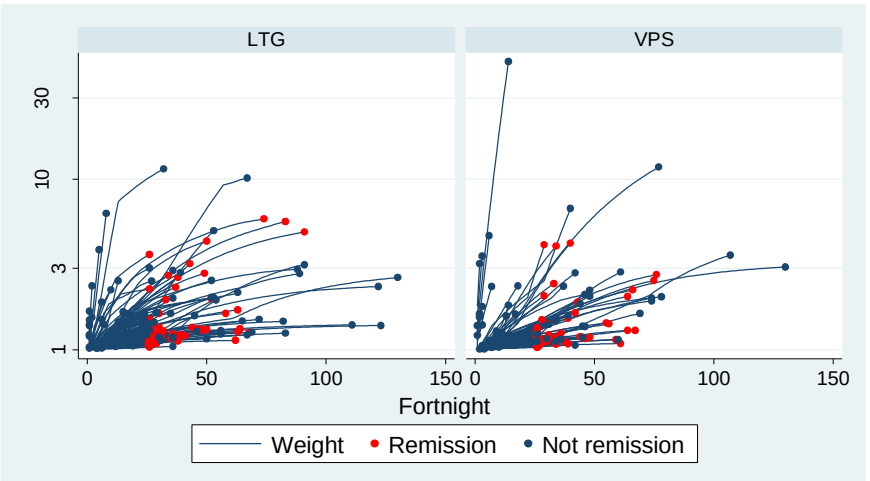


MRC | Medical Research Council

untransformed range: 1-50
CV: 0.77

55

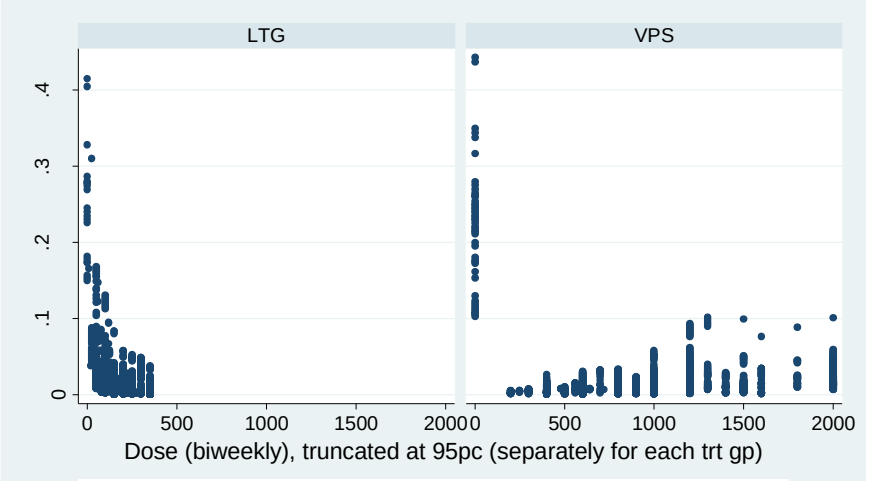
SANAD: weights by time



MRC | Medical Research Council

56

Large weights arise from low doses



Very low doses have led to large p(withdraw)
→ large weights after several periods on low dose

MRC | Me

57

Modelling stategy: outcome model

- Recall: outcome is 52-week remission
- Outcome can't occur before 52 weeks
- 60% of outcomes occur at 52 weeks
- So we
  - exclude intervals before 52 weeks
  - include a dummy for 52 weeks
  - include a linear term for time after 52 weeks
- And randomised group
- And we apply the weights

MRC | Medical Research Council

58

SANAD: results

Method	Odds ratio for remission (LTG vs. VPS)	Comment
ITT	0.76 (0.62 - 0.94)*	Doesn't estimate effect of treatment itself
Per-protocol	0.65 (0.47 - 0.90)	Subject to selection bias
IPCW	0.68 (0.47 - 0.98)	Removes selection bias due to observed confounders

* from Marson et al. *Lancet* 2007; 369: 1016–1026

MRC | Medical Research Council

59

Plan

1. IPW for missing data
2. IPW for artificially missing data
3. Introduction to IPCW (= IPW for longitudinal data)
4. Assumptions in IPCW
5. Modelling switching
6. Estimating and using the weights
7. Illustration using SANAD trial
8. **Complexities - for future reference**
   a) **Stabilised weights**
   b) **Large weights**
   c) **Continuous time**
   d) **IPTW**
9. Summary

MRC | Medical Research Council

60

Stabilised weights (1)

V = baseline covariates
Z = randomised group
$A(k)$ = treatment in interval k
$L(k)$ = time-dependent covariates
Overbar = history

- Suppose the time-dependent covariates $\bar{L}(k)$ don't predict switching
- So baseline covariates V are the only confounders
- But we can adjust for V in the outcome model
  - big difference between baseline and time-dependent covariates: regression adjustment can handle confounding by baseline covariates
- So we don't need to weight for V (if we are going to adjust for V in the analysis)
- This suggests stabilised weights…

MRC | Medical Research Council

61

Stabilised weights (2)

V = baseline covariates
Z = randomised group
$A(k)$ = treatment in interval k
$L(k)$ = time-dependent covariates
Overbar = history

- Weights above were

$$\widehat{W}(t) = \prod_{k=0}^{t} \frac{1}{1 - p_{ik}}$$

where $p_{ik} = Pr[A(k) = 1 \mid Z = 0, V, A(k-1) = 0, \bar{L}(k), T \geq k]$

- We replace these with the stabilised weights

$$\widehat{W}_s(t) = \prod_{k=0}^{t} \frac{1}{1 - p_{ik}} \Big/ \prod_{k=0}^{t} \frac{1}{1 - p_{0ik}} = \prod_{k=0}^{t} \frac{1 - p_{0ik}}{1 - p_{ik}}$$

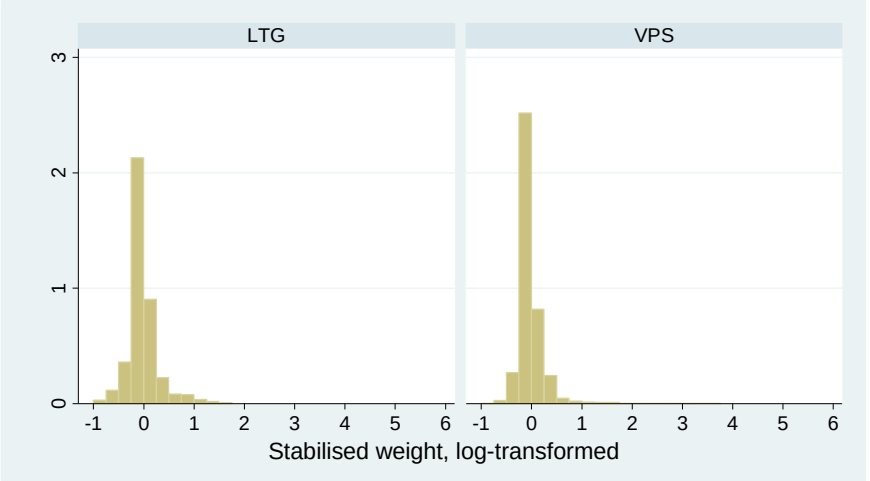
where $p_{0ik} = Pr[A(k) = 1 \mid Z = 0, V, A(k-1) = 0, T \geq k]$

- If the time-dependent covariates $\bar{L}(k)$ don't predict switching then $p_{0ik} = p_{ik}$ and $\widehat{W}_s(t) = 1$
- To estimate the p_{0ik} we need to fit a second model that is just like the first model but doesn't use the time-dependent covariates $\bar{L}(k)$.

MRC | Medical Research Council

62

SANAD: stabilised weights

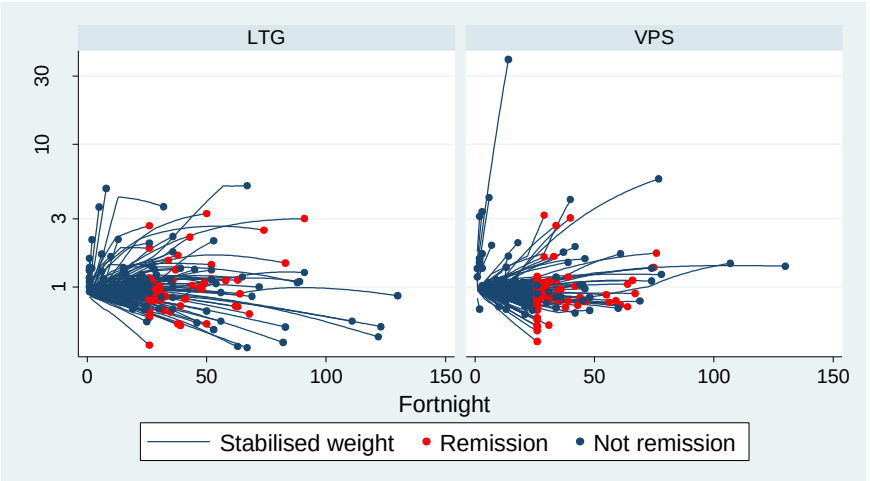


MRC | Medical Research Council

untransformed range: 0.37-39
CV: 0.67 (smaller than unstabilised)

63

SANAD: stabilised weights by time



MRC | Medical Research Council

64

SANAD: results

Method	Weights	Odds ratio for remission (LTG vs. VPS)	Comment
ITT	None	0.76 (0.62 - 0.94)*	Doesn't estimate effect of treatment itself
Per-protocol	None	0.65 (0.47 - 0.90)	Subject to selection bias
IPCW	Weights	0.68 (0.47 - 0.98)	Removes selection bias due to observed confounders
IPCW	Stabilised weights	0.68 (0.48 - 0.97)	Ditto, and slightly improves efficiency

* from Marson et al. *Lancet* 2007; 369: 1016–1026

MRC | Medical Research Council

65

Large weights

- The main pitfall in IPCW is getting very large weights
- This represents a fundamental problem
- For example, suppose 99% of progressing patients immediately switch
  - IPCW assumes the remaining 1% are representative of the missing 99%
    - » *you might well be suspicious of this, but let's assume it*
  - IPCW then gives the 1% a weight of 100
  - this means you have very little data getting a very large role, hence large standard errors
- Now suppose 100% of progressing patients immediately switch
  - you have no non-switchers to weight up
  - IPCW fails ("non-positivity")

MRC | Medical Research Council

66

How should we tackle large weights? 1. Prevention

The key is to be cautious in building switching models

- Variable selection
  - we want confounders
    - » *i.e. variables which predict both switching and counterfactual time-to-event*
  - we don't want variables which only predict switching
    - » *since these unnecessarily inflate the weights*
- Functional form
  - sometimes a linear form leads to very large weights for outlying observations
  - consider other functional forms

MRC | Medical Research Council

67

How should we tackle large weights? 2. Treatment

- Some authors propose "truncating" weights
  - e.g. if weight > 100, replace weight = 100.
- No good theoretical rationale
- Seaman & White (2013):
  - "If the missingness model is correctly specified and the large weights arise because the predictors of missingness are highly informative, truncation may re-introduce some of the bias IPW was used to eliminate. However, when large weights are likely due to model misspecification, truncating them is a reasonable measure."
- Still worse is excluding individuals with large weights

MRC | Medical Research Council

68

Continuous time

Can alternatively

- Model time to switching in continuous time
  - Cox model with time-dependent covariates
- Compute fully time-dependent weights
  - in practice, only need them at the outcome event times
- Apply these weights in a Cox model analysis
  - Cox model with time-dependent weights
  - hard to fit in some software

MRC | Medical Research Council

69

IPTW vs. IPCW

- IPCW = inverse probability of censoring weighting
  - is "per-protocol analysis done well"
  - it censors treatment switchers and constructs weights to avoid selection bias
  - we built models to predict switching, and hence estimating the probabilities of remaining unswitched
- IPTW = inverse probability of treatment weighting
  - is "as-treated analysis done well"
  - it analyses by treatment received and constructs weights to avoid selection bias
  - we would use the same models, but also use data after switch with estimated probabilities of being switched

MRC | Medical Research Council

70

Role of IPTW with treatment switching

- IPTW requires a marginal structural model (MSM) that describes how outcome relates to treatment history
  - NB this is a causal model so it compares counterfactual outcomes in same individuals, not actual outcomes between different individuals
- Here, the MSM would describe the hazard of death for never treated, always treated and switchers
- Would we model the hazard for switchers as
  - the same as for those always on treatment?
    - » *seems a very strong assumption: requires past treatment to be irrelevant*
  - or not the same?
    - » *then the switchers don't inform estimation of the contrast of interest, always treated vs. never treated*
    - » *so we might as well go back to IPCW*

MRC | Medical Research Council

71

IV methods for survival data

- Can we do something like a CACE analysis for data like SANAD?
- Yes! There is an IV method called the rank-preserving structural nested failure time model (RPSFTM)
  - e.g. Watkins et al. Adjusting overall survival for treatment switches: commonly used methods and practical application. *Pharm Stat*. 2013;12(6):348-357
- However, IV methods struggle with complex treatment changes such as we have in SANAD, because they must model post-switch outcomes
- Assumption trade-off tips towards IPCW (Lecture 1):
  - time-dependent variables → no unmeasured confounders more plausible
  - complex switching → IV methods less feasible

MRC | Medical Research Council

72

Plan

1. IPW for missing data
2. IPW for artificially missing data
3. Introduction to IPCW (= IPW for longitudinal data)
4. Assumptions in IPCW
5. Modelling switching
6. Estimating and using the weights
7. Illustration using SANAD trial
8. Complexities - for future reference
9. **Summary**

4 computational steps in IPCW analysis using pooled logistic regression (1)

1. Censor observations and re-format the data
   - one record per individual per interval
   - baseline covariates repeated across intervals
   - time-dependent covariates
   - time-dependent outcomes: switching, death
   - exclude (or flag) data after switch

   Check the re-format by repeating the ITT analysis
2. Model the probability of being censored over time
   - using logistic regression
   - first time conditional on baseline covariates
   - second time conditional on baseline and time-dependent covariates
   - use a spline for interval
   - usually requires model building and checking

4 computational steps in IPCW analysis using pooled logistic regression (2)

3. For each individual at each time, compute the inverse probability of remaining uncensored
   - extract the fitted probabilities of switching (one per individual per interval), say p_{0it} and p_{ik} for the two models
   - combine the fitted probabilities to get the stabilised weights $sw_{it} = \prod_{k \le t} \frac{1-p_{0ik}}{1-p_{ik}}$
   - check the range of the weights; if large, consider changing model or truncating weights
   - repeat steps 2 & 3 for the treatment arm, if switching occurs there; otherwise let its weights be 1
4. Use these weights in a weighted analysis of the outcome model

MRC | Medical Research Council

75

Key messages

- IPW can adjust for selection bias when certain individuals (typically those who have changed treatment) are excluded
- IPCW does the same job in longitudinal data where an individual's follow-up after a switch is excluded
- Both methods assume no unmeasured confounders and positivity
- No unmeasured confounders is often more plausible for IPCW because time-dependent variables can be adjusted for

MRC | Medical Research Council

76