Agenda

- Concept of Incidence
- Cumulative Incidence
- Incidence Rate
- Relationship between Cumulative Incidence and Incidence Rate
- Relationship between Incidence and Prevalence
- Problems

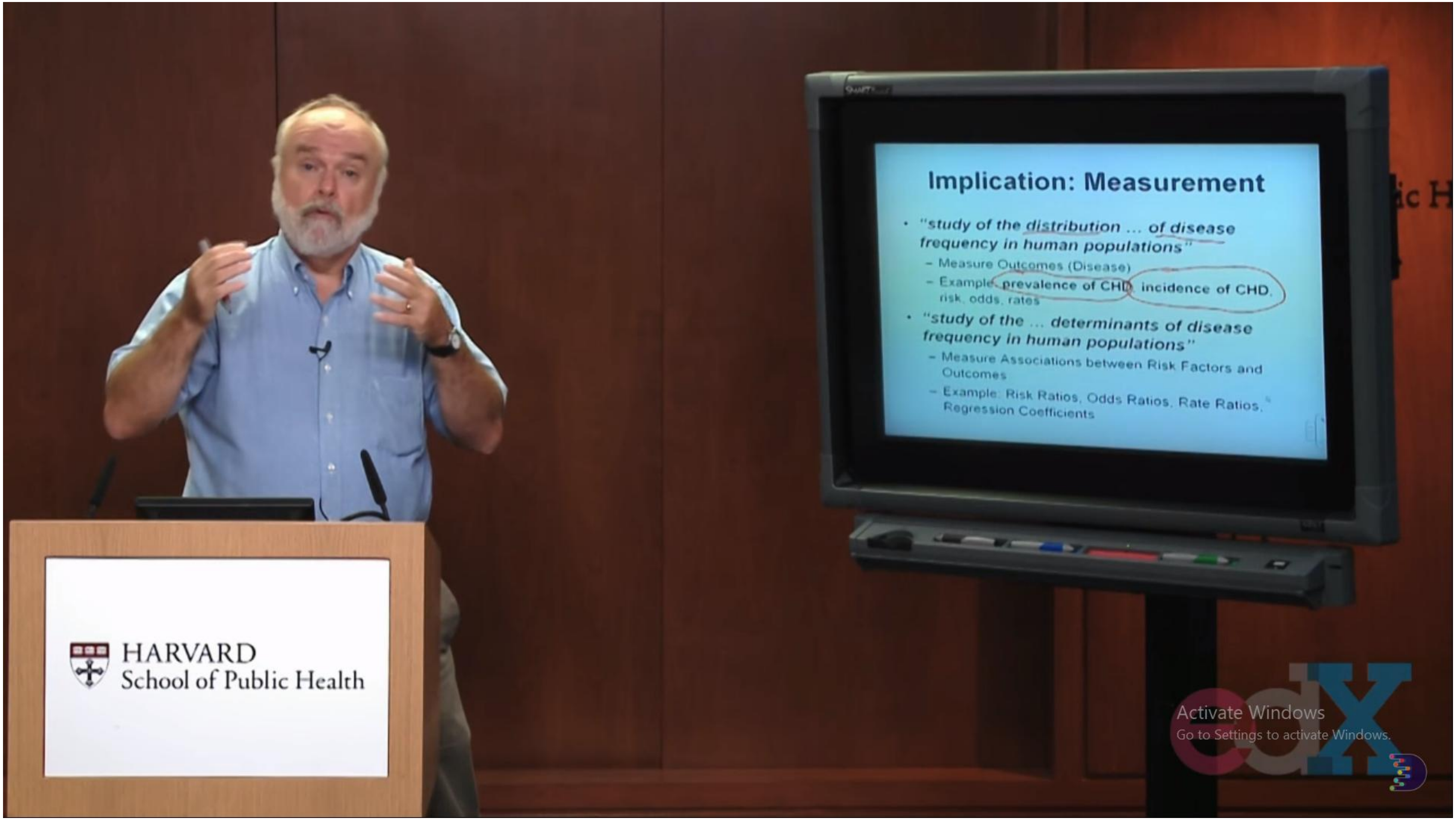
Definition of Epidemiology

" The study of the distribution and determinants of disease frequency in man"

MacMahon B, Pugh TF. Epidemiology: principles and methods. Little Brown. 1970

Implication: Measurement

- ***"study of the distribution … of disease frequency in human populations"***
  - Measure Outcomes (Disease)
  - Example: **prevalence of CHD**, **incidence of CHD**, risk, odds, rates
- ***"study of the … determinants of disease frequency in human populations"***
  - Measure Associations between Risk Factors and Outcomes
  - Example: Risk Ratios, Odds Ratios, Rate Ratios, Regression Coefficients


Implication: Measurement
• "study of the distribution ... of disease frequency in human populations"
– Measure Outcomes (Disease)
– Example: prevalence of CHD, incidence of CHD, risk, odds, rates
• "study of the ... determinants of disease frequency in human populations"
– Measure Associations between Risk Factors and Outcomes
– Example: Risk Ratios, Odds Ratios, Rate Ratios, Regression Coefficients
HARVARD
School of Public Health
Activate Windows
Go to Settings to activate Windows.

Why is Prevalence of CHD higher among Non-Smokers?

- Non-Smokers have higher risk of developing CHD (**higher incidence – unlikely**)
- Non-Smokers live longer after developing CHD (**longer duration – unlikely**)
- CHD causes people to stop smoking (**reversed causation - likely**)
- **Bias**
- **Confounding**
- **Chance**

Conclusion

- Prevalence data are not good for identify risk factors for developing disease
- Alternative: Measure **Incidence** – development of disease in a population followed over time

Incidence

- Measures development of disease in a population over a time
- **Population-at-Risk:** Free of disease at the start of the follow-up period and at risk for developing the disease
- **Cumulative Incidence (Estimated Risk, Average Risk)**
  - (# Cases)/(# Subjects at Risk)
- **Incidence Rate**:
  - (#Cases)/(Person-Time)

Incidence

- Measures development of disease in a population over a time
- **Population-at-Risk:** Free of disease at the start of the follow-up period and at risk for developing the disease
- **Cumulative Incidence (Estimated Risk, Average Risk)**
  - (# Cases)/(# Subjects at Risk)
- **Incidence Rate**:
  - (#Cases)/(Person-Time)

Incidence

- Measures development of disease in a population over a time
- **Population-at-Risk**: Free of disease at the start of the follow-up period and at risk for developing the disease
- **Cumulative Incidence (Estimated Risk, Average Risk)**
  - (# Cases)/(# Subjects at Risk)
- **Incidence Rate**:
  - (#Cases)/(Person-Time)

Example: Incidence of First Headaches (D) among 100 students over 7 days

ID#								Days of Follow-up
1.	D							1
2.	D							1
3.	D							1
4.	D							1
5.	X	D						2
6.	X	X	D					3
7.	X	X	D					3
8.	X	X	D					3
9.	X	X	D					3
10.	X	X	X	D				4
11.	X	X	X	D				4
12.	X	X	X	X	D			5
13.	X	X	X	X	X	D		6
14.	X	X	X	X	X	D		6
15.	X	X	X	X	X	X	D	7
16.	X	X	X	X	X	X	X	7
...								
100.	X	X	X	X	X	X	X	7
Day	1	2	3	4	5	6	7	

Example: Incidence of First Headaches (D) among 100 students over 7 days

ID#		Days of Follow-up
1.	→ D	1
2.	→ D	1
3.	→ D	1
4.	→ D	1
5.	→ D	2
6.	→ D	3
7.	→ D	3
8.	→ D	3
9.	→ D	3
10.	→ D	4
11.	→ D	4
12	→ D	5
13.	→ D	6
14.	→ D	6
15.	→ D	7
16.	→	7
100.	→	7

Day 1 2 3 4 5 6 7

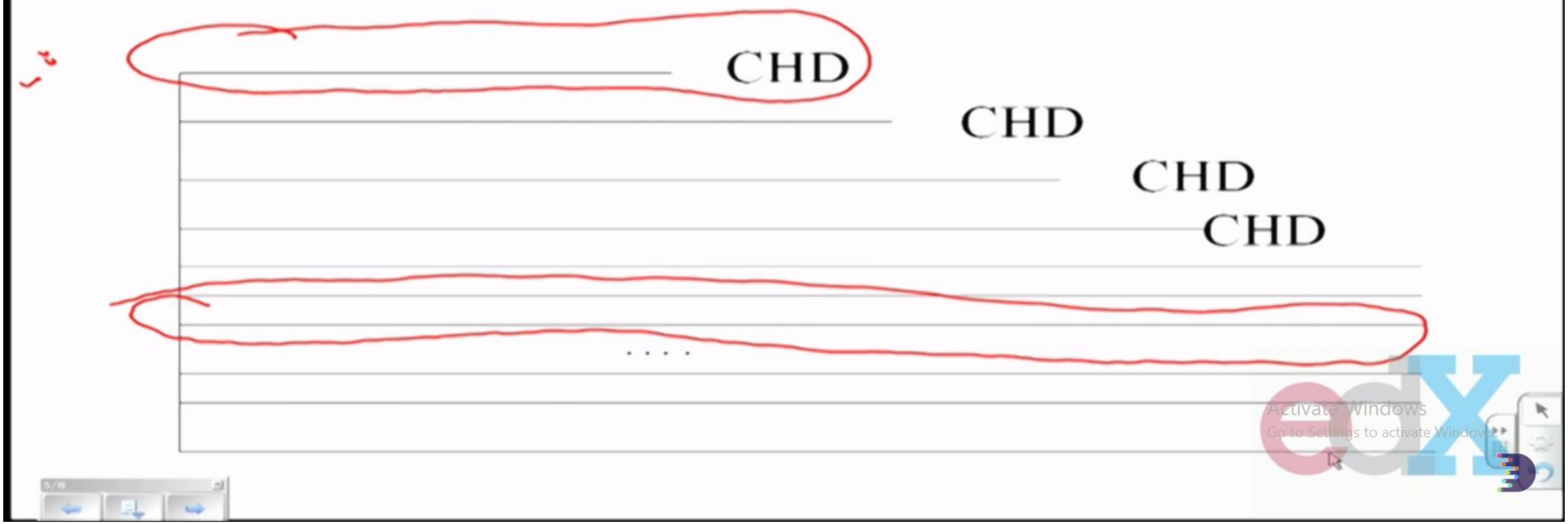
Options for Describing the Incidence of Headaches

- **15 students developed headaches**
  - **Problem**: doesn't account for the population at risk
  - **Problem**: doesn't account for length of follow-up
  - Need a point of reference

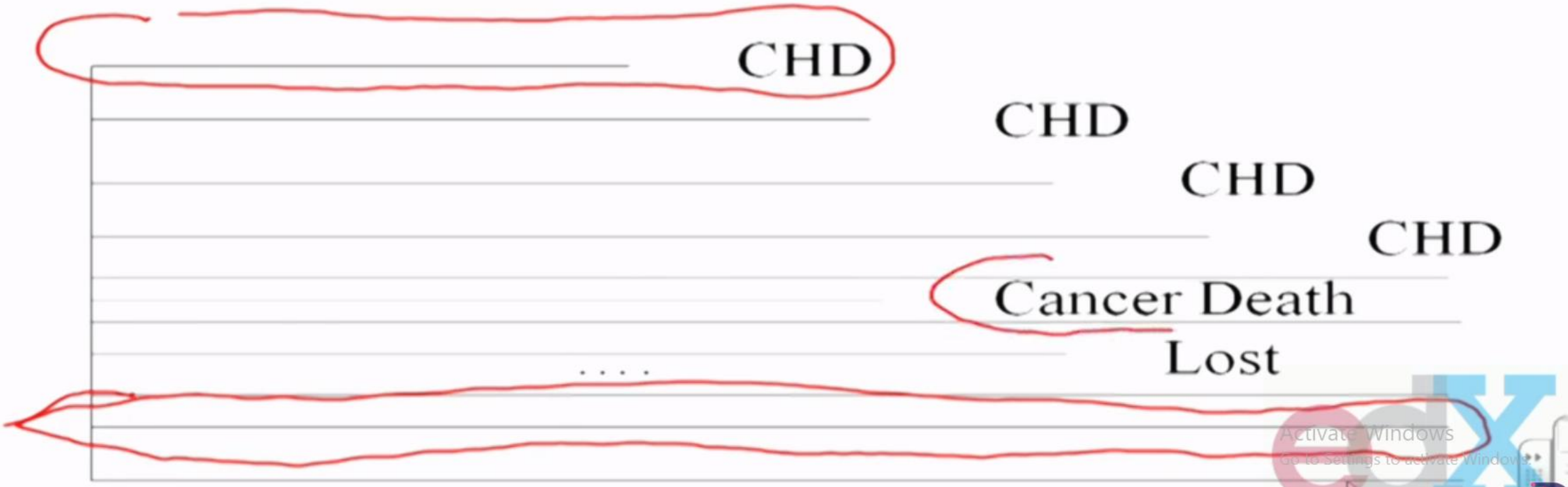
Option #1 for Describing the Incidence of Headaches

- 0.15 developed Headaches within 7 days
  - Proportion
- **Cumulative incidence (Estimated Risk)**
  - **(#new cases of disease)/(# in Pop.-at-Risk)**
- Problem: Value depends on chosen length of follow-up
  - 0.11 developed Headaches within 4 days
  - 0.04 developed Headaches within 1 day

Data Needed for Calculating Cumulative Incidence of CHD

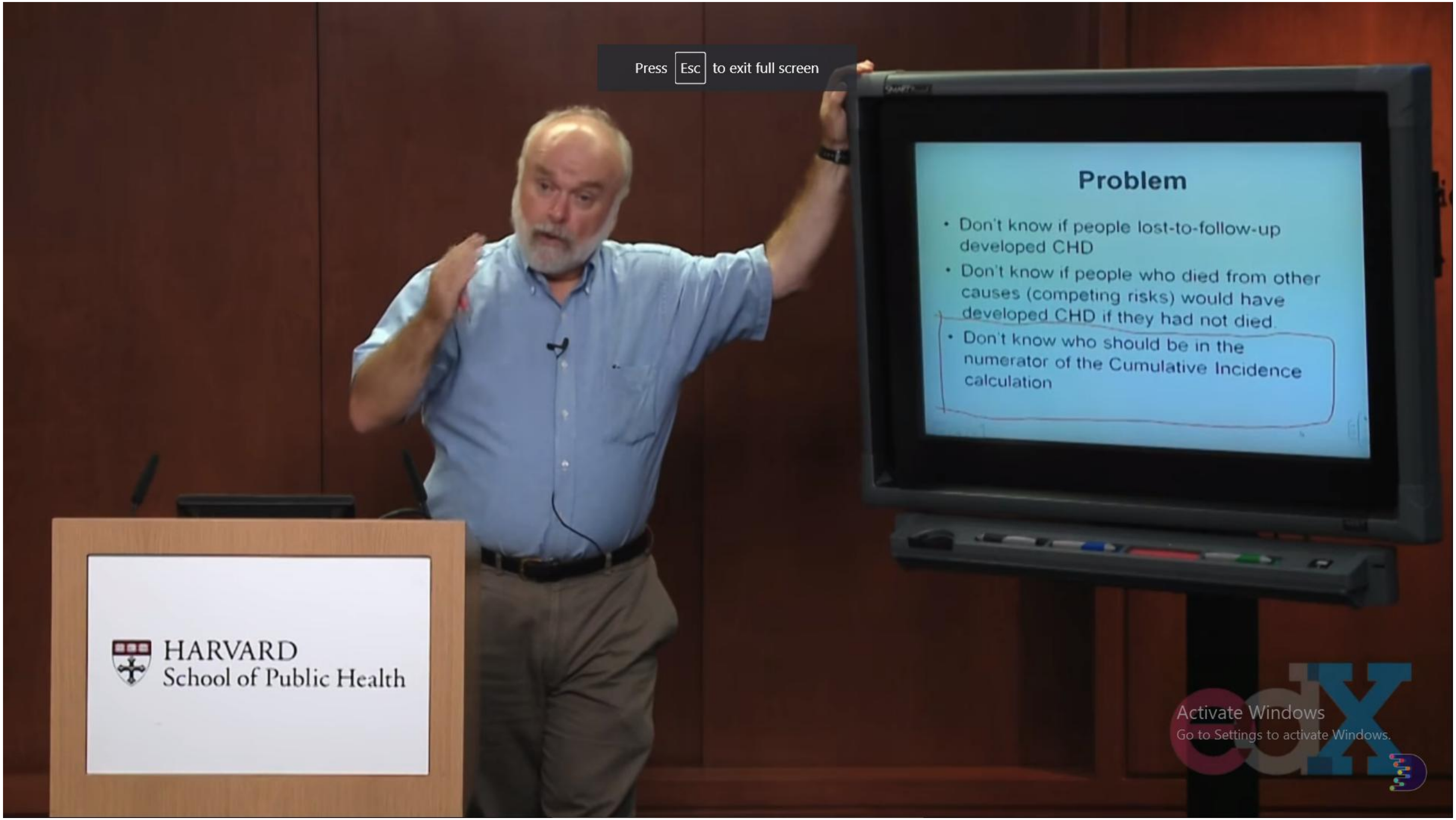


Problem: Competing Risks and Losses-to-Follow-up



Problem

- Don't know if people lost-to-follow-up developed CHD
- Don't know if people who died from other causes (competing risks) would have developed CHD if they had not died.
- Don't know who should be in the numerator of the Cumulative Incidence calculation


Press Esc to exit full screen
Problem
• Don't know if people lost-to-follow-up developed CHD
• Don't know if people who died from other causes (competing risks) would have developed CHD if they had not died.
• Don't know who should be in the numerator of the Cumulative Incidence calculation
HARVARD
School of Public Health
Activate Windows
Go to Settings to activate Windows.

Solution

- If reason for incomplete follow-up is unrelated to risk of CHD, then survival analysis methods provide valid estimates of the risk for developing CHD during a fixed time period
- Special Case: **Incidence Rates**

Option # 2 for Describing the Incidence of Headaches

- 15 students developed Headaches from 645 person-days of observation
- **Incidence Rate** = 15/(645 person-days)
  - 15/645pd = .0233/1pd = 2.33/100pd
- **Problem**: Summary measure and may not reflect high or low incidence rates on certain days

Incidence Rate Calculation

- IR(t) = (# cases that develop during the interval around t)/(amount of person-time observed during the interval)
- Length of the time interval should reflect constancy of IR(t)

Incidence Rate Calculation

Incidence Rate at time "t' = IR(t)

= (# of new cases occurring during interval around 't")
/(amount of person-time observed during the interval)

Question: Why divide by person-time?

Incidence Rate Calculation

Suppose you have 100 people who are "at-risk" during the one-week period around "t" and 5 develop the disease

$$IR(t) \approx$$

Incidence Rate Calculation

Suppose you have 100 people who are "at-risk" during the one-week period around "t" and 5 develop the disease

$$IR(t) \approx \frac{5 \text{ CASES}}{700 \text{ p-days}}$$

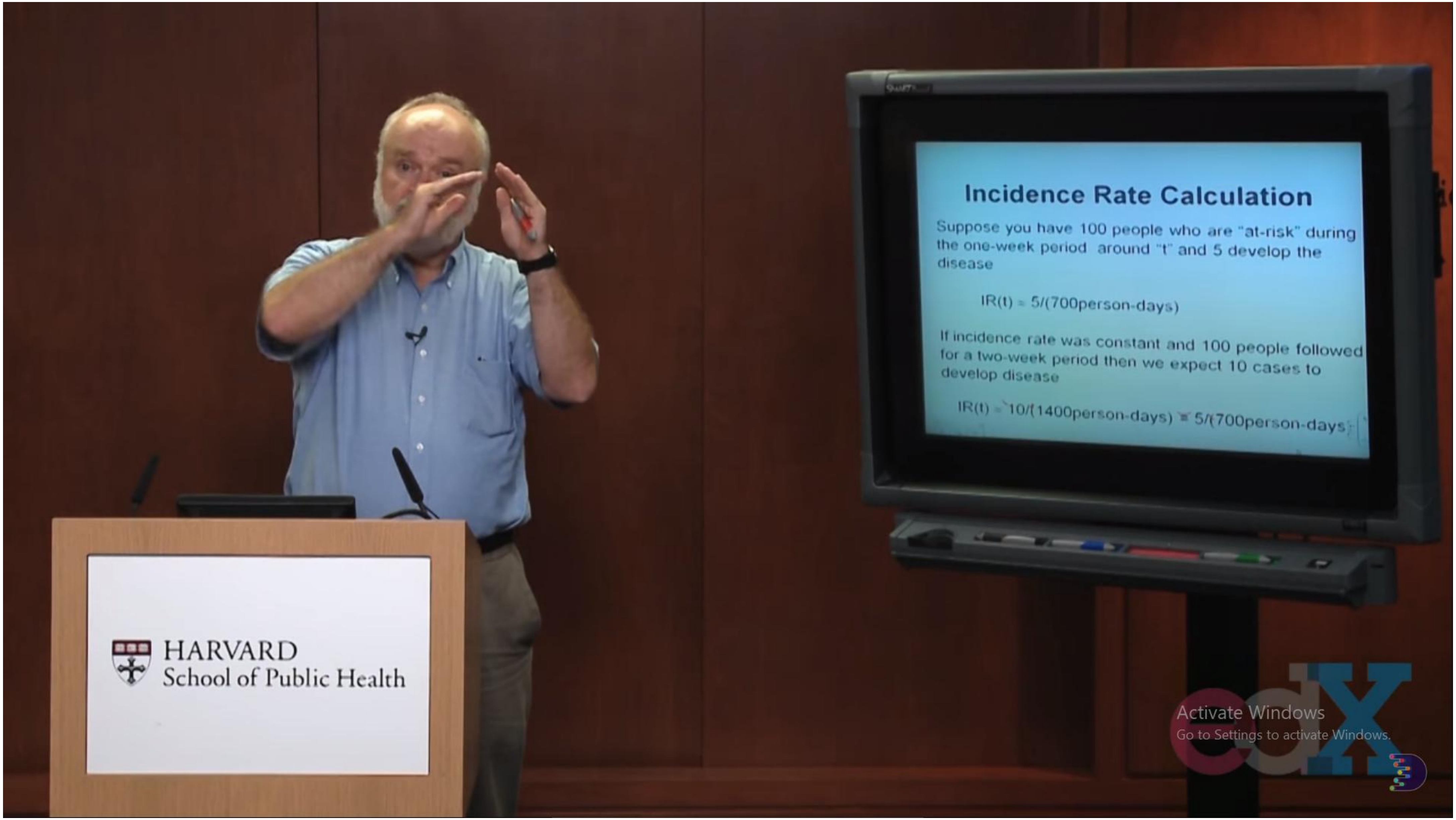
Incidence Rate Calculation

Suppose you have 100 people who are "at-risk" during the one-week period around "t" and 5 develop the disease

$$IR(t) \approx 5/(700\text{person-days})$$

If incidence rate was constant and 100 people followed for a two-week period then we expect 10 cases to develop disease

$$IR(t) \approx \frac{10}{1400P}$$


Incidence Rate Calculation
Suppose you have 100 people who are "at-risk" during the one-week period around "t" and 5 develop the disease
IR(t) = 5/(700person-days)
If incidence rate was constant and 100 people followed for a two-week period then we expect 10 cases to develop disease
IR(t) = 10/(1400person-days) = 5/(700person-days)
HARVARD
School of Public Health
Activate Windows
Go to Settings to activate Windows.

Incidence Rate Calculation

- Special case: constant IR(t) during the entire follow-up period

  IR = (total number of cases)/
  (total amount of follow-up)

- General Case: **IR** = Average IR(t)

Incidence Rate Calculation

- Special case: constant IR(t) during the entire follow-up period

  IR = (total number of cases)/ (total amount of follow-up)

- General Case: **IR** = Average IR(t)

Mortality Experience for Three Populations Followed for Four Years

	Population		
	A	B	C
Number-at-Risk	100	100	100
# Deaths in first year	10	2	10
# Deaths in second year	10	2	9
# Deaths in third year	10	2	8
# Deaths in fourth year	10	34	7
Total # Deaths	40	40	34
4-yr Cum. Incidence			

Mortality Experience for Three Populations Followed for Four Years

	Population		
	A	B	C
Number-at-Risk	100	100	100
# Deaths in first year	10	2	10
# Deaths in second year	10	2	9
# Deaths in third year	10	2	8
# Deaths in fourth year	10	34	7
Total # Deaths	40	40	34
4-yr Cum. Incidence	0.40	0.40	0.34

Incidence Rates: Population A

Year	Population-at-Risk at start of year	# Deaths	Person-Years	IR (deaths/py)
1	100	10	95	.11
2	90	10	85	.12
3	80	10	75	.13
4	70	10	65	.15
Total		40	320	.13

Week 3 : INCIDENCE RATE OF CHD

Incidence Rates: Population A

Year	Population-at-Risk at start of year	# Deaths	Person-Years	IR (deaths/py)
1	100	10	95	10/95py = .11
2	90	10	85	10/85py = .12
3	80	10	75	.13
4	70	10	65	.15
Total		40	320	.13

32:07 / 40:29

Incidence Rates: Population B

Year	Population-at-Risk at start of year	# Deaths	Person-Years	IR (deaths/py)
1	100	2	99	2/99py ≃ .02
2	98	2	97	.02
3	96	2	95	.02
4	94	34	77	34/77py = .44
Total		40	368	.11

Incidence Rates: Population C

Year	Population-at-Risk at start of year	# Deaths	Person-Years	IR (deaths/py)
1	100	10	95	.11
2	90	9	85.5	.11
3	81	8	77	.10
4	73	7	69.5	.10
Total		34	327	.10

Incidence Rates: Population C

Year	Population-at-Risk at start of year	# Deaths	Person-Years	IR (deaths/py)
1	100	10	95	.11
2	90	9	85.5	.11
3	81	8	77	.10
4	73	7	69.5	.10
Total		34	327	34/327 .10

Comments:

For **Population C**, IR(t) appears to be constant over time,

IR = (total no. of new cases)/
(total amount of person-time of follow-up)

total no. of new cases = 10 + 9 + 8 + 7 = 34

total amount of follow-up = 95 + 85.5 + 77 + 69.5
= 327

IR = (34 cases)/(327 person-years)
= (.10 case)/(1 person-year)
= (10 cases)/(100 person years)

If the Incidence Rate function, IR(t), is not constant over time then IR is a **weighted** average of IR(t) for different time periods.

Example: Average Incidence Rate for **Population B**

Average IR = 40 / 368py

= [(99py)**(2/99py)** + (97py)**(2/97py)**
+ (95py)**(2/95py)** + (77py)**(34/77py)**]
/ [99py + 97py + 95py + 77py]

Example

Suppose you followed 1000 people for 1 year and the **Incidence Rate** (IR) was 10/(100y)

How much person-time would you observe?

How many cases of disease would you observe?

What is one-year **Cumulative Incidence**?

Example

Suppose you followed 1000 people for 1 year and the **Incidence Rate** (IR) was 10/(100y)

How much person-time would you observe?

≈ 1000PY

How many cases of disease would you observe?

100

What is one-year **Cumulative Incidence**?

Example

Suppose you followed 1000 people for 1 year and the **Incidence Rate** (IR) was 10/(100y)

- How much person-time would you observe?

  ≈ 1000 PY

- How many cases of disease would you observe?

  100

- What is one-year **Cumulative Incidence**?

  100/1000 = .10

Relationship Between Cumulative Incidence (Risk) and Incidence Rate

- Approximation (for time period Δt) :
  Risk $\approx$ (IR)(Δt))
- Example:
  - IR = 10/100y
  - One-year Risk = (10/100y)(1y) = .10
- Problem: Formula doesn't take into account correct amount of person-time
- Approximation works well for small IR and short Δt

Example (**Population C**): IR = .10/y = 1/10y

Implications from formula:

One-year Risk	=	(1/10y)(1y)	=	$\boxed{0.10}$
Four-year Risk	=	(1/10y)(4y)	=	0.40
Ten-year Risk	=	(1/10y)(10y)	=	1.0
Twelve-year Risk	=	(1/10y)(12y)	=	**1.2 ???**

Problem: Formula does not account for decreasing population-at-risk

Solution: Survival Analysis Methods

Example (**Population C**): IR = .10/y = 1/10y

Implications from formula:

One-year Risk	= (1/10y)(1y)	=	0.10
Four-year Risk	= (1/10y)(4y)	=	0.40
Ten-year Risk	= (1/10y)(10y)	=	1.0
Twelve-year Risk	= (1/10y)(12y)	=	**1.2 ???**

Problem: Formula does not account for decreasing population-at-risk

Solution: Survival Analysis Methods

Incidence Rates: Population C

Year	Population-at-Risk at start of year	# Deaths	Person-Years	IR (deaths/py)
1	100	10	95	.11
2	90	9	85.5	.11
3	81	8	77	.10
4	73	7	69.5	.10
Total		34	327	.10

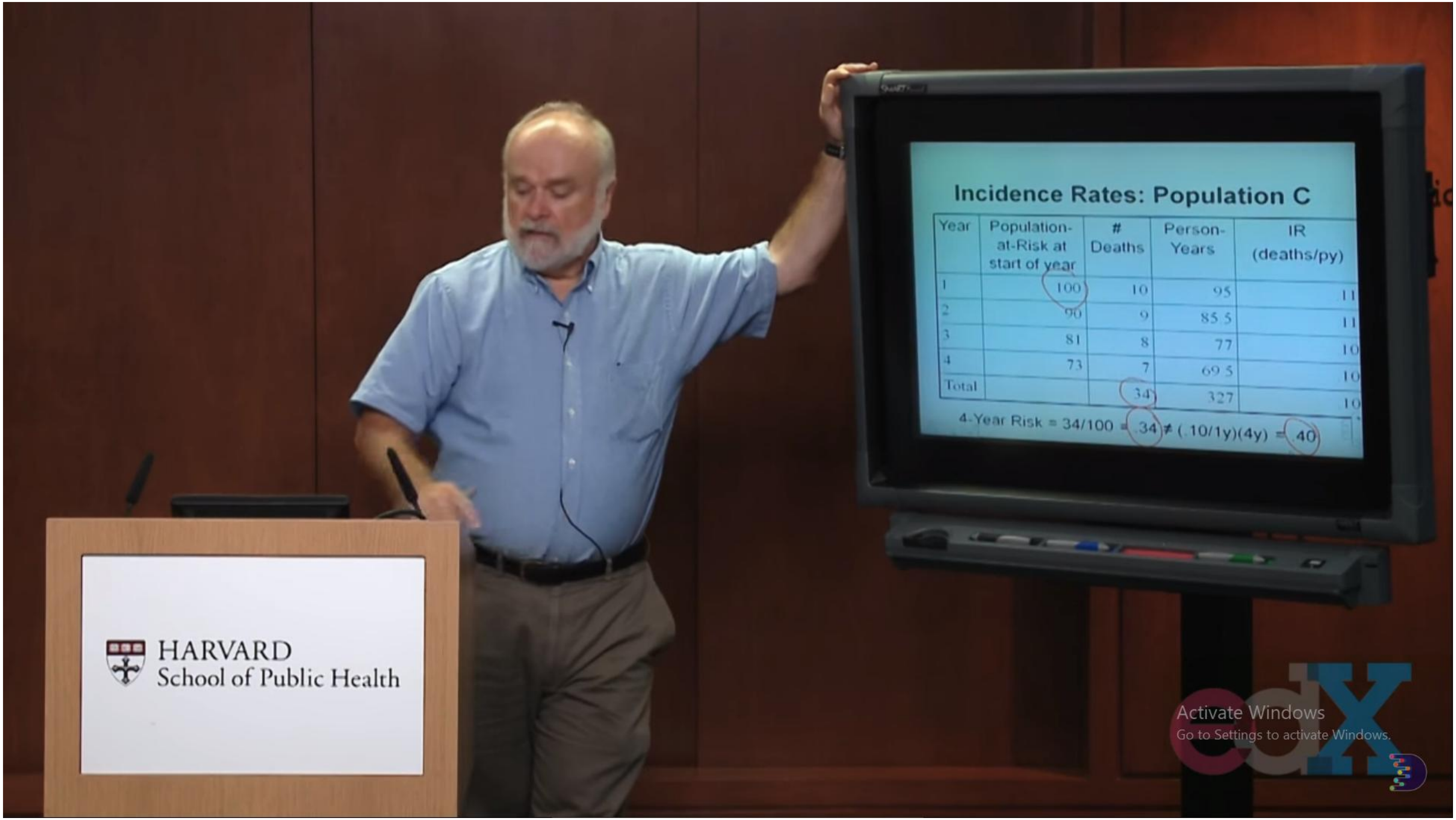
4-Year Risk = 34/100 = .34 ≠ (.10/1y)(4y) = .40

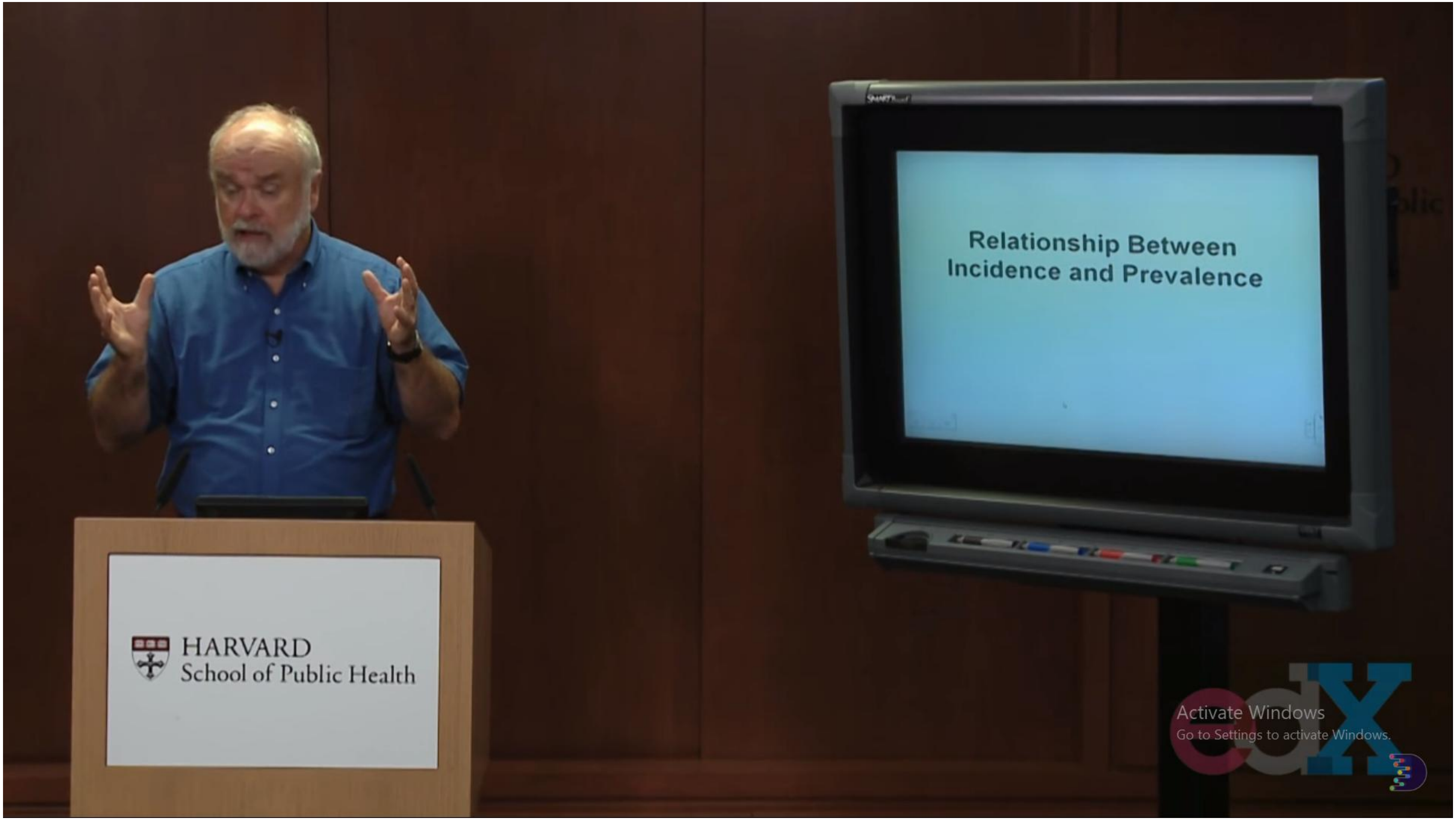
Relationship Between Cumulative Incidence (Risk) and Incidence Rate

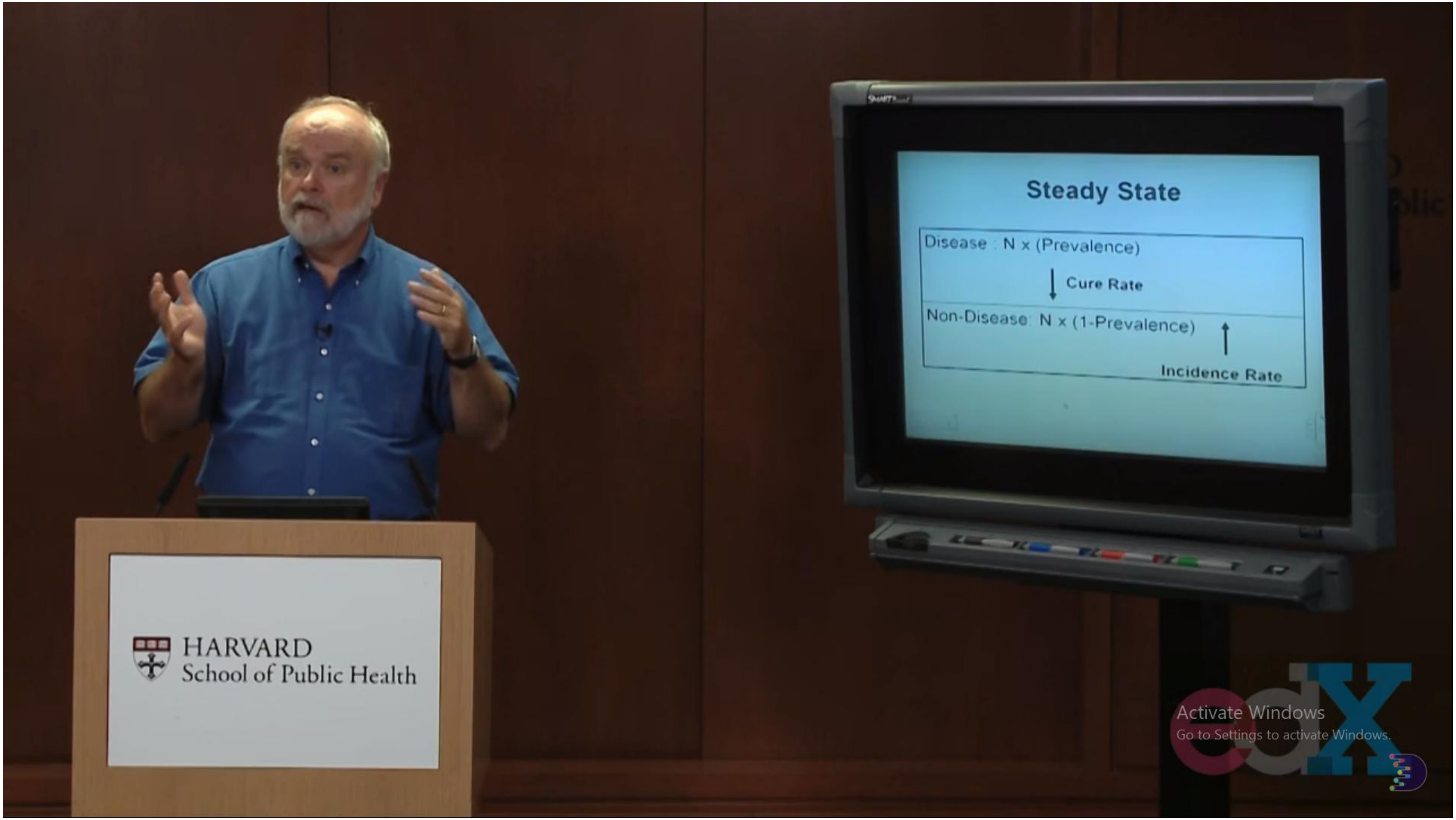
- General Formula: $R = 1 - \exp(-\int IR(x)dx)$
- Special Case: $R = 1 - \exp(-(IR)(\Delta t))$
  - Assumes Constant Incidence Rate
- Example: **Population C**
  - $IR = .10/1y$
  - Four-year Risk $= 1 - \exp(-(.10/1y)(4y)) = 0.33$

Relationship Between Cumulative Incidence (Risk) and Incidence Rate

- General Formula: **$R = 1 - \exp(-\int IR(x)dx)$**
- Special Case: **$R = 1 - \exp(-(IR)(\Delta t))$**
  - Assumes Constant Incidence Rate
- Example: **Population C**
  - $IR = .10/1y$
  - Four-year Risk $= 1 - \exp(-(.10/1y)(4y)) = 0.33$


Incidence Rates: Population C
Year
Population-at-Risk at start of year
Deaths
Person-Years
IR (deaths/py)
1 100 10 95
2 90 9 85.5
3 81 8 77
4 73 7 69.5
Total 34 327
4-Year Risk = 34/100 = .34 ≠ (.10/1y)(4y) = .40
HARVARD
School of Public Health
Activate Windows
Go to Settings to activate Windows.


Relationship Between
Incidence and Prevalence
HARVARD
School of Public Health
Activate Windows
Go to Settings to activate Windows.



HARVARD
School of Public Health
Steady State
Disease : N x (Prevalence)
Cure Rate
Non-Disease: N x (1-Prevalence)
Incidence Rate
Activate Windows
Go to Settings to activate Windows.

Steady State

Disease : N x (Prevalence)

↓ **Cure Rate**

Non-Disease: N x (1-Prevalence) ↑

Incidence Rate

Steady State

Disease : N x (Prevalence)

↓ **Cure Rate**

Non-Disease: N x (1-Prevalence) ↑

Incidence Rate

$$[N \times \text{Prevalence}] \times \text{Cure Rate} = [N \times (1-\text{Prevalence})] \times \text{Incidence Rate}$$

Steady State

[N x Prevalence] x Cure Rate =
[N x (1-Prevalence)] x Incidence Rate

Steady State

$$[N \times \text{Prevalence}] \times \text{Cure Rate} = [N \times (1-\text{Prevalence})] \times \text{Incidence Rate}$$

$$\text{Prevalence}/(1-\text{Prevalence}) = (\text{Incidence Rate})/(\text{Cure Rate})$$

Steady State

Prevalence/(1-Prevalence)
= (Incidence Rate)/(Cure Rate)

Cure Rate = (#cures)/(# disease-days)
= 1/(Average Duration of Disease)
= 1/D

Prevalence/(1-Prevalence) = (Incidence Rate)/(1/D)
= (Incidence Rate) x D

Steady State

Prevalence/(1-Prevalence)
= (Incidence Rate)/(Cure Rate)

Cure Rate = (#cures)/(# disease-days)
= 1/(Average Duration of Disease)
= 1/D

Prevalence/(1-Prevalence) = (Incidence Rate)/(1/D)
= (Incidence Rate) x D

Steady State

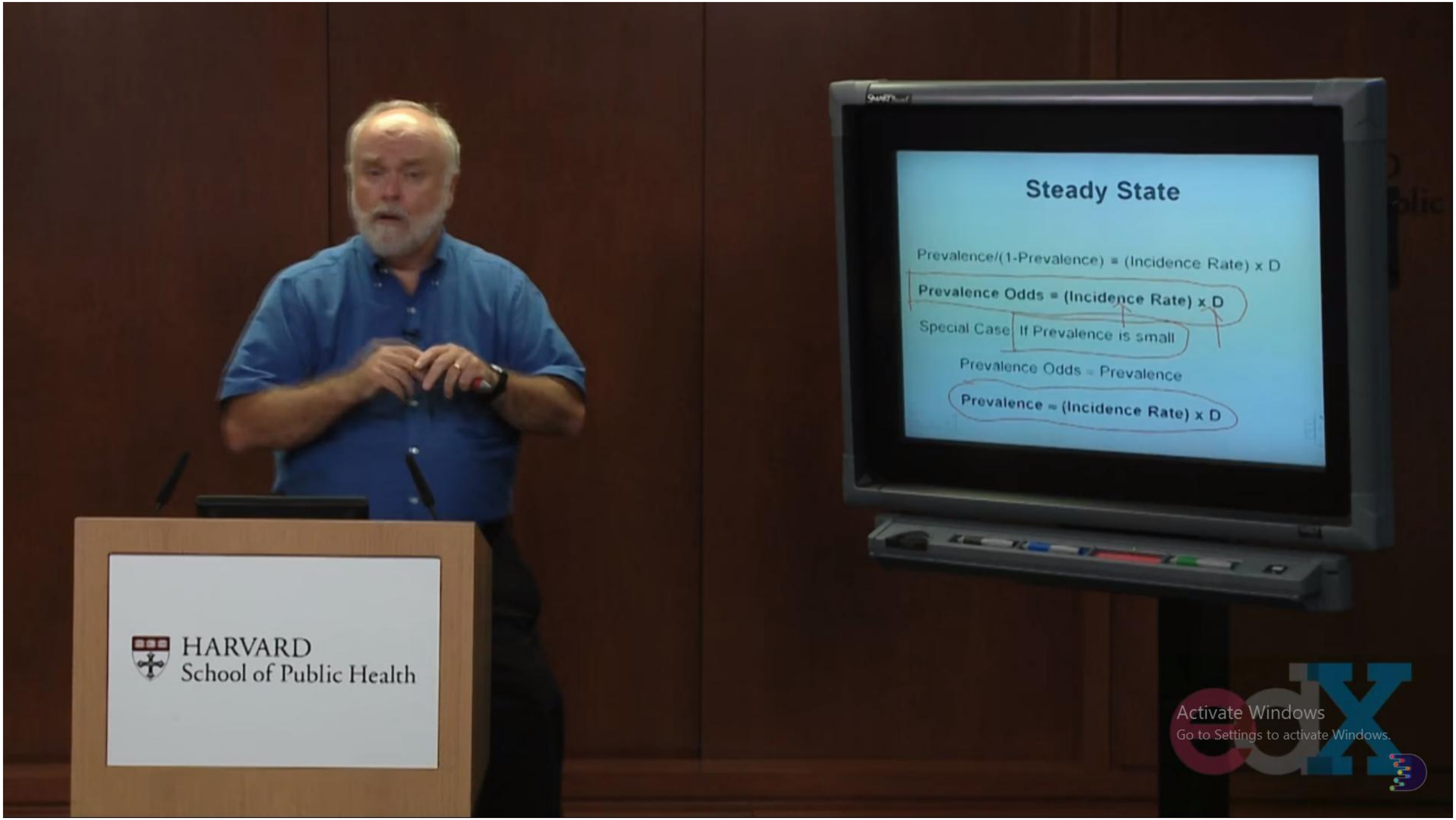
Prevalence/(1-Prevalence) = (Incidence Rate) x D

Prevalence Odds = (Incidence Rate) x D

Special Case: If Prevalence is small

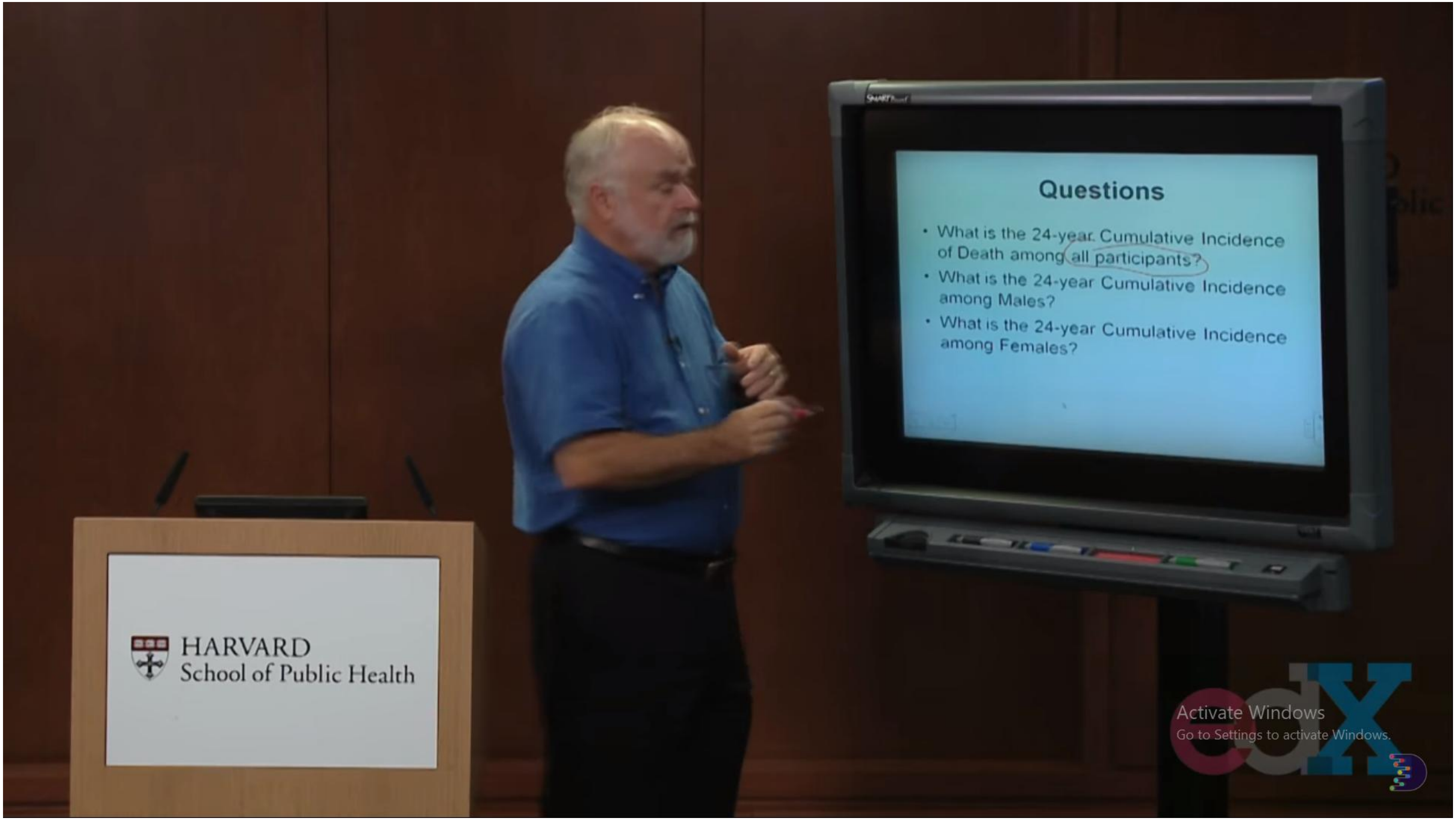
Prevalence Odds ≈ Prevalence

Prevalence ≈ (Incidence Rate) x D


Steady State
Prevalence/(1-Prevalence) = (Incidence Rate) x D
Prevalence Odds = (Incidence Rate) x D
Special Case
If Prevalence is small
Prevalence Odds ≈ Prevalence
Prevalence ≈ (Incidence Rate) x D
HARVARD
School of Public Health
Activate Windows
Go to Settings to activate Windows.

24-Year Cumulative Incidence of Death for Males and Females

Sex	Died	Survived	Total	Cumulative Incidence
Males	843	1101	1944	
Females	707	1783	2490	
Total	1550	2884	4434	


Questions
What is the 24-year Cumulative Incidence of Death among all participants?
What is the 24-year Cumulative Incidence among Males?
What is the 24-year Cumulative Incidence among Females?
HARVARD
School of Public Health
Activate Windows
Go to Settings to activate Windows.

Questions

- What is the 24-year Cumulative Incidence of Death among all participants?
- What is the 24-year Cumulative Incidence among Males?
- What is the 24-year Cumulative Incidence among Females?

Week 3 : SOLUTION: PROBLEM 1

24-Year Cumulative Incidence of Death for Males and Females

Sex	Died	Survived	Total	Cumulative Incidence
Males	843	1101	1944	
Females	707	1783	2490	
Total	1550	2884	4434	1550/4

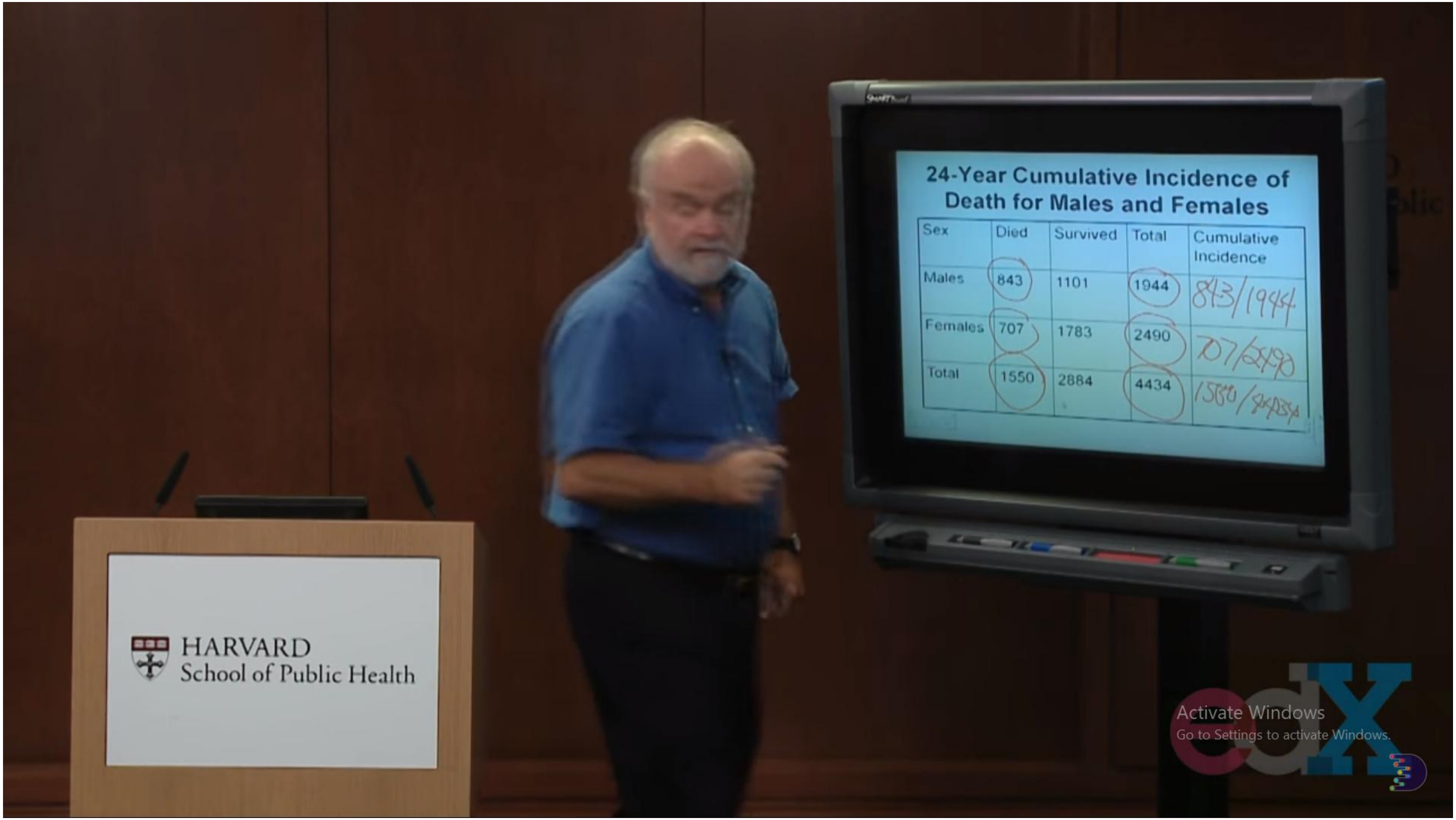
0:47 / 2:01

24-Year Cumulative Incidence of Death for Males and Females

Sex	Died	Survived	Total	Cumulative Incidence
Males	843	1101	1944	8
Females	707	1783	2490	
Total	1550	2884	4434	1550/4434

24-Year Cumulative Incidence of Death for Males and Females

Sex	Died	Survived	Total	Cumulative Incidence
Males	843	1101	1944	843/194
Females	707	1783	2490	
Total	1550	2884	4434	1550/4434


24-Year Cumulative Incidence of Death for Males and Females
Sex
Died
Survived
Total
Cumulative Incidence
Males
843
1101
1944
843/1944
Females
707
1783
2490
707/2490
Total
1550
2884
4434
HARVARD
School of Public Health
Activate Windows
Go to Settings to activate Windows.

24-Year Cumulative Incidence of Death for Males and Females

Sex	Died	Survived	Total	Cumulative Incidence
Males	843	1101	1944	843/1944 = 0.43
Females	707	1783	2490	707/2490 = 0.28
Total	1550	2884	4434	1550/4343 = 0.35

Incidence Rate of Death for Males and Females

Sex	Died	Person-Years of Follow-up	Incidence Rate
Males	843	38287.33	
Females	707	52828.26	
Total	1550	91115.59	

Questions

- What is the Incidence Rate of Death among all participants?
- What is the Incidence Rate of Death among Males?
- What is the Incidence Rate of Death among Females?

Incidence Rate of Death for Males and Females

Sex	Died	Person-Years of Follow-up	Incidence Rate
Males	843	38287.33	
Females	707	52828.26	
Total	1550	91115.59	1550/(91115

Incidence Rate of Death for Males and Females

Sex	Died	Person-Years of Follow-up	Incidence Rate
Males	843	38287.33	843/38287.33py = 2.20/100py
Females	707	52828.26	707/52828.26py = 1.34/100py
Total	1550	91115.59	1550/91115.59py = 1.70/100py

Incidence of CHD for Males and Females

Sex	CHD	Population-at-Risk	Person-Years of Follow-up
Males	710	1944	32763.95
Females	530	2490	48161.20
Total	1240	4434	80925.15

Question

- How would you measure the incidence of CHD among men?
- How would you measure the incidence of CHD among women?

Problem

- Cannot measure Cumulative Incidence directly
- Not all participants are "at-risk" for developing (and recording) CHD during 24 years of follow-up
- Reasons
  – Deaths from other causes
  – Losses-to-follow-up
- Solution: Report Incidence Rate

Questions

- What is the Incidence Rate of CHD among all participants?
- What is the Incidence Rate of CHD among Males?
- What is the Incidence Rate of CHD among Females?

Incidence Rate of CHD for Males and Females

Sex	CHD	Person-Years of Follow-up	Incidence Rate
Males	710	32763.95	
Females	530	48161.20	
Total	1240	80925.15	1240/(80,925.15

Incidence Rate of CHD for Males and Females

Sex	CHD	Person-Years of Follow-up	Incidence Rate
Males	710	32763.95	710/32763.95 = 2.17/100py
Females	530	48161.20	530/48161.20 = 1.10/100py
Total	1240	80925.15	1240/80925.15 = 1.53/100py