Effect Modification

1

Thursday afternoon

- What is an effect modifier?
- Caveats about its interpretation:
  - What is the connection to risk factors that we should adjust for?
  - Dependence on the effect measure of choice
  - Why does effect modification take place?
- Why care about effect modification?
- How do we analyze the data to examine the presence of effect modifiers?
- Marginal structural models

2

Effect modification

- **What is an effect modifier?**
- Caveats about its interpretation:
  - What is the connection to risk factors that we should adjust for?
  - Dependence on the effect measure of choice
  - Why does effect modification take place?
- Why care about effect modification?
- How do we analyze the data to examine the presence of effect modifiers?
- Marginal structural models

3

Patient	M	Y_1	Y_0
1. Tom		1	0
2. John		1	0
3. Mary		1	0
4. Amalia		1	0
5. Rick		0	0
6. Rose		1	1
7. Tim		1	1
8. Peter		0	0
9. Ruth		0	0
10. Doris		0	0
11. Anna		0	1
12. Helen		0	1
13. Oscar		0	1
14. Ruben		0	1
15. George		0	0
16. Andrea		1	1
17. Richard		1	1
18. Betsy		0	0
19. Alicia		1	1
20. Phoebe		1	1

$P(Y_1 = 1) = 10/20$

$P(Y_0 = 1) = 10/20$

No population average effects

4

Patient	M	Y_1	Y_0
1. Tom	1	1	0
2. John	1	1	0
3. Mary	1	1	0
4. Amalia	1	1	0
5. Rick	1	0	0
6. Rose	1	1	1
7. Tim	1	1	1
8. Peter	1	0	0
9. Ruth	1	0	0
10. Doris	1	0	0
11. Anna	0	0	1
12. Helen	0	0	1
13. Oscar	0	0	1
14. Ruben	0	0	1
15. George	0	0	0
16. Andrea	0	1	1
17. Richard	0	1	1
18. Betsy	0	0	0
19. Alicia	0	1	1
20. Phoebe	0	1	1

M = 1 South African
M = 0 Canadian

$P(Y_1 = 1 | M=1) = 6/10$

$P(Y_0 = 1 | M=1) = 2/10$

CRD in South Africans = 4/10

$P(Y_1 = 1 | M=0) = 4/10$

$P(Y_0 = 1 | M=0) = 8/10$

CRD in Canadians= - 4/10

5

Patient	M	Y_1	Y_0
1. Tom	1	1	0
2. John	1	1	0
3. Mary	1	1	0
4. Amalia	1	1	0
5. Rick	1	0	0
6. Rose	1	1	1
7. Tim	1	1	1
8. Peter	1	0	0
9. Ruth	1	0	0
10. Doris	1	0	0
11. Anna	0	0	1
12. Helen	0	0	1
13. Oscar	0	0	1
14. Ruben	0	0	1
15. George	0	0	0
16. Andrea	0	1	1
17. Richard	0	1	1
18. Betsy	0	0	0
19. Alicia	0	1	1
20. Phoebe	0	1	1

Causal risk differences
CRD in South Africans = 4/10
CRD in Canadians= - 4/10

Nationality is an effect modifier on the additive scale.

Definition: M is an effect modifier of the effect of A on Y on the additive scale if M is a pre-trx variable and when we stratify by M the causal average effect (ATE or CRD) differs accross strata

6

Patient	M	Y_1	Y_0
1. Tom	1	1	0
2. John	1	1	0
3. Mary	1	1	0
4. Amalia	1	1	0
5. Rick	1	0	0
6. Rose	1	1	1
7. Tim	1	1	1
8. Peter	1	0	0
9. Ruth	1	0	0
10. Doris	1	0	0
11. Anna	0	0	1
12. Helen	0	0	1
13. Oscar	0	0	1
14. Ruben	0	0	1
15. George	0	0	0
16. Andrea	0	1	1
17. Richard	0	1	1
18. Betsy	0	0	0
19. Alicia	0	1	1
20. Phoebe	0	1	1

Causal risk ratios

CRR in South Africans = 3

CRR in Canadians= 1/2

Nationality is an effect modifier on the multiplicative scale.

Definition: M is an effect modifier of the effect of A on Y on the multiplicative scale if M is a pre-trx variable and when we stratify by M the causal risk ratio (CRR) differs accross strata.

7

Effect modification

- What is an effect modifier?
- **Caveats about its interpretation:**
  - **What is the connection to risk factors that we should adjust for?**
  - Dependence on the effect measure of choice
  - Why does effect modification take place?
- Why care about effect modification?
- How do we analyze the data to examine the presence of effect modifiers?
- Marginal structural models

8

Lesson 1: Effect modifiers and conditional exchangeability

- An **effect modifier** may or may not be a variable L needed to make your observational study emulate a conditionally randomized study.
- And conversely …
- A variable L that is needed to make your observational study emulate a conditionally randomized study may or may not be an **effect modifier.**

9

Example of a case in which:

M is an effect modifier

but

you do not need to condition on M to make your study emulate a randomized trial

10

Patient	M	A	Y	Y_1	Y_0
1. Tom	1	1	1	1	0
2. John	1	1	1	1	0
3. Mary	1	0	0	1	0
4. Amalia	1	0	0	1	0
5. Rick	1	0	0	0	0
6. Rose	1	0	1	1	1
7. Tim	1	0	1	1	1
8. Peter	1	0	0	0	0
9. Ruth	1	1	0	0	0
10. Doris	1	1	0	0	0
11. Anna	0	1	0	0	1
12. Helen	0	1	0	0	1
13. Oscar	0	0	1	0	1
14. Ruben	0	0	1	0	1
15. George	0	0	0	0	0
16. Andrea	0	0	1	1	1
17. Richard	0	1	1	1	1
18. Betsy	0	1	0	0	0
19. Alicia	0	1	1	1	1
20. Phoebe	0	1	1	1	1

$P(Y_1=1|A=1)$ = 5 / 10
$= P(Y_1=1|A=0)$

$P(Y_0=1|A=0)$ = 5/10
$= P(Y_0=1|A=1)$

There is exchangeability.

No need to standardize by M.

But M is nevertheless an effect modifier.

M is not like L

11

Example of a case in which:

you need to condition on L to make your study emulate a randomized trial

but

L is an not effect modifier

12

Patient	L	A	Y	Y_1	Y_0
1. Tom	1	1	1	1	0
2. John	1	1	1	1	0
3. Mary	1	1	1	1	1
4. Amalia	1	1	1	1	1
5. Rick	1	1	1	1	1
6. Rose	1	1	1	1	1
7. Tim	1	0	0	1	0
8. Peter	1	0	1	1	1
9. Ruth	1	0	1	1	1
10. Doris	0	1	1	1	0
11. Anna	0	1	0	0	0
12. Helen	0	1	0	0	0
13. Oscar	0	0	0	1	0
14. Ruben	0	0	0	1	0
15. George	0	0	0	0	0
16. Andrea	0	0	0	0	0
17. Richard	0	0	0	0	0
18. Betsy	0	0	0	0	0

$P(Y_1=1) = 12 / 18$ but

$P(Y_1 = 1 \mid A = 1) = 1$

Also,

$P(Y_0=1) = 6 / 18$ but

$P(Y_0=1 \mid A = 0) = 0$

No marginal exchangeability

However, there is conditional exchangeability given L

13

Patient	L	A	Y	Y_1	Y_0
1. Tom	1	1	1	1	0
2. John	1	1	1	1	0
3. Mary	1	1	1	1	1
4. Amalia	1	1	1	1	1
5. Rick	1	1	1	1	1
6. Rose	1	1	1	1	1
7. Tim	1	0	0	1	0
8. Peter	1	0	1	1	1
9. Ruth	1	0	1	1	1
10. Doris	0	1	1	1	0
11. Anna	0	1	0	0	0
12. Helen	0	1	0	0	0
13. Oscar	0	0	0	1	0
14. Ruben	0	0	0	1	0
15. George	0	0	0	0	0
16. Andrea	0	0	0	0	0
17. Richard	0	0	0	0	0
18. Betsy	0	0	0	0	0

$P(Y_1=1 \mid A=1, L=1) =$
$P(Y_1=1 \mid A=0, L=1) = 1$
and
$P(Y_0=1 \mid A=1, L=1) =$
$P(Y_0=1 \mid A=0, L=1) = 2/3$

Exchangeability in the stratum L=1

Likewise

$P(Y_1=1 \mid A=1, L=0) =$
$P(Y_1=1 \mid A=0, L=0) = 1/3$
and
$P(Y_0=1 \mid A=1, L=0) =$
$P(Y_0=1 \mid A=0, L=0) = 0$

Exchangeability in the stratum L=0

14

Patient	L	A	Y	Y_1	Y_0
1. Tom	1	1	1	1	0
2. John	1	1	1	1	0
3. Mary	1	1	1	1	1
4. Amalia	1	1	1	1	1
5. Rick	1	1	1	1	1
6. Rose	1	1	1	1	1
7. Tim	1	0	0	1	0
8. Peter	1	0	1	1	1
9. Ruth	1	0	1	1	1
10. Doris	0	1	1	1	0
11. Anna	0	1	0	0	0
12. Helen	0	1	0	0	0
13. Oscar	0	0	0	1	0
14. Ruben	0	0	0	1	0
15. George	0	0	0	0	0
16. Andrea	0	0	0	0	0
17. Richard	0	0	0	0	0
18. Betsy	0	0	0	0	0

L is not an effect modifier on the additive scale.

Causal risk difference for those in stratum L=1

$P(Y_1=1|L=1) - P(Y_0=1|L=1) = 1 - 2/3 = 1/3$

is the same as

Causal risk difference for those in stratum L=0

$P(Y_1=1|L=0)-P(Y_0=1|L=0)=1/3 - 0 = 1/3$

15

Effect modification

- What is an effect modifier?
- **Caveats about its interpretation:**
  - What is the connection to risk factors that we should adjust for?
  - **Dependence on the effect measure of choice**
  - Why does effect modification take place?
- Why care about effect modification?
- How do we analyze the data to examine the presence of effect modifiers?
- Marginal structural models

16

Lesson 2: effect modification and the scale of the effect measure

- The presence or absence of effect modification depends on the scale of the effect measure

- This is why we should not just say M is an effect modifier …

  we should say, e.g., M is an effect modifier on the multiplicative scale or on the additive scale, etc.

17

Patient	M	A	Y	Y_1	Y_0
1. Tom	1	1	1	1	0
2. John	1	1	1	1	0
3. Mary	1	1	1	1	1
4. Amalia	1	1	1	1	1
5. Rick	1	1	1	1	1
6. Rose	1	1	1	1	1
7. Tim	1	0	0	1	0
8. Peter	1	0	1	1	1
9. Ruth	1	0	1	1	1
10. Doris	0	1	1	1	0
11. Anna	0	1	0	0	0
12. Helen	0	1	0	0	0
13. Oscar	0	0	0	1	0
14. Ruben	0	0	0	1	0
15. George	0	0	0	0	0
16. Andrea	0	0	0	0	0
17. Richard	0	0	0	0	0
18. Betsy	0	0	0	0	0

M is not an effect modifier on the additive scale.

Causal risk difference for those in stratum M=1

$P(Y_1=1|M=1) - P(Y_0=1|M=1) = 1 - 2/3 = 1/3$

is the same as

Causal risk difference for those in stratum M=0

$P(Y_1=1|M=0) - P(Y_0=1|M=0) = 1/3 - 0 = 1/3$

18

Patient	M	A	Y	Y_1	Y_0
1. Tom	1	1	1	1	1
2. John	1	1	1	1	0
3. Mary	1	1	1	1	1
4. Amalia	1	1	1	1	1
5. Rick	1	1	1	1	1
6. Rose	1	1	1	1	1
7. Tim	1	0	0	1	0
8. Peter	1	0	1	1	1
9. Ruth	1	0	1	1	1
10. Doris	0	1	1	1	0
11. Anna	0	1	0	0	0
12. Helen	0	1	0	0	0
13. Oscar	0	0	0	1	0
14. Ruben	0	0	0	1	0
15. George	0	0	0	0	0
16. Andrea	0	0	0	0	0
17. Richard	0	0	0	0	0
18. Betsy	0	0	0	0	0

M IS an effect modifier on the multiplicative scale.

Causal risk ratio for those in stratum M=1

$P(Y_1=1|M=1) / P(Y_0=1|M=1)= 1/ (2/3) = 3/2 = 1.5$

is NOT the same as

Causal risk ratio for those in stratum M=0

$P(Y_1=1|M=0)/P(Y_0=1|M=0)= (1/3)/ 0 = \text{infinite}$

19

Qualitative effect modification

- **Definition**: M is a **qualitative effect modifier** on the additive scale for the effect of A on Y if
  - in some strata of M the CRD (or ATE) < 0
  - In other strata of M the CRD (or ATE) > 0

- **Definition**: M is a **qualitative effect modifier** on the multiplicative scale for the effect of A on Y if
  - in some strata of M the CRR < 1
  - In other strata of M the CRR > 1

20

Patient	M	Y_1	Y_0
1. Tom	1	1	0
2. John	1	1	0
3. Mary	1	1	0
4. Amalia	1	1	0
5. Rick	1	0	0
6. Rose	1	1	1
7. Tim	1	1	1
8. Peter	1	0	0
9. Ruth	1	0	0
10. Doris	1	0	0
11. Anna	0	0	1
12. Helen	0	0	1
13. Oscar	0	0	1
14. Ruben	0	0	1
15. George	0	0	0
16. Andrea	0	1	1
17. Richard	0	1	1
18. Betsy	0	0	0
19. Alicia	0	1	1
20. Phoebe	0	1	1

M = 1 South African
M = 0 Candadian

$P(Y_1 = 1 | M=1) = 6/10$

$P(Y_0 = 1 | M=1) = 2/10$

CRD in South Africans = 4/10 > 0

If and only if

$P(Y_1 = 1 | M=0) = 4/10$

$P(Y_0 = 1 | M=0) = 8/10$

CRD in Canadians = - 4/10 <0

21

Patient	M	Y_1	Y_0
1. Tom	1	1	0
2. John	1	1	0
3. Mary	1	1	0
4. Amalia	1	1	0
5. Rick	1	0	0
6. Rose	1	1	1
7. Tim	1	1	1
8. Peter	1	0	0
9. Ruth	1	0	0
10. Doris	1	0	0
11. Anna	0	0	1
12. Helen	0	0	1
13. Oscar	0	0	1
14. Ruben	0	0	1
15. George	0	0	0
16. Andrea	0	1	1
17. Richard	0	1	1
18. Betsy	0	0	0
19. Alicia	0	1	1
20. Phoebe	0	1	1

M is a qualitative effect modifier on the additive scale

$P(Y_1 = 1 | M=1) = 6/10$

$P(Y_0 = 1 | M=1) = 2/10$

CRD in South Africans = 4/10 > 0

If and only if

$P(Y_1 = 1 | M=0) = 4/10$

$P(Y_0 = 1 | M=0) = 8/10$

CRD in Canadians = - 4/10 <0

22

Patient	M	Y_1	Y_0
1. Tom	1	1	0
2. John	1	1	0
3. Mary	1	1	0
4. Amalia	1	1	0
5. Rick	1	0	0
6. Rose	1	1	1
7. Tim	1	1	1
8. Peter	1	0	0
9. Ruth	1	0	0
10. Doris	1	0	0
11. Anna	0	0	1
12. Helen	0	0	1
13. Oscar	0	0	1
14. Ruben	0	0	1
15. George	0	0	0
16. Andrea	0	1	1
17. Richard	0	1	1
18. Betsy	0	0	0
19. Alicia	0	1	1
20. Phoebe	0	1	1

M is a qualitative effect modifier on the multiplicative scale

$P(Y_1 = 1 | M=1) = 6/10$

$P(Y_0 = 1 | M=1) = 2/10$

CRR in South Africans = 3 > 1

If and only if

$P(Y_1 = 1 | M=0) = 4/10$

$P(Y_0 = 1 | M=0) = 8/10$

CRR in Canadians = ½ < 1

23

Patient	M	Y_1	Y_0
1. Tom	1	1	0
2. John	1	1	0
3. Mary	1	1	0
4. Amalia	1	1	0
5. Rick	1	0	0
6. Rose	1	1	1
7. Tim	1	1	1
8. Peter	1	0	0
9. Ruth	1	0	0
10. Doris	1	0	0
11. Anna	0	0	1
12. Helen	0	0	1
13. Oscar	0	0	1
14. Ruben	0	0	1
15. George	0	0	0
16. Andrea	0	1	1
17. Richard	0	1	1
18. Betsy	0	0	0
19. Alicia	0	1	1
20. Phoebe	0	1	1

Mathematical Fact:

M is a qualitative effect modifier on the additive scale

If and only if

M is a qualitative effect modifier on the multiplicative scale

24

Patient	M	A	Y	Y_1	Y_0
1. Tom	1	1	1	1	0
2. John	1	1	1	1	0
3. Mary	1	1	1	1	1
4. Amalia	1	1	1	1	1
5. Rick	1	1	1	1	1
6. Rose	1	1	1	1	1
7. Tim	1	0	0	1	0
8. Peter	1	0	1	1	1
9. Ruth	1	0	1	1	1
10. Doris	0	1	1	1	0
11. Anna	0	1	0	0	0
12. Helen	0	1	0	0	0
13. Oscar	0	0	0	1	0
14. Ruben	0	0	0	1	0
15. George	0	0	0	0	0
16. Andrea	0	0	0	0	0
17. Richard	0	0	0	0	0
18. Betsy	0	0	0	0	0

M is not an effect modifier on the additive scale.

Causal risk difference for those in stratum M=1

$P(Y_1=1|M=1) - P(Y_0=1|M=1) = 1 - 2/3 = 1/3$

is the same as

Causal risk difference for those in stratum M=0

$P(Y_1=1|M=0)-P(Y_0=1|M=0)=1/3 - 0 = 1/3$

25

Patient	M	A	Y	Y_1	Y_0
1. Tom	1	1	1	1	0
2. John	1	1	1	1	0
3. Mary	1	1	1	1	1
4. Amalia	1	1	1	1	1
5. Rick	1	1	1	1	1
6. Rose	1	1	1	1	1
7. Tim	1	0	0	1	0
8. Peter	1	0	1	1	1
9. Ruth	1	0	1	1	1
10. Doris	0	1	1	1	0
11. Anna	0	1	0	0	0
12. Helen	0	1	0	0	0
13. Oscar	0	0	0	1	0
14. Ruben	0	0	0	1	0
15. George	0	0	0	0	0
16. Andrea	0	0	0	0	0
17. Richard	0	0	0	0	0
18. Betsy	0	0	0	0	0

M IS an effect modifier on the multiplicative scale.

Causal risk ratio for those in stratum M=1

$P(Y_1=1|M=1) / P(Y_0=1|M=1)= 1/ (2/3) = 3/2 = 1.5$

is NOT the same as

Causal risk ratio for those in stratum M=0

$P(Y_1=1|M=0)/P(Y_0=1|M=0)= (1/3)/ 0$
= infinite

26

More on Lesson 2…

- If M is a **non-qualitative** effect modifier on the **multiplicative** scale, then
  - M is a **non-qualitative** effect modifier on the **additive** scale,

  or
  - M is NOT an effect modifier on the **additive** scale

- If M is a **non-qualitative** effect modifier on the **additive** scale, then
  - M is a **non-qualitative** effect modifier on the **multiplicative** scale,

  or
  - M is NOT an effect modifier on the **multiplicative** scale

27

Even more on Lesson 2…

- If M is a **NOT** an effect modifier on the **multiplicative** scale,

but

- the baseline risk (i.e. risk under no trx), **differs** across strata of M, then
  - M is a non-qualitative effect modifier on the **additive** scale

- If M is a **NOT** an effect modifier on the **additive** scale,

but

- the baseline risk (i.e. risk under no trx), **differs** across strata of M, then
  - M is a non-qualitative effect modifier on the **multiplicative** scale

28

Effect modification

- What is an effect modifier?
- **Caveats about its interpretation:**
  - What is the connection to risk factors that we should adjust for?
  - Dependence on the effect measure of choice
  - **Why does effect modification take place?**
- Why care about effect modification?
- How do we analyze the data to examine the presence of effect modifiers?
- Marginal structural models

29

Patient	M	Y_1	Y_0
1. Tom	1	1	0
2. John	1	1	0
3. Mary	1	1	0
4. Amalia	1	1	0
5. Rick	1	0	0
6. Rose	1	1	1
7. Tim	1	1	1
8. Peter	1	0	0
9. Ruth	1	0	0
10. Doris	1	0	0
11. Anna	0	0	1
12. Helen	0	0	1
13. Oscar	0	0	1
14. Ruben	0	0	1
15. George	0	0	0
16. Andrea	0	1	1
17. Richard	0	1	1
18. Betsy	0	0	0
19. Alicia	0	1	1
20. Phoebe	0	1	1

M = 1 South Africans
M = 0 Canadians

CRD in South Africans = 4/10
CRD in Canadians = - 4/10

We should not ascribe a causal interpretation to the effect modifiers.

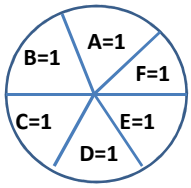
Effect modifiers may be **proxies** for other factors that **interact** with exposure to either bring about or prevent the disease .

30

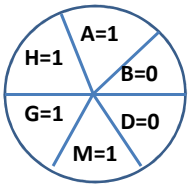
The sufficient cause model

31

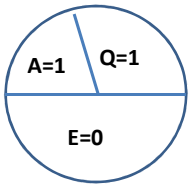
Sufficient cause model
(Mackie, 1965, Rothman, 1976)



Sufficient cause: one causal mechanism with **component causes**: A=1, B=1, C=1, D=1, E=1, F=1



Sufficient cause with **component causes** A=1, B=0, M=1, D=0 G=1, H=1



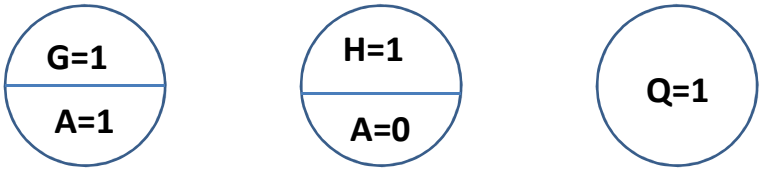
Sufficient cause with component causes A=1, Q=1, E=0

1. All components causes of a given sufficient cause are necessary for the activation of that sufficient cause. i.e. no component cause is redundant.

2. Activation of a sufficient cause yields the onset of the event or outcome of interest, say, death.

32

A simplified sufficient cause model in our transplant example



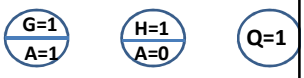
G =1 if surgery lacks special monitoring system and patient has delicate baseline conditions
G = 0 otherwise

H = 1 if patient has ejection fraction < 0.20
H = 0 otherwise

Q = 1 if patient has certain co-morbidities
Q = 0 otherwise

33

Patient	M	Y_1	Y_0	Type	Component Causes
1. Tom	1	1	0	causative	G=1 and H=0 and Q=0
2. John	1	1	0	causative	G=1 and H=0 and Q=0
3. Mary	1	1	0	causative	G=1 and H=0 and Q=0
4. Amalia	1	1	0	causative	G=1 and H=0 and Q=0
5. Rick	1	0	0	immune	G=0 and H=0 and Q=0
6. Rose	1	1	1	doomed	(G=1 and H=1) or Q=1
7. Tim	1	1	1	doomed	(G=1 and H=1) or Q=1
8. Peter	1	0	0	immune	G=0 and H=0 and Q=0
9. Ruth	1	0	0	immune	G=0 and H=0 and Q=0
10. Doris	1	0	0	immune	G=0 and H=0 and Q=0
11. Anna	0	0	1	preventative	G=0 and H=1 and Q=0
12. Helen	0	0	1	preventative	G=0 and H=1 and Q=0
13. Oscar	0	0	1	preventative	G=0 and H=1 and Q=0
14. Ruben	0	0	1	preventative	G=0 and H=1 and Q=0
15. George	0	0	0	immune	G=0 and H=0 and Q=0
16. Andrea	0	1	1	doomed	(G=1 and H=1) or Q=1
17. Richard	0	1	1	doomed	(G=1 and H=1) or Q=1
18. Betsy	0	0	0	immune	G=0 and H=0 and Q=0
19. Alicia	0	1	1	doomed	(G=1 and H=1) or Q=1
20. Phoebe	0	1	1	doomed	(G=1 and H=1) or Q=1



G =1 if surgery lacks special monitoring system and patient has delicate baseline conditions

H = 1 if patient has ejection fraction < 0.20

Q = 1 if patient has certain co-morbidities

Effect modification arises because the types are imbalanced across strata of the effect modifier.
The imbalance reflects different distributions of component causes.

34

Effect modification

- What is an effect modifier?
- Caveats about its interpretation:
  - What is the connection to risk factors that we should adjust for?
  - Dependence on the effect measure of choice
  - Why does effect modification take place?
- **Why care about effect modification?**
- How do we analyze the data to examine the presence of effect modifiers?
- Marginal structural models

35

Why care about effect modification?

- To learn about who will
  - most benefit from treatment, and
  - be most harmed by treatment.
- To facilitate calculations of treatment effects in different populations with different distributions of the effect modifiers.
- As a mean to get clues about the mechanisms ("interactions") leading to disease.

36

Conditional vs marginal effects

- **Conditional effects**:
  - **Stratified analysis**: you compute the effect measure of your choice separately for each level of M.
    - E.g. One CRD for South Africans, one CRD for Canadians
- **Marginal effects:**
  - **Overall analysis**: you compute the effect measure of your choice over the entire population
    - E.g. One CRD for the conglomerate of all South Africans and Canadians.

37

Which causal questions is relevant for your problem?

- Sometimes, you may not care about the variations in trx effects across subpopulations, so the **marginal** trx effect is your target:
  - E.g. if you are a policy maker evaluating the benefits of implementing a water fluoridation program.
- Other times, you will want to compute **conditional** trx effects:
  - E.g. when the intervention can be targeted to different subsets of the population. For example, who should be prescribed with estrogen replacement therapy?

38

Effect modification

- What is an effect modifier?
- Caveats about its interpretation:
  - What is the connection to risk factors that we should adjust for?
  - Dependence on the effect measure of choice
  - Why does effect modification take place?
- Why care about effect modification?
- **How do we analyze the data to examine the presence of effect modifiers?**
- Marginal structural models

39

Analyst's table					Oracle's table	
Patient	**M**	**L**	**A**	**Y**	**Y1**	**Y0**
1. Tom	1	1	1	1	1	0
2. John	1	0	1	1	1	0
3. Mary	1	1	1	1	1	0
4. Amalia	1	0	1	1	1	0
5. Rick	1	1	0	0	0	0
6. Rose	1	0	0	1	1	1
7. Tim	1	1	0	1	1	1
8. Peter	1	1	0	0	0	0
9. Ruth	1	1	1	0	0	0
10. Doris	1	1	1	0	0	0
11. Anna	0	0	1	0	0	1
12. Helen	0	0	1	0	0	1
13. Oscar	0	0	1	0	0	1
14. Ruben	0	1	1	0	0	1
15. George	0	0	0	0	0	0
16. Andrea	0	0	0	1	1	1
17. Richard	0	0	0	1	1	1
18. Betsy	0	1	0	0	0	0
19. Alicia	0	0	0	1	1	1
20. Phoebe	0	1	0	1	1	1

40

Analyst's table				
Patient	**M**	**L**	**A**	**Y**
1. Tom	1	1	1	1
2. John	1	0	1	1
3. Mary	1	1	1	1
4. Amalia	1	0	1	1
5. Rick	1	1	0	0
6. Rose	1	0	0	1
7. Tim	1	1	0	1
8. Peter	1	1	0	0
9. Ruth	1	1	1	0
10. Doris	1	1	1	0
11. Anna	0	0	1	0
12. Helen	0	0	1	0
13. Oscar	0	0	1	0
14. Ruben	0	1	1	0
15. George	0	0	0	0
16. Andrea	0	0	0	1
17. Richard	0	0	0	1
18. Betsy	0	1	0	0
19. Alicia	0	0	0	1
20. Phoebe	0	1	0	1

41

Stratification to examine effect modification

- **In an observational study, in order to examine the presence of effect modification by M, ideally you should:**
  - First, stratify by the levels of M, i.e. divide your population into subpopulations, one for each level of M.
  - Next, separately in each stratum, standardize or compute IPTW adjusting by the risk factors L that you believe are needed to make your study emulate a conditionally randomized trial.

42

Being realistic…

- In the last slide I said "ideally you should …" because in reality you will hardly ever be able to do it…
  - If M takes many levels, you will not have enough data to carry out the stratified analysis
  - If L takes many levels, you will not have enough data to carry out the standardization
- I will postpone until later today the discussion of marginal structural models to handle the sparse data problem.

43

Stratification as a form of adjustment

- We have used **stratification** to examine **effect modification** by a variable **M**
- But **stratification** is often used as an alternative to **standardization** as a form of **adjustment for risk factors L**
- In fact, many think of the terms **stratification** and **adjustment** as synonymous.

44

Stratification as a form of adjustment

- Suppose you decided that, say,

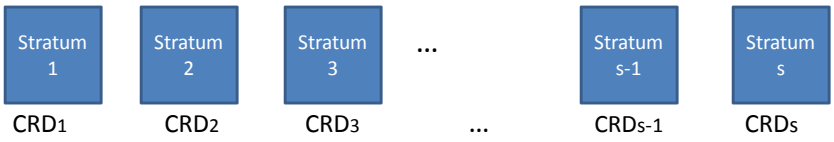
  L = (age, sbp, race)

  are all the risk factors that you need to adjust for in order to make your study emulate a conditionally randomized trial. All together there are s different strata defined by the different combinations of age, sbp and race categories

1. in each stratum j you compute the crude risks in the exposed and unexposed and compare them with the effect measure of your choice, say with the causal risk difference, CRDj
2. because you believe that **L** are all the relevant risk factors, you interpret the stratum specific crude effect measures as causal effect measures.
3. So… you have actually computed one causal effect measure per stratum of **L**. Now, you can examine if these differ across strata. That is, you can examine if there is effect modification by **L**.
4. The distinction between **M** and **L** is that because you believe that conditional on **L** there is exchangeability, you do not need to standardize or use iptw weighting in each stratum of **L** in order to compute the "causal effect measures". The crude measures "are" your causal effect measure estimators.

45

Stratification vs standardization by L

- 

- You can either …
  - stop your analysis here, in which case, you would be reporting **conditional effects**, i.e. the separate effects for each stratum of **L**
  - or you can go one further step and take the weighed average of the stratum specific effect measures, i.e. compute the standardized CRD in the entire population

    $CRD_1 \times P(L=1) + CRD_2 \times P(L=2) + \ldots + CRD_s \times P(L=s)$

    in which case you would be reporting the **marginal effects**.

46

Stratification vs Standardization Summary

- **You believe M** is **not enough** to emulate a cond. randomized trial.
- To examine **effect modification** by **M** you stratify on **M** but …
  - in each stratum you **standardize** adjusting for **L** pretending that the reference population is the subpopulation in the stratum.
  - The **standardized** risks in each stratum are the counterfactual risks. Now you combine them using the effect measure of your choice to yield the causal effect measure in each stratum.
- **You believe L is enough** to emulate a cond. randomized trial.
- To examine **effect modification** by **L** you **stratify** on **L** and …
  - in each stratum you **do NOT standardize,** you just compute the crude effect measures and you interpret them as causal.
- If you care to compute the **marginal effect in the entire population**, you need to do **standardization** with reference the entire population , or IPTW, adjusting for **L**.

47

Dependence of the relationship between marginal and conditional effect measures on the scale of measurement

48

Patient	M	Y_1	Y_0
1. Tom	1	1	1
2. John	1	1	0
3. Mary	1	1	0
4. Amalia	1	0	0
5. Rick	1	1	1
6. Rose	1	0	1
7. Tim	1	1	0
8. Peter	1	1	0
9. Ruth	1	0	1
10. Doris	1	0	0
11. Anna	0	1	0
12. Helen	0	0	0
13. Oscar	0	0	0
14. Ruben	0	0	0
15. George	0	0	1
16. Andrea	0	1	0
17. Richard	0	0	0
18. Betsy	0	0	0
19. Alicia	0	1	0
20. Phoebe	0	1	1

When M = 1

$CRD_1 = P(Y_1=1|M=1) - P(Y_0=1|M=1)$

$= 6/10 - 4/10 = 2/10$

When M = 0

$CRD_0 = P(Y_1=1|M=0) - P(Y_0=1|M=0)$

$= 4/10 - 2/10 = 2/10$

No effect modification on the additive scale

49

Patient	M	Y_1	Y_0
1. Tom	1	1	1
2. John	1	1	0
3. Mary	1	1	0
4. Amalia	1	0	0
5. Rick	1	1	1
6. Rose	1	0	1
7. Tim	1	1	0
8. Peter	1	1	0
9. Ruth	1	0	1
10. Doris	1	0	0
11. Anna	0	1	0
12. Helen	0	0	0
13. Oscar	0	0	0
14. Ruben	0	0	0
15. George	0	0	1
16. Andrea	0	1	0
17. Richard	0	0	0
18. Betsy	0	0	0
19. Alicia	0	1	0
20. Phoebe	0	1	1

When M = 1

$CRD_1 = P(Y_1=1|M=1) - P(Y_0=1|M=1)$

$= 6/10 - 4/10 = 2/10$

When M = 0

$CRD_0 = P(Y_1=1|M=0) - P(Y_0=1|M=0)$

$= 4/10 - 2/10 = 2/10$

No effect modification on the additive scale

Stratum specific CRD equal to overall (marginal) CRD:

In the entire population

$CRD = P(Y_1=1) - P(Y_0=1)$

$= 10/20 - 6/20 = 2/10$

50

Patient	M	Y_1	Y_0
1. Tom	1	1	1
2. John	1	1	0
3. Mary	1	1	0
4. Amalia	1	0	0
5. Rick	1	1	1
6. Rose	1	0	1
7. Tim	1	1	0
8. Peter	1	1	0
9. Ruth	1	0	1
10. Doris	1	0	0
11. Anna	0	1	0
12. Helen	0	0	0
13. Oscar	0	0	0
14. Ruben	0	0	0
15. George	0	0	1
16. Andrea	0	0	0
17. Richard	0	0	0
18. Betsy	0	0	0
19. Alicia	0	1	0
20. Phoebe	0	1	1

When M = 1

CRR1 = $P(Y_1=1|M=1)$ / $P(Y_0=1|M=1)$

= (6/10) / (4/10) = 1.5

When M = 0

CRD_0 = $P(Y_1=1|M=0)$ / $P(Y_0=1|M=0)$

= (3/10) / (2/10) = 1.5

No effect modification on the multiplicative scale

51

Patient	M	Y_1	Y_0
1. Tom	1	1	1
2. John	1	1	0
3. Mary	1	1	0
4. Amalia	1	0	0
5. Rick	1	1	1
6. Rose	1	0	1
7. Tim	1	1	0
8. Peter	1	1	0
9. Ruth	1	0	1
10. Doris	1	0	0
11. Anna	0	1	0
12. Helen	0	0	0
13. Oscar	0	0	0
14. Ruben	0	0	0
15. George	0	0	1
16. Andrea	0	0	0
17. Richard	0	0	0
18. Betsy	0	0	0
19. Alicia	0	1	0
20. Phoebe	0	1	1

When M = 1

CRR1 = $P(Y_1=1|M=1)$ / $P(Y_0=1|M=1)$

= (6/10) / (4/10) = 1.5

When M = 0

CRD_0 = $P(Y_1=1|M=0)$ / $P(Y_0=1|M=0)$

= (3/10) / (2/10) = 1.5

No effect modification on the multiplicative scale

Stratum specific CRR equal to overall (marginal) CRR:

In the entire population

CRR = $P(Y_1=1)$ / $P(Y_0=1)$

= (9/20) / (6/20) = 1.5

52

Patient	M	Y_1	Y_0
1. Tom	1	1	1
2. John	1	1	0
3. Mary	1	1	1
4. Amalia	1	0	0
5. Rick	1	1	1
6. Rose	1	1	0
7. Tim	1	1	0
8. Peter	1	1	1
9. Ruth	1	1	0
10. Doris	1	0	0
11. Anna	0	1	0
12. Helen	0	0	0
13. Oscar	0	1	0
14. Ruben	0	0	0
15. George	0	1	1
16. Andrea	0	1	0
17. Richard	0	0	0
18. Betsy	0	0	0
19. Alicia	0	1	0
20. Phoebe	0	1	1

When M = 1

$$COR_1 = \frac{P(Y_1=1|M=1) / [1- P(Y_1=1|M=1)]}{P(Y_0=1|M=1) / [1- P(Y_0=1|M=1)]} = \frac{(8/10) / (2/10)}{(4/10) / (6/10)} = 6$$

When M = 0

$$COR_0 = \frac{P(Y_1=1|M=0) / [1- P(Y_1=1|M=0)]}{P(Y_0=1|M=0) / [1- P(Y_0=1|M=0)]} = \frac{(6/10) / (4/10)}{(2/10) / (8/10)} = 6$$

53

Patient	M	Y_1	Y_0
1. Tom	1	1	1
2. John	1	1	0
3. Mary	1	1	1
4. Amalia	1	0	0
5. Rick	1	1	1
6. Rose	1	1	0
7. Tim	1	1	0
8. Peter	1	1	1
9. Ruth	1	1	0
10. Doris	1	0	0
11. Anna	0	1	0
12. Helen	0	0	0
13. Oscar	0	1	0
14. Ruben	0	0	0
15. George	0	1	1
16. Andrea	0	1	0
17. Richard	0	0	0
18. Betsy	0	0	0
19. Alicia	0	1	0
20. Phoebe	0	1	1

When M = 1

$$COR_1 = \frac{P(Y_1=1|M=1) / [1- P(Y_1=1|M=1)]}{P(Y_0=1|M=1) / [1- P(Y_0=1|M=1)]} = \frac{(8/10) / (2/10)}{(4/10) / (6/10)} = 6$$

When M = 0

$$COR_0 = \frac{P(Y_1=1|M=0) / [1- P(Y_1=1|M=0)]}{P(Y_0=1|M=0) / [1- P(Y_0=1|M=0)]} = \frac{(6/10) / (4/10)}{(2/10) / (8/10)} = 6$$

But stratum specific COR NOT equal to overall (marginal) COR:

In the entire population

$$COR = \frac{P(Y_1=1) / [1- P(Y_1=1)]}{P(Y_0=1) / [1- P(Y_0=1)]} = \frac{(14/20) / (6/20)}{(6/20) / (14/20)} = 5.44$$

54

Mathematical fact

Additive scale

If
$CRD_1 = CRD_2 = \ldots = CRD_S =$ some number c

then CRD = c

Also,
If
$ATE_1 = ATE_2 = \ldots = ATE_S =$ some number c

then ATE = c

Multiplicative scale

If
$CRR_1 = CRR_2 = \ldots = CRR_S =$ some number c

then CRR = c

Odds ratio scale

If
$COR_1 = COR_2 = \ldots = COR_S =$ some number c

then COR is closer to 1

- Causal odds ratio in the entire population **closer to 1 than the** common stratum specific causal odds ratio

55

Patient	M	Y_1	Y_0
1. Tom	1	1	1
2. John	1	1	0
3. Mary	1	1	0
4. Amalia	1	0	0
5. Rick	1	1	1
6. Rose	1	0	1
7. Tim	1	1	0
8. Peter	1	1	0
9. Ruth	1	0	1
10. Doris	1	0	0
11. Anna	0	1	0
12. Helen	0	0	0
13. Oscar	0	1	0
14. Ruben	0	1	0
15. George	0	0	1
16. Andrea	0	1	0
17. Richard	0	1	0
18. Betsy	0	0	1
19. Alicia	0	1	0
20. Phoebe	0	1	1

When M = 1

$CRD_1 = P(Y_1=1|M=1) - P(Y_0=1|M=1)$
$= 6/10 - 4/10 = 2/10$

When M = 0

$CRD_0 = P(Y_1=1|M=0) - P(Y_0=1|M=0)$
$= 7/10 - 3/10 = 4/10$

$CRD_1 \times P(M=1) + CRD_0\ P(M=0) =$
$(2/10) \times (10/20) + (4/10) \times (10/20) = 3/10$

56

Patient	M	Y_1	Y_0
1. Tom	1	1	1
2. John	1	1	0
3. Mary	1	1	0
4. Amalia	1	0	0
5. Rick	1	1	1
6. Rose	1	0	1
7. Tim	1	1	0
8. Peter	1	1	0
9. Ruth	1	0	1
10. Doris	1	0	0
11. Anna	0	1	0
12. Helen	0	0	0
13. Oscar	0	1	0
14. Ruben	0	1	0
15. George	0	0	1
16. Andrea	0	1	0
17. Richard	0	1	0
18. Betsy	0	0	1
19. Alicia	0	1	0
20. Phoebe	0	1	1

When M = 1

$CRD_1 = P(Y_1=1|M=1) - P(Y_0=1|M=1)$

$= 6/10 - 4/10 = 2/10$

When M = 0

$CRD_0 = P(Y_1=1|M=0) - P(Y_0=1|M=0)$

$= 7/10 - 3/10 = 4/10$

$CRD_1 \times P(M=1) + CRD_0 P(M=0) =$

$(2/10) \times (10/20) + (4/10) \times (10/20) = 3/10$

Overall (marginal) CRD equal to weighted average of stratum specific CRD

In the entire population

$CRD = P(Y_1=1) - P(Y_0=1)$

$= 13/20 - 7/20 = 3/10$

57

Patient	M	Y_1	Y_0
1. Tom	1	1	1
2. John	1	1	0
3. Mary	1	1	0
4. Amalia	1	0	0
5. Rick	1	1	0
6. Rose	1	0	0
7. Tim	1	1	0
8. Peter	1	1	0
9. Ruth	1	0	1
10. Doris	1	0	0
11. Anna	0	1	0
12. Helen	0	1	0
13. Oscar	0	1	0
14. Ruben	0	1	0
15. George	0	0	1
16. Andrea	0	1	0
17. Richard	0	1	0
18. Betsy	0	0	0
19. Alicia	0	1	0
20. Phoebe	0	1	1

When M = 1

$CRR_1 = P(Y_1=1|M=1) - P(Y_0=1|M=1)$

$= (6/10) / (2/10) = 3$

When M = 0

$CRR_0 = P(Y_1=1|M=0) / P(Y_0=1|M=0)$

$= (8/10) / (2/10) = 4$

$CRD_1 \times P(M=1) + CRD_0 P(M=0) =$

$3 \times (10/20) + 4 \times (10/20) = 3.5$

58

Patient	M	Y_1	Y_0
1. Tom	1	1	1
2. John	1	1	0
3. Mary	1	1	0
4. Amalia	1	0	0
5. Rick	1	1	0
6. Rose	1	0	0
7. Tim	1	1	0
8. Peter	1	1	0
9. Ruth	1	0	1
10. Doris	1	0	0
11. Anna	0	1	0
12. Helen	0	1	0
13. Oscar	0	1	0
14. Ruben	0	1	0
15. George	0	0	1
16. Andrea	0	1	0
17. Richard	0	1	0
18. Betsy	0	0	0
19. Alicia	0	1	0
20. Phoebe	0	1	1

When M = 1

$CRR_1 = P(Y_1=1|M=1) - P(Y_0=1|M=1)$

$= (6/10) / (2/10) = 3$

When M = 0

$CRR_0 = P(Y_1=1|M=0) / P(Y_0=1|M=0)$

$= (8/10) / (2/10) = 4$

$CRD_1 \times P(M=1) + CRD_0 \, P(M=0) =$

$3 \times (10/20) + 4 \times (10/20) = 3.5$

Overall (marginal) CRD equal to weighted average of stratum specific CRD

In the entire population

$CRR = P(Y_1=1) - P(Y_0=1)$

$= (14/20) / (4/20) = 3.5$

59

Patient	M	Y_1	Y_0
1. Tom	1	1	1
2. John	1	1	0
3. Mary	1	1	1
4. Amalia	1	0	0
5. Rick	1	1	1
6. Rose	1	1	0
7. Tim	1	1	0
8. Peter	1	1	1
9. Ruth	1	1	0
10. Doris	1	0	0
11. Anna	0	1	0
12. Helen	0	0	0
13. Oscar	0	1	0
14. Ruben	0	0	0
15. George	0	1	1
16. Andrea	0	1	0
17. Richard	0	0	0
18. Betsy	0	0	0
19. Alicia	0	0	0
20. Phoebe	0	1	1

When M = 1

$COR1 = \frac{P(Y1=1|M=1) / [1- P(Y1=1|M=1)]}{P(Y0=1|M=1) / [1- P(Y0=1|M=1)]}$

$= \frac{(8/10) / (2/10)}{(4/10) / (6/10)} = 6$

When M = 0

$COR0 = \frac{P(Y1=1|M=0) / [1- P(Y1=1|M=0)]}{P(Y0=1|M=0) / [1- P(Y0=1|M=0)]}$

$= \frac{(5/10) / (5/10)}{(2/10) / (8/10)} = 4$

COR1 x P(L=1) + COR0 x P(L=0) =

6 x 10/20 + 4 x 10/20 = 5

60

Patient	M	Y_1	Y_0
1. Tom	1	1	1
2. John	1	1	0
3. Mary	1	1	1
4. Amalia	1	0	0
5. Rick	1	1	1
6. Rose	1	1	0
7. Tim	1	1	0
8. Peter	1	1	1
9. Ruth	1	1	0
10. Doris	1	0	0
11. Anna	0	1	0
12. Helen	0	0	0
13. Oscar	0	1	0
14. Ruben	0	0	0
15. George	0	1	1
16. Andrea	0	1	0
17. Richard	0	0	0
18. Betsy	0	0	0
19. Alicia	0	0	0
20. Phoebe	0	1	1

When M = 1

$$COR1 = \frac{P(Y1=1|M=1)\ /\ [1- P(Y1=1|M=1)]}{P(Y0=1|M=1)\ /\ [1- P(Y0=1|M=1)]} = \frac{(8/10)\ /\ (2/10)}{(4/10)\ /\ (6/10)} = 6$$

When M = 0

$$COR0 = \frac{P(Y1=1|M=0)\ /\ [1- P(Y1=1|M=0)]}{P(Y0=1|M=0)\ /\ [1- P(Y0=1|M=0)]} = \frac{(5/10)\ /\ (5/10)}{(2/10)\ /\ (8/10)} = 4$$

COR1 x P(L=1) + COR0 x P(L=0) =
6 x 10/20 + 4 x 10/20 = 5

But overall (marginal) COR is not the average of stratum specific CORs

In the entire population

$$COR = \frac{P(Y_1=1)\ /\ [1- P(Y_1=1)]}{P(Y_0=1)\ /\ [1- P(Y_0=1)]} = \frac{(13/20)\ /\ (7/20)}{(6/20)\ /\ (14/20)} = 4.33$$

61

Mathematical fact

Additive scale

CRD = causal risk diff in the entire population

CRD_j = causal risk diff in stratum j

$CRD = \sum CRD_j\ P(L=j)$

Multiplicative scale

CRR = causal risk ratio in the entire population

CRR_j = causal risk ratio in stratum j

$CRR = \sum CRR_j\ P(L=j)$

Odds ratio scale

COR = causal odds ratio in the entire poplulation

COR_j = causal odds ration in stratum j

$COR \neq \sum COR_j\ P(L=j)$

62

Collapsible vs non-collapsible effect measures

- An effect measure is **collapsible** when the effect measure in the entire population is an average of the effect measures of all the strata
- We have seen that the **causal** risk difference and the causal risk ratio are collapsible, but the causal odds ratio is not
- It is also a mathematical fact that the **crude** risk difference and the crude risk ratio are collapsible measures, but the crude odds ratio is not.

63

Patient	L	A	Y
1. Tom	1	0	0
2. John	1	0	0
3. Mary	1	0	0
4. Amalia	1	0	0
5. Rick	1	0	1
6. Rose	1	1	0
7. Tim	1	1	1
8. Peter	1	1	1
9. Ruth	1	1	0
10. Doris	1	1	1
11. Anna	0	0	0
12. Helen	0	0	0
13. Oscar	0	0	1
14. Ruben	0	0	1
15. George	0	0	0
16. Andrea	0	1	1
17. Richard	0	1	1
18. Betsy	0	1	1
19. Alicia	0	1	0
20. Phoebe	0	1	1

Data comes from a randomized trial. There is exchangeability, no need to adjust for L

Crude odds ratio in stratum L = 1

$$\frac{(3/5)}{(2/5)} = 6$$
$$\frac{(1/5)}{(4/5)}$$

Crude odds ratio in stratum L = 0

$$\frac{(4/5)}{(1/5)} = 6$$
$$\frac{(2/5)}{(3/5)}$$

Crude marginal odds ratio

$$\frac{(7/10)}{(3/10)} = 5.44$$
$$\frac{(3/10)}{(7/10)}$$

64

Being realistic…

- In the last slide I said "ideally you should …" because in reality you will hardly ever be able to do it…
  - If M takes many levels, you will not have enough data to carry out the stratified analysis
  - If L takes many levels, you will not have enough data to carry out the standardization

65

Effect modification

- What is an effect modifier?
- Caveats about its interpretation:
  - What is the connection to risk factors that we should adjust for?
  - Dependence on the effect measure of choice
  - Why does effect modification take place?
- Why care about effect modification?
- How do we analyze the data to examine the presence of effect modifiers?
- **Marginal structural models**

66

The EFFECT study revisited

- Suppose that we want to investigate if **age** is an **effect modifier** for the effect of **Statin** on, say, **three year post-hospitalization serum lipids** (on the additive scale).
- Here
  M = age (measured on years)
  L = age, heart rate, creatinine level, potasium level
  A = statin
  Y = three-year post hospitalization serum lipids
- We want to compute for each age group

$$ATE(age) = E(Y_1|age) - E(Y_0|age)$$

67

Marginal structural models

- Two problems in the EFFECT STUDY
  1. We may run out of people in some age groups
  2. We can't do standardization by L in each stratum of age, because L has several continuous components.
- Solutions
  1. To resolve issue (1) we specify a marginal structural model.
  2. To resolve issue (2) we fit the marginal structural model by IPW using parametric models for the propensity score.

68

IPW estimation of β when L has many components: rational

- Weighting each person by

  w = 1/P(A|L)

- creates a pseudo study in which everybody receives A=1 and everybody receives A=0. So, in the pseudo study there is exchangeability and

$$E(Y_a|age) = E(Y|A = a, age)$$

- So, under the MSM model, in the pseudo-study

$$E(Y|A, age) = \beta_0 + \beta_1 A + \beta_2 age + \beta_3 age * A$$

- The last equation is a standard linear regression model.
  - The method of least squares is known to give unbiased estimators of the linear regression parameters.
- So we should estimate β by **weighted** least squares with weights w.
- Since we don't know P(A|L) and we can't estimate it by stratification on L due to L having many components, we must first **estimate P(A|L) under a parametric model**

69

IPW estimation of β when L has many components: algorithm

- ***STEP 1. Estimate the propensity score under parametric models for P(A|L), e.g.***

$$\log\frac{P(A=1|age,hr,creat,potas)}{P(A=0|age,hr,creat,potas)} = \alpha 0 + \alpha 1 age + \alpha 2 hr + \alpha 3 creat + \alpha 4 potas$$

  - ***For each treated person compute w=1/fitted value and estimate $E(Y_1)$ with the weighted mean of the treated.***
  - ***For each untreated person compute w=1/(1-fitted value) and estimate E(Y0) with the weighted mean of the untreated.***

- ***STEP 2. Run weighted least squares with outcome Y and covariates A and L (according to your MSM model) and weights w.***
- The estimators of the model parameters are your IPW estimators of the parameters of the marginal structural model

70

Stabilized: IPW estimation of β when L has many components: algorithm

- ***STEP 1. Estimate the propensity score under parametric models for P(A|L), e.g.***

$$\log\frac{P(A=1|age,hr,creat,potas)}{P(A=0|age,hr,creat,potas)} = \alpha 0 + \alpha 1 age + \alpha 2 hr + \alpha 3 creat + \alpha 4 potas$$

- ***Step 2. Estimate the propensity score under parametric models for P(A|M), e.g.***

$$\log\frac{P(A=1|age)}{P(A=0|age)} = \delta 0 + \delta 1 age$$

- ***Step 3.***
  - ***For each treated person compute***

$$sw = \frac{fitted\ value\ of\ P(A=1|M)}{fitted\ value\ of\ P(A=1|L)}$$

  - ***For each untreated person compute***

$$sw = \frac{fitted\ value\ of\ P(A=0|M)}{fitted\ value\ of\ P(A=0|L)}$$

- ***STEP 4. Run weighted least squares with outcome Y and covariates A and M (according to your MSM model) and weights sw.***
- ***The estimators of the model parameters are your IPW estimators of***

71

Validity of IPTW with stabilized weights

- If there is
  1. Unconditional exchangeability, or
  2. Conditional exchangeability given M,
- It holds that

$$E(Y_a \mid M) = E(Y \mid A=a, M)$$

So fitting standard regression models is ok. under (1) or (2)

- Weighting by 1/w = 1/f(A|L) creates a pseudo-population in which (1) holds. ***So plain IPTW ok***
- Weighting by 1/sw = f(A|M)/f(A|L) creates a pseudo-population in which (2) holds. ***So stabilized IPTW ok***

Why stabilized IPTW should reduce variability?

- Variability of plain IPTW estimators caused by occasional presence small weights relative to the rest
- Stabilized IPTW makes the distribution of the weights less spread. Idea:

- John has w small. This means that John's estimated P(A=1|L), i.e. his $f_1(L)$, is small
- Then it is likely that his estimated p(A=1|M), i.e. $f_2(M)$ is also small
- Since $f_1(L)/f_2(M)$ = small/small = not so small, then ***John's sw is larger than John's w***

- Peter has w large. This means that Peter's estimated P(A=1|L), i.e. his $f_1(L)$, is large
- Then it is likely that his estimated p(A=1|M), i.e. $f_2(M)$ is also large
- Since $f_1(L)/f_2(M)$ = large/large = not so large, then ***John's sw is smaller than Peter's w***

Peter and John's w's very different but their sw's are not so different

Inverse weighted probability (IPW) Stabilized- Linear

Use t_ipw.ado

```
t_ipw,outcome(string) treatvar(string) pvars(varlist)
[ovars(varlist)] [pvarstab(varlist)] [stabilized] [binary]
[cvars(varlist) censored] [bootstrap]
```

```
• t_ipw, outcome(cardbill) pvars(stent acutemi ejecfrac ves1proc p_inter*) treatvar( abcix) ///
  ovars(ejecfrac ) bootstrap rep(200)

• t_ipw, outcome(cardbill) pvars(stent acutemi ejecfrac ves1proc p_inter*) ///
treatvar( abcix) ovars(ejecfrac ) pvarstab(ejecfrac ) stabilized bootstrap rep(200)
```

Without Stabilization

beta1*(abcix)*	Boots. sdt Err.	[95% Conf.	Interval]	(N)	z
700.199	1025.885	-1322.802	2723.200		0.683

beta2*(ejecfrac)*	Boots. sdt Err.	[95% Conf.	Interval]	(N)	z
-175.455	52.780	-279.534	-71.375		-3.324

Stabilized

beta1*(abcix)*	Boots. sdt Err.	[95% Conf.	Interval]	(N)	z
861.719	924.710	-961.769	2685.207		0.932

beta2*(ejecfrac)*	Boots. sdt Err.	[95% Conf.	Interval]	(N)	z
-133.303	40.388	-212.946	-53.661		-3.30

MSM model $E(Y_{abcix}|\text{ejecfrac}) = \beta_0 + \beta_1\,\text{abcix} + \beta_2\,\text{ejecfrac}$

How do you interpret the parameters β_1 and β_2 ? What assumptions are needed? Do the parameters have the same interpretation in the two analysis?

Regression models for evaluating effect modification when L = M

- If

$$L=M$$

- M = set of effect modifiers
- L = set of variables needed to control confounding bias
- Then, you do not need to weight because already conditionally on M, the treated and untreated are exchangeable.

75