* Kresge 502 Cart: Alright, so people are already starting to respond. So as we talked about on Tuesday, we always have a trivia question at the start of class. So this particular question talks about which of the following is not true. So one of these 4 answers is not true about

0:15

* Kresge 502 Cart: breast cancer. You can either join by the web you can join by text, or there's also a QR. Code. I don't know how close you have to be to scan it or not. But you have 3 different ways to answer. So the question is, which of the following is not true. Is it a. That the first mastectomy for breast cancer treatment was 1,500 years ago. B. That obesity be before puberty is associated with an increased risk of breast cancer.

0:39

* Kresge 502 Cart: Breast cancer has a higher occurrence in the left than the right breast, or that breast cancer can occur in cats and dogs. And it looks like we have 15 people who have responded for people just coming in. We'll give you about maybe 25 s to think about it.

1:09

* Kresge 502 Cart: You can talk with your colleagues

1:30

* Kresge 502 Cart: this is, can be a collaboration.

1:34

* Kresge 502 Cart: 3

1:46

* Kresge 502 Cart: of most people who wanted to give it maybe 10 more seconds.

1:51

* Kresge 502 Cart: 5 more seconds. If anybody wants more time. Just raise your hand. We're going to call it.

2:13

* Kresge 502 Cart: and and you just hit escape

2:22

* Kresge 502 Cart: to see the answers alright. And the winner is

2:24

* Kresge 502 Cart: response, history over here. Yeah, got it? Oh, interesting. Okay. So so many people thought that this was not correct, that breast cancer has a higher, higher occurrence in the left than the right breast. For some reason that's actually true. So about 51% of tumors arise in the left breast versus the right breast. According to my reading of the literature, there's no real explanation of why that is. But that

2:30

* Kresge 502 Cart: is not true right? So that is an actual, true statement. So the actual, true answer is B.

3:02

* Kresge 502 Cart: So what you'll learn about breast cancer when Dr. Eliason gives the lecture is that excess body weight early in life?

3:11

* Kresge 502 Cart: Well, that is.

3:24

* Kresge 502 Cart: that's true. That is true.

3:27

* Kresge 502 Cart: This is a typo. Oh, darn! Oh, I'm so sorry, so sorry. Oh, goodness, Ok, Major, fail.

3:29

* Kresge 502 Cart: Sorry about that, guys. So, okay, so these are all actually, true answers. So, okay, that's that's too bad. But you did. But really interesting thought it was like a extra trick. Luckily it's it's doesn't count towards the grade. Okay, so sorry about that. I think. I remember that now from last year that it was a typo, and I think I carried it forward. So I apologize.

3:44

* Kresge 502 Cart: So first we wanted to talk a little bit about upcoming assignments. So your first assignment is going to be due next Thursday. So a week from today is that correct? Is that next Thursday, Tuesday. So we actually a week and a half. And it will be due by the end of the day, and it's

4:09

* Kresge 502 Cart: most helpful for all of us if you upload. Your assignment onto the Harvard canvas page. And so again, you're gonna be thinking in the mindset as if you're a reviewer for a journal. So you wanna think about reviewing the epidemiologic aspects of the study. So as I mentioned earlier, it's it's gonna be a study using the all of Us research program

4:34

* Kresge 502 Cart: to look at the association between air quality and cancer risk. And there's a detail word document on Harvard canvas that goes through. How do you approach

4:59

* Kresge 502 Cart: reviewing a a paper for a journal article? Usually you wanna look through and think about the study design? Are there potential sources of bias. Is the hypothesis sound? Have they taken, undertaken the study with with rigor? And so that the results? And and have they interpreted the results

5:09

* Kresge 502 Cart: like correctly, or have they? Sometimes you'll see authors will overstate results. So there's a really detailed set of criteria in how to write up the critique, and it shouldn't be really long to be no more than 3 pages double space. And so our wonderful teaching fellows are gonna have an office hour on February first

5:31

* Kresge 502 Cart: from 3, 30 to 4, 30 in Cresc. 205. I wonder do you think it's possible to record it or have a zoom option? Would that be possible? And then for people who might not be able to attend, but might wanna hear with the discussion, so we'll post that onto the Harvard canvas page. So if you, if any questions come up, take a look at the Harvard canvas page about the structure and and way to think about this

5:54

* Kresge 502 Cart: really want to kind of go through with the mind of an epidemiologist in thinking through the critique. And then. kind of later in February. The the one of the bigger assignments is the cancer epidemiology pro projects. And remember, there's 2 parts to this. There's gonna be an individual. Write up on the descriptive epidemiology of that cancer. And then, you, you're gonna have a group presentation where you present, not only on the descriptive epidemiology of cancer, but also on the association between a risk factor and cancer.

6:18

* Kresge 502 Cart: So starting at 50'clock today, so start thinking about which of these cancer sites you'd like to take part in. And do people sign up for more than one?

6:53

* Kresge 502 Cart: Okay? And what happens if we have too many people, there's a cap. There is a cap. Okay? So if you're really excited about a certain cancer definitely, try to log on as soon as you can at 50'clock. So you can get in on that cancer. Honestly, each of the cancers that is listed here is really fascinating from an epidemiology perspective, both in terms of the descriptive epidemi epidemiology which really is like the patterns of disease

7:05

* Kresge 502 Cart: in different populations across time, across different groups as well as risk factors. So I think there's a lot of interest in

7:32

* Kresge 502 Cart: in these. so we'll talk a little bit more in detail about what each of the components looks like. But just to let you know that 50'clock today go to Harvard canvas and sign up self. Select for one of the cancers any questions on either of the assignments. Yes.

7:40

* Kresge 502 Cart: Per group. Yeah, is it 5 or

7:59

* Kresge 502 Cart: exactly. Yeah. Great. Any other questions.

8:13

* Kresge 502 Cart: Okay, fantastic.

8:32

* Kresge 502 Cart: Okay, so we're gonna delve into a little bit more into some of the the core methods and cancer epidemology and on Tuesday we talked a little bit about the role that family history plays in cancer. Which is interesting both in terms of thinking of it as

8:34

* Kresge 502 Cart: a risk factor and try to understand the ideology of cancer. But also, things like family history can be used. You know, for example, clinically and risk prediction models. And looking at you know, which groups of people might need to have more regular surveillance, more regular screening, for example. So how do we? How, when you think about family history.

8:51

* Kresge 502 Cart: of in doing an epidemiology study. How might you collect information on family history?

9:15

* Kresge 502 Cart: What are some thoughts on where you might get that kind of information if you were doing an epidemiology, study

9:21

* Kresge 502 Cart: questionnaire. Yeah? Yeah. So you could send out a questionnaire to. Let's say you have a case control study or cohort and just ask people, and how, what kind of questions would you want to ask? Do you think about if you were putting it on the questionnaire?

9:29

* Kresge 502 Cart: Sorry I want to put you on the spot. But yeah, any thoughts about how you might ask some questions. 498, sex. Maybe specifically about family history. How might you ask the question?

9:46

* Kresge 502 Cart: Answer.

10:08

* Kresge 502 Cart: yeah, no. That's great. Yeah, really. Yes, in terms of shared exposures within the family kind of, yeah, yeah, absolutely. So, if you wanted to go into more detail about. You know what might be non genetic factors in family history. You might want to collect that.

10:26

* Kresge 502 Cart: But right? So in the in the questionnaire. Right? You might. If you just wanna collect family history, you might wanna ask not only did has anybody in your family had breast cancer, prostate cancer, pancreatic cancer, etc. You wanna might wanna know what age they were, what relationship they were to you, and maybe probably with cancer. And if you think about genetic relatedness.

10:51

* Kresge 502 Cart: you probably aren't as interested in, maybe say, cousins, etc. So you really might be thinking about primary like first degree relatives, whether it's your parents, whether it's your siblings, maybe even your kids. And then you might also wanna know if they died. And then, you know, with with the advent of of genetic testing, or even things like 23 me, it might be interesting to know, did you have any of that

11:14

* Kresge 502 Cart: testing done, and could we get those results back? So if you were, let's say you had a cohort of

11:39

* Kresge 502 Cart: you know, 10,000 people, ages 30 to 45 at the start of the study. What might be one limitation if you just asked at at the baseline about family history, can you think about about. Maybe I'm not asking the question exactly right. But so you have a group of people sort of early

11:46

* Kresge 502 Cart: in midlife.

12:10

* Kresge 502 Cart: and you're asking them about their family history. Yeah, I feel like some of their family might not be old enough to have undisturbed cancer. Exactly

12:12

* Kresge 502 Cart: right? So you might not capture what really is family history? Just because maybe your siblings haven't gotten old enough. Even your parents might at that point might not be old enough to really be at the highest incidence of cancer. Exactly. Any other limitations you can think about of questionnaire-based information on family history

12:23

* Kresge 502 Cart: if you wanted to.

12:43

* Kresge 502 Cart: Yes, maybe to have a call one.

12:46

* Kresge 502 Cart: Yeah. But observing, like you're not, there could be mistakes

12:49

* Kresge 502 Cart: right? And so like. So there's 2 kind of interesting things with that. So maybe with a cohort study. Maybe just you can't remember right, you know. I think I think my dad had something. But don't remember, I don't. I don't remember. If they did, they really have prostate, or it was pancreatic cancer, right? So in a cohort study would be non differential right? So they would be as likely to over and underreport it. And so, yeah.

12:53

* Kresge 502 Cart: yeah, exactly right. So underestimate the genetic risk, because people are dying from other things it would be interesting to think about is also, are there shared genes between those 2 causes of or the other way which would be interesting. And then, in terms of the recall bias.

13:17

* Kresge 502 Cart: go back to your point. If it's a case control study. And so you're asking people who have cancer who don't have cancer. The people who have cancer may remember things differently. I often talk about it as rumination. Really. Why did I get this cancer? And

13:42

* Kresge 502 Cart: you may think about it differently. Sorry about that right? Exactly. And so was there something somebody else said.

14:01

* Kresge 502 Cart: So there's that. And then the other thing is, there may be people who don't know their family's history right? They may be they were adopted. Maybe there was a family member like a parent who was never part of that person's life. And so in certain populations, understanding family history is much, more, much more challenging. Well, there was not necessarily to get around this, but really interesting opportunity

14:09

* Kresge 502 Cart: for investigating family history is in Scandinavia, where they're able to track

14:34

* Kresge 502 Cart: cancer, hospitalization, death in the entire population. These things are required by law to be registered into these national databases, and they were able to create something called the multigenerational Register. This one particular. This data is from Sweden. So basically, they were able to go from the time the cancer registry first started in Sweden, which was in the 1950 s.

14:42

* Kresge 502 Cart: And then track multiple generations so they could link the birth registers. So they knew who was related to whom in the register, and then they were able to tag on the how those people, whether the people had cancer or not. And then look at multigenerational family history. So I just wanna I think this is an interesting study design and interesting example. So this particular study included 2.2 million

15:07

* Kresge 502 Cart: of women who are leading living in Sweden, in whom 54 0 cases of breast cancer occurred, and then these are the relative risk

15:32

* Kresge 502 Cart: breast cancer in these women. If they had a sister only who had breast cancer, a mother, only 2 sisters, a mother and a sister and or a mother, and a brother, and all of those the reference group is no family history of cancer.

15:41

* Kresge 502 Cart: So what? When you're looking at the relative risk estimates and 95% confidence intervals for risk of breast cancer with family history. What is this data

15:58

* Kresge 502 Cart: kind of say to you?

16:12

* Kresge 502 Cart: Right? So it looks like the the, you know. There's a stronger association. If you have 2 sisters, a mother, sister, and even a stronger with mother and brother. So why do you think that might be why do you think it might be that it's a little bit stronger with having more family members.

16:27

* Kresge 502 Cart: What might that be saying about family history as an exposure?

16:48

* Kresge 502 Cart: Yeah, it could be more may indicate when you have more family members that have the cancer, it might be more indicative that family history isn't just that shared environment that we talked about, but also, maybe more due to genetic factors. Exactly.

16:52

* Kresge 502 Cart: So so one of the challenges with you know, family based studies looking at family history is exactly what we were just talking about, which is, is a family history of genetic factors, or is it not genetic factors? And one interesting epidemiologic approach? That people have taken is, instead of studying family members to really, specifically look at twins.

17:10

* Kresge 502 Cart: What do you think is interesting? We I know we had. We talked about the twin astronauts. Being simple. So what do you think might be interesting about studying twins and the different types of twin pairs that might be able to kind of disentangle the effect of genes from environment in cancer. If you look at family history

17:35

* Kresge 502 Cart: when you think about the types of twins there are. Bless you.

17:59

* Kresge 502 Cart: genetics. Yeah, exactly. So. So if you you think there's 2 main types of twins, others, these on identical or monozygotic twins who share in theory 100% of their inherited genome. And then you have 50% on average share genetic factors in the fraternal twins or the dizygotic twins

18:07

* Kresge 502 Cart: twins. And so you can basically like, you can look at what's your risk of breast cancer, for example, if your twin also had breast cancer, and then you can ask the question, how does that differ? If you're an identical twin versus if you're a fraternal twin, and then you can almost take away the environmental component by subtracting out the effect that you see in the dizzy twins and any excess

18:32

* Kresge 502 Cart: concordance that you see in the monozygotic twins.

18:57

* Kresge 502 Cart: then, you say is is likely due to inherited genetic factors. So this this was I was part of a project. I did my postdoc in in Sweden, and I was part of a really cool study of twins in the Nordic country, where they were able. If anybody's looking for interesting research projects, I think doing work in Sweden's came in. It was really fascinating.

19:01

* Kresge 502 Cart: They were able to in the 1950 s. Basically, they wanted to build this twin cohort so they were able to go back and get the records. These were handwritten records from each of the individual parishes throughout Scandinavia and identify when a twin birth

19:25

* Kresge 502 Cart: was recorded, and at the time they didn't have genetics, so they would go then to the family and say, Do the did the twins look identical or or not, and to try to get at this fraternal versus identical twins. And then they built this registry of of over 200,000 twin pairs from the Scandinavian countries, and then again, because they can link it

19:46

* Kresge 502 Cart: with cancer registries and mortality registries. They can follow up these cohorts for decades and see who develops cancer. And so this is a plot of the age looking at age of diagnosis on the X axis, and then cumulative risk of cancer on the Y axis. And then the gray line represents what's just the cumulative incidence of cancer in that population. And you can see it's about 33%.

20:11

* Kresge 502 Cart: So it's kind of like the lifetime. Risk one in 3.

20:39

* Kresge 502 Cart: And then, if your twin also had some sort of cancer and your dysa got a twin, the likelihood that you also would

20:43

* Kresge 502 Cart: develop cancer during your lifetime is closer to 40%. And then, if you're an identical twin and your twin pair developed some form of cancer. Your own risk, was closer to about 46%. So sort of again, you can sort of see, there seems to be both a genetic and shared environment component that leads to the overall risk of cancer.

20:52

* Kresge 502 Cart: So then you can one of the the concepts they use in in twin studies by sort of parsing out the effects of dizygotic versus monozygotic twins is called heritability, which is a simple measure, which is what's the proportion of the variability in cancer incidence that you can attribute to inherited genetic factors, and so overall about

21:16

* Kresge 502 Cart: about 35% of the heritability of any cancer is due to inherited genetic factors. And then you can see this range. You have some cancers like colorectal cancer. We'll talk about colorectal cancer a lot about. We know a lot about modifiable lifestyle factors. And so genetic factors have a small role, but much more so environmental and lifestyle factors play a role

21:41

* Kresge 502 Cart: and colorectal cancer. In contrast, prostate cancer. And then melanoma seem to have these very strong genetic factors. It doesn't mean it's only genes. It could be genes interacting with the environment, but they both have a strong genetic predisposition.

22:06

* Kresge 502 Cart: So then, the question is, if anybody has any questions about any of these, just feel free to raise your hand. So then the question is, yes, the risk of cancer increases when the genetics and environmental

22:24

* Kresge 502 Cart: well, I think Melanova is an interesting example, because it also, you know, some aspects of melanoma are due to exposure to UV radiation. And so the question is, maybe there's an underlying genetic susceptibility to which people are going to be more or less susceptible to UV radiation. So that's that's an example. One example of this idea where genes and environment

22:38

* Kresge 502 Cart: interact together. Another example is in bladder cancer, where there's

23:04

* Kresge 502 Cart: certain genes that modify. Your ability to metabolize alcohol. And so depending, 2 people can drink the same amount of alcohol, but you get differences in how that gets metabolized, and then that then leads to an effect on cancer. So it's together the genes and the environment together. So that's an example of gene environment interaction.

23:13

* Kresge 502 Cart: So

23:40

* Kresge 502 Cart: so when we think about genetic factors, and we talked about this a little bit. So remember, about 99.6% of the inherited genome is similar across everybody in the room.

23:42

* Kresge 502 Cart: And and so the the variability that exists is actually quite small. The majority of that variability in genes are these

23:58

* Kresge 502 Cart: things that are more common. Single nucleotide polymorphism changes just a single variant change in the DNA and these common variants.

24:10

* Kresge 502 Cart: What we mean by common is usually that the prevalence in some population is probably 5% or more so, maybe not in every population, but in some populations about 5% or more. So these are very common, relatively common variation in the population. The flip side, though, is that most of them have very, very small effects on

24:23

* Kresge 502 Cart: diseases. So when we look at cancer and we'll talk about a genome, wide association study. These individual genetic variants that are seem to be associated with cancer across multiple studies. Sometimes their odds, ratios, or relative risk estimates are very, very small

24:46

* Kresge 502 Cart: at the other extreme. You have very, very rare

25:03

* Kresge 502 Cart: variants that cause disease in a Mendelian fashion, meaning, if you have the gene. Unfortunately, you're gonna get the disease. So Huntington's is a prime example of that, that. If you have the very rare, it's a very, that very, very high likelihood that you're gonna if not completely likelihood, that you're gonna get the disease I can't think of.

25:10

* Kresge 502 Cart: Maybe in cancer.

25:34

* Kresge 502 Cart: If there's that sort of Mendelian, maybe some of the cancer syndromes get close, but they're much more in this middle phase. I don't know Ed or anybody. If you can think of an example of a true Mendelian trait for a genetic variant in cancer.

25:36

* Kresge 502 Cart: There's some that still like, I think we think a lot about Brca. One and Brca. 2 mutations in the germ line being strongly associated with the risk of breast cancer, ovarian cancer. But they're not fatalistic. So not everybody gets them. So

25:55

* probably a fake

26:14

* Kresge 502 Cart: he was for? Yeah. So it gets close to a Mendelian trait.

26:16

* Kresge 502 Cart: Yeah. So so yeah, but you can kinda get this sense. There's this gradient and and the more rare the variants are the stronger the effect size is, and then more common variance tend to have very modest to low effect sizes.

26:25

* Kresge 502 Cart: I already talked about this. So about 5 to 10% of

26:50

* Kresge 502 Cart: cancers that have an inherited cause are due to these family syndromes or these more.

26:57

* Kresge 502 Cart: we look at it, these lower frequency variants, or even

27:08

* Kresge 502 Cart: the very rare variants causing Mendelian related traits.

27:13

* Kresge 502 Cart: But for these families that have these

27:19

* Kresge 502 Cart: germline mutations, they're pretty awful. The risk of certain cancers occurring is very, very high, and it gets passed from generation to generation. And we'll talk about a couple of these hereditary cancer syndromes specifically

27:22

* Kresge 502 Cart: and what they look like in the types of cancers that they have. But there's there can be many cases in a family with the same type of cancer, especially if it's uncommon or a rare type of cancer. And we looked at, we talked about that with the family history in the Swedish generational cohort, that when you saw 2 sisters, or you had a mother and sister, or mother and father a brother, your risk was was

27:43

* Kresge 502 Cart: matiring is probably more likely these were due to a hereditary cancer. Other features of hereditary cancer syndromes are cancers occurring in the sex not usually affected. I think breast cancer is an example of that. And if it's occurring in multiple generations.

28:07

* Kresge 502 Cart: so these are probably some of the most common of these hereditary cancer syndromes. I'm going to be talking about Leif Ramini, and a little bit about lynch syndrome, and I think when Dr. Song comes in lectures he'll talk more about lynch syndrome and its role in colorectal cancer and other cancers as well. And then I'm going to talk a little bit about Lee Rermini.

28:27

* Kresge 502 Cart: Dr. Lee was a faculty member here at the Harvard School of Public Health, together with Dr. Fermini, who was at the National Cancer Institute. We'll talk about as well. And then, of course. Brca, one and Brca 2, and breast cancer and ovarian cancer are probably pretty well known to you. Interestingly, they seem to be playing a role in other cancers as well. So again, these are the more

28:52

* Kresge 502 Cart: lower prevalence, more moderate moderate effects. So they're they're playing a role in the development of male breast cancer prostate cancer, probably pancreatic cancer as well.

29:16

* Kresge 502 Cart: So leaf Ramini, how do I get rid of this? Sorry? Yeah. Got it? Thank you.

29:28

* Kresge 502 Cart: So the lifer-mini syndrome was really identified by Doctors Lien formini back in 1,969. And they were really interested in the Gene, p. 53 which plays a broad role in cell growth in cell cycle arrest in DNA repair-related mechanisms.

29:38

* Kresge 502 Cart: It can play a role in apoptosis which is basically programmed cell death, and it can eliminate any damage cells. So it plays a really important role in cellular growth as well as making sure it gets rid of any bad.

30:02

* Kresge 502 Cart: This cell division, just as Dr. Weinberg was kind of talking about. It helps scan and make sure that any of these cell divisions are not resulting in mutations.

30:22

* Kresge 502 Cart: So however, in cancer, when you look, remember the difference between germline and somatic when you look somatically at tumors. So if you looked at the Germline, DNA. Take a blood specimen and look at the DNA that's going to really reflect the DNA that came from the parents.

30:33

* Kresge 502 Cart: You might not see at all probably won't, for in the most case, unless there's this family syndrome in the tumors, however, it can be quite common, and it's probably one of the most commonly mutated genes in in tumors across an array of cancers. So it's realized that when you basically have these mutations, all of a sudden, all these great repair DNA repair pathways, apoptosis, etc. Go away.

30:50

* Kresge 502 Cart: But what lean from Media, I realized, was actually there are people in which their germ line has a mutation already, and so, if you remember the 2 hypothesis

31:18

* Kresge 502 Cart: carcinogenesis. So if you, if you get one bad allele

31:30

* Kresge 502 Cart: with a p. 53 mutation. All you need is one more in order for cancer to be occurring. So so you're already kind of

31:37

* Kresge 502 Cart: more than halfway there in getting cancer. So how they identified this was, they studied families, and there was really an interesting array of cancers ranging from high prevalence of osteosarcoma, melanoma, brain breast cancer, leukemia, few others in these populations. Very early on set before the ages of 45,

31:46

* Kresge 502 Cart: and the risk of cancer in the populations was quite high, and they were able to show that this was due to p. 53 mutations that were inherited. So germline mutations were in this family.

32:09

* Kresge 502 Cart: and so unfortunately, while you can't probably prevention, for these families may be hard. These aren't populations that you'd wanna do much more intensive surveillance at much younger ages. So a very interesting kind of tidbit about in cancer was identified, probably, I think, about 7 or 8 years ago. When somebody asked the question, why do elephants have such a low

32:26

* Kresge 502 Cart: mortality from cancer. So with other animals, the larger the size, and the older they live, the greater the mortality from cancer. And so, if you compare the mortality from cancer in humans to that of elephants, you can see the lifetime mortality from cancers between, say, 10 and 25% in elephants in 4.8

32:54

* Kresge 502 Cart: in terms of average lifespan. It's pretty similar body weight much much higher. So there's more cells. So just by chance, you think you might develop more cancer in elephants? So you can see, this number of cells is is exponentially higher. What was determined, though when they did sequencing of the elephant was that while humans have 2 copies, right? We have 2 copies, alleles of elephants actually have 40.

33:19

* Kresge 502 Cart: So that's pretty interesting. So they all are slightly have slight modifications in the proteins they produce, but there's so much redundancy in this pathway, so, instead of having to knock out 2 alleles in elephant, you'd have to knock out 40, actually, as a tidbit.

33:48

* Kresge 502 Cart: I learned that mammoths. They also mammoths which are now extinct. They're able to get DNA from mammoths also show they have a very high copy of p. 53 alleles in their

34:06

* Kresge 502 Cart: in their DNA oops.

34:23

* Kresge 502 Cart: Okay, so lynch syndrome. And then feel free to just hop in on this. And I know, as I said, Dr. Song is gonna have a lot more detail. So lynch syndrome occurs when there's inherited mutations in a type of DNA repair called mismatch repair genes. So these are specific genes that are involved in a type of DNA repair, called mismatch, repair. And this leads to a syndrome

34:28

* Kresge 502 Cart: hereditary non polyposis, colorectal cancer and as as you were saying, egg what you can see here is that while the risk of colon cancer in the general population is about, say, 5 in a family with lynch syndrome, the risk of colorectal cancer is about 80. It's also quite elevated. The lifetime risk for indonesal cancer about 60% compared to about 2%

34:55

* Kresge 502 Cart: in the population. Also, elevated risks of stomach, ovarian cancer, and some other cancers as well.

35:20

* Kresge 502 Cart: Anything else sad?

35:33

* Kresge 502 Cart: Yeah, we'll be talking more about it. So in terms of the discovery. So all of those family syndromes are really again, these more moderate effect, size, genetic mutations, but whose prevalence is pretty low. So again, at a population level.

35:35

* account for so much of the genetic susceptibility to cancer. So a lot of what we know now about what causes cancer in terms of inherited genetics has come from doing genome-wide association studies. And so this is where you take.

35:56

* Kresge 502 Cart: Now, we understand, we really need to have sort of tens of thousands or hundreds of thousands of cases of cancer, you know, tens to 100,000 controls. And then look kind of it's really a discovery effort. You're looking at millions of individual, single, nucleotide polymorphisms and sort of saying, are any of these

36:12

* Kresge 502 Cart: much more common in cancer cases, or even less common in cancer cases? And the way we look at that is in something called the Manhattan plot. And the idea is, it looks like, you know, a cityscape. So basically, you're plotting on the X axis is what chromosome the Snps are on, and then on. Y-axis, you take the negative log, 10

36:32

* Kresge 502 Cart: p. Value. And so, and then, because you're looking at millions of single nucleotide polymorphisms. Usually you set your P value fairly, conservatively at 10 to the minus 8 here. And so any of those variants that are above that P value threshold are, you would say, are associated with the cancer you don't know

36:57

* Kresge 502 Cart: is the prevalence higher or lower, because it's just a P. Value. But these are the variants that seem to be kind of potentially interesting. But also this isn't meant to be a course on genetic epidemiology. But the question then becomes like, are these the variants that are causing the cancer to occur? Or are they just really close by to another genetic variant that maybe isn't tagged as well in the array, etc. So there's additional work you'd want

37:21

* Kresge 502 Cart: to do here. Now I can say so. This is actually a Manhattan plot from prostate cancer. And so all of these variants that are popping up here are variants that in multi-ethnic and in multiple populations are significantly and reproducibly associated with the risk of prostate cancer.

37:46

* Kresge 502 Cart: The individual effect of these variants, though, is really small. So the relative risk, if you looked, if you have this variant versus. If you don't, the relative risk might be 1.1 1.2. So really, really small.

38:04

* Kresge 502 Cart: But now, as of a couple of weeks ago, or maybe a few weeks ago, the latest publication, we've identified over 450 of these individual genetic variants. So then, the question is, well, maybe if one variant so you're not getting much of an effect. What if you took all of the variants together and created some sort of score? And that's what we do with apologetic risk score. So you're basically

38:17

* Kresge 502 Cart: looking at each of these 450 barons. And you're basically like, right, here's

38:40

* variant one. Do you have it? Yes or no.

38:46

* Kresge 502 Cart: plus variant 2,

38:50

* Kresge 502 Cart: yes or no, and maybe you weight it by the strength of the

38:53

* Kresge 502 Cart: you take your betas from the odds, ratios, and maybe so, if area one had an odds, ratio of 1.3 and variant 2 had an odds ratio of 1.1. This one's gonna get weighted a little bit stronger. And then you add those up. So now what you can see, though then in this Polygenic risk score. You can look at people who have the highest genetic risk, because maybe they have. You know, 100 of the 400 variants versus those people who have the lowest genetic risk. Maybe they have

38:58

* Kresge 502 Cart: 0. Maybe they have 5 of those risk variants. And you can really start to see big differences in risk prediction. And so this, I guess, is just to say, now,

39:28

* Kresge 502 Cart: you know, in terms of the these Gwoss studies have really helped to. Maybe maybe not so much. We can't think about prevention per se. But you can really start to think, could you use this like for early detection of cancer, for example. So if you knew somebody had they were in the top. This is the top one percentile of the polygenic risk, for so they probably have the most variance. Their lifetime risk of prostate cancer is 65.

39:39

* Kresge 502 Cart: If you're in the bottom tenth percentile of the genetic risk, your lifetime, risk of of prostate cancer is closer to to 2.

40:05

* Kresge 502 Cart: So you can really start to see this ability to risk stratify over a person's lifetime. All right. So we'll take it. Yes. understand how.

40:14

* Kresge 502 Cart: So I'm trying to connect it to the previous slide the man has. Are they related somehow? They are

40:29

* Kresge 502 Cart: this, I mean. This is an older version of of the Manhattan plot, but in this was an older. Anyway, it doesn't matter. Let's just say there are on this plot, 450 unique genetic variants, with a P. Value greater than 10 to the minus eighth.

40:39

* Kresge 502 Cart: that are associated with prostate cancer. And then you make a simple score, and you take each variant.

40:59

* Kresge 502 Cart: and for each person say, Do you have it or not? Yes or no? If you don't have it.

41:06

* Kresge 502 Cart: then that doesn't go contribute to the score if you have this one, and then let's say you don't have any of the others. Your total score is gonna be based solely on this one variant. But some people have 50 variants or 100 variants. And so, depending on how many variants you have, you can create these polygenic risk scores. And then, basically, all you're doing is dividing it into 10 equal groups.

41:13

* Kresge 502 Cart: It's a very simple quartiles or quantiles. For instance, if someone has less variance. Then there's going to be lower, right. If they have less variance, their score is going to be lower and their lifetime risk in this case of prostate cancer is gonna be much, much lower. So you might not see as strong of an a curve like this is basically looking at the absolute risk of cancer over someone's lifetime. Your risk

41:39

* Kresge 502 Cart: of of colorectal cancer with a polygenic score. It's probably gonna look pretty, not gonna look as impressive. Remember, prostate and melanoma were the 2 cancers in the twin studies as well as in family based studies that have the strongest inherited genetic component. Now we've identified a lot of genetic variants. And you can really start with these polygenic risk scores to be able to say something about future risk.

42:05

* Kresge 502 Cart: Do we use them, and do they copy? Not yet.

42:30

* Kresge 502 Cart: But people are doing that to say, could you use this as part of an early detection strategy? That's the question for cancer.

42:33

* Kresge 502 Cart: but we aren't using them clinically. Now.

42:44

* Kresge 502 Cart: now this that remember, this is snps. Now that's different than say for lynch syndrome.

42:48

* Kresge 502 Cart: where, if you are a family member and who's part of a family that's known to have Lynn syndrome? Then you are, Gonna use that genetic information. You're gonna go to a genetic counselor hopefully at a affiliated with a cancer center. And you're gonna be followed closely. And for a range of of potentially different cancers. So that's it is being used clinically in these kind of cancer syndromes.

42:56

* Kresge 502 Cart: There's other examples, too, but the Ji-wa snips, I would say in general, I haven't been translated yet. Any other questions.

43:26

* Kresge 502 Cart: Okay? So the next 2 slides are really getting at

43:37

* Kresge 502 Cart: basic descriptive measures of cancer.

43:41

* Kresge 502 Cart: Yep, so we have an incidence rate a mortality rate and a fatality rate. So pulling back from intro epie, what is an incidence rate? And how would you calculate it?

43:47

* Kresge 502 Cart: What are the 3 components that would go into calculating an incidence rate

44:03

* Kresge 502 Cart: number of new cases in

44:14

* Kresge 502 Cart: yep, right? And then the third piece of it. So you have. So the interest rate.

44:17

* Kresge 502 Cart: The numerator is the number of cases of cancer in the population in the population usually usually present it

44:23

* Kresge 502 Cart: per 100,000 people. What's the other element. That's important time. Exactly. Right? So usually in a year, you might say, so, what's the annual incidence rate of cancer, or every 5 years, or something.

44:32

* Kresge 502 Cart: Okay. So mortality and fatality, what's the difference?

44:47

* Kresge 502 Cart: What's mortality rate? Yep. And again, with time? Exactly right? So it's very similar to incidents, because the it's the number of cancer deaths.

44:52

* and then denominator instead, the number of cancers

45:10

* Kresge 502 Cart: divided by the population and then

45:15

* Kresge 502 Cart: over time.

45:21

* Kresge 502 Cart: Here. Ca, so then, what's fatality? What's the difference. There? I'm talking about fatality. Which group of people am I looking at? Am I looking at the whole population? No, what am I looking at?

45:23

* Kresge 502 Cart: Only the cancer cases. Exactly right? So fatality, then, and this is this gets

45:35

* Kresge 502 Cart: confused all the time. In fact, there was this paper that was like a response, like a letter to the editor. Actually, that was written about this exact thing because there was a paper that kept talking about mortality rates. But actually, but they were actually talking about fatality rates. So fatality, the denominator are just people cases with cancer.

45:42

* Kresge 502 Cart: And in the numerator there's the cancer deaths. And again, it's over time.

46:05

* Kresge 502 Cart: So can you think of a situation where you might have a high fatality.

46:11

* Kresge 502 Cart: but a low mortality?

46:18

* Kresge 502 Cart: Can you think of a a cancer

46:27

* Kresge 502 Cart: when the cancer has a very low incidence rate, but a high fatality. Exactly right? So cancers like pancreatic cancer liver cancer, they're kind of on the rare side. So in terms of

46:30

* Kresge 502 Cart: the population level, the mortality rate looks much lower than for other cancers. But really, when you look at the vitality, so what happens in people who have the cancer, it can look pretty pretty awful. And it is so that's an example where the mortality is telling you something a little bit different. It's really a population level statistic versus the fatality which is really telling you, what's the prognosis in people who have cancer?

46:45

* Kresge 502 Cart: Okay, so what's problems. Then

47:13

* Kresge 502 Cart: what's the prevalence of cancer? If you were trying to?

47:19

* Kresge 502 Cart: What statistic does that refer to

47:24

* Kresge 502 Cart: existing cases at that time? Yeah, yeah, exactly. And so, what are the factors that go into prevalence

47:31

* Kresge 502 Cart: like? Is there an example where you could have

47:43

* Kresge 502 Cart: trying to think like a high incidence, but a low prevalence.

47:48

* Kresge 502 Cart: Do you remember the formula for calculating prevalence?

48:00

* Kresge 502 Cart: Yes, there's yes, exactly right. So if you have a highly fatal cancer, the prevalence may look kind of low because the duration someone's living with that cancer is is very short. So even if the incidence of size. So basically, this is a simplified formula, incidence, times duration.

48:03

* Kresge 502 Cart: Looking at the

48:28

* Kresge 502 Cart: I was gonna write that. And I was like, that's right, right. It's a simplified formula. Yeah. So if your incidence is high, but your duration is low, your prevalence is gonna end up being pretty low. And then have you ever heard of the concept of years of life lost?

48:30

* Kresge 502 Cart: Yeah. And what is what if you could just like, sort of give a descriptor of what that is? Here's a life lost that's close. That's that's Dali's or disability adjusted life here. So here's a life loss literally is you know you not only count

48:49

* Kresge 502 Cart: that it's a death, but the age at which someone died. So it's sort of for for childhood cancers. These can rank very, very high because the amount of years of life lost can be quite significant. So that may be a figure. Oh, yes, yeah.

49:15

* Kresge 502 Cart: exactly so in the formula here they usually use 75. II don't know if they've changed it now, but they generally think like, if someone had lived a full life they would have lived to 75 years. How many years did they lose because they were a child, and they died at say age 8, or something like that for them they would have lost

49:34

* Kresge 502 Cart: 67 years of life, and then, for cancers like prostate cancer, which the age at which people were dying might be 75, 80, even though the mortality might be high. Actually the years of life. Loss is not as much because you're dying around this time you would have died anyway.

49:54

* Kresge 502 Cart: Yup, I think it was based on the on what the expect. Life expectancy was probably 2020, 30 years ago. So that's why I'm saying I don't know if they changed it, or if they always still use 75 but that was sort of, I think, based on the actuary tables in in within the setting of the United States, right? Because other countries have different age life expectancies.

50:16

* Kresge 502 Cart: Alright. So here, this is data. And we're gonna in the descriptive epidemiology project, the group project that you're gonna work on. You're gonna get to work with. Some of these international cancer epidemiology, databases that will provide you with incidence mortality, prevalence across different populations. So. This is the crude

50:46

* Kresge 502 Cart: incidence rate per 100,000 people of cancer in different countries, and the darker the blue, the higher the incidence rate, the lighter, the blue, the lower the incidence rate. So what might be a goal of looking at a map like this? What? What? Why might you want to

51:11

* Kresge 502 Cart: make comparisons like this across countries.

51:30

* Kresge 502 Cart: Or or maybe you could even just say, just looking at this map. What what does this tell you about this particular. I think this is total cancer incident. So what is the rate of cancer? Yeah, so it's it's, where is it? Maybe more concentrated? So this is just the crude rate per 100,000 people. Yup.

51:37

* Kresge 502 Cart: it can help visualize disparities. Yeah, exactly right. So are there populations that seem like there are greater risk for this cancer.

52:02

* Kresge 502 Cart: If I told you that the average life expectancy, let's say, in parts of Africa, might be 15 or 20 years lower than say in North America.

52:11

* Kresge 502 Cart: and how might you? How like? What might you want to try to do to

52:26

* Kresge 502 Cart: say, how much is there still a difference. We're independent of the fact that there's a different age of distribution.

52:31

* Kresge 502 Cart: Yeah, age adjust. Exactly. So you'd wanna account for those age distrib. And you can see this really clearly here. This is looking at the age distribution in countries that are considered to be low and medium resource, countries versus those that are high resource. So you have in the lower and medium resource countries. You have much younger

52:40

* Kresge 502 Cart: distribution of ages and more people in those age groups than you do in the high resource country. So if you wanna make these kind of comparisons, sometimes you wanna take away the effective age, and to do that, you can do age standardization. And it's just sort of a very simple concept, which is you basically would take the population into different age groupings.

53:03

* Kresge 502 Cart: And you take the actual number of cancer deaths in those age groups, the actual number of people in those age groups and for a specific year. And that's how you would get the actual crude incidence rate per 100,000 people.

53:25

* Kresge 502 Cart: That's generally how it's presented. And then you'd say, Well, I'm going to take the same

53:42

* Kresge 502 Cart: distribution of ages and apply it to each country. So that the basic question, like, What would the if I assume that the rates of cancer. In these age groups are

53:47

* Kresge 502 Cart: true rates. Then I apply a different distribution of age. So everybody has the same age distribution. What does that age adjusted rate look like. So does that make sense kind of what the distribution is doing. So basically, like.

54:01

* Kresge 502 Cart: now, every country is going to have this same distribution of age in it. But I'm taking the actual rates of cancer in the individual countries for those ages. And then I'm getting an adjusted rate.

54:15

* Kresge 502 Cart: Does that make sense.

54:33

* Kresge 502 Cart: Alright. So we're gonna take a, we're gonna have a little breakout now. So this is a real example. So there was a group in upstate New York that wanted to look at

54:34

* Kresge 502 Cart: prostate cancer mortality rates to see if there were racial disparities in in prostate cancer mortality. So they had the number of cancer specific deaths in white men and black men. They then had the number of men living at the time. Who were either white or black, and then per 100,000 people per year, and

54:47

* Kresge 502 Cart: they calculated the prostate cancer mortality rate in white men to be 28.7, and in black men to be 28.9 per 100,000.

55:15

* Kresge 502 Cart: So 2 pieces of data to think about one is that here's the distribution of age

55:28

* Kresge 502 Cart: in white men.

55:35

* Kresge 502 Cart: And here's the distribution of age in black men. So just to say that black men had generally a younger age overall in the population compared to white men.

55:37

* Kresge 502 Cart: Right? So you can see that you have more older

55:49

* Kresge 502 Cart: men who are white than in the black men. And the other fact is that mortality from prostate cancer increases really exponentially with age, so that the rates are about 100 times higher in men 65, and older compared to younger men.

55:54

* Kresge 502 Cart: Alright. So the question is to have a quick discussion is, if you were to adjust the rates between black and white men. So they have the same age distribution. What do you think would happen? And what do you think you might see in terms of whether or not there might be differences in mortality between white and black men. So take just a minute or 2 and talk about how you think the mortality rates adjusted for age would look like between these 2 groups of individuals.

56:11

* Kresge 502 Cart: logic of what you think

59:13

* Kresge 502 Cart: could potentially happen. So so the question that we're trying to answer is.

59:40

* Kresge 502 Cart: are there racial differences in prostate cancer mortality that are not due to age? Right? So that the question is really is that we don't want it to be about age differences, right? Because right now, on average, the black men in the whole population are younger than the white men. So we wanna say, is there a difference?

59:47

* Kresge 502 Cart: A racial difference? We see that is separate from any difference that there might be from age. So what what were your thoughts qualitatively on, on what might happen if you did an age adjustment or age standardization?

1:00:07

* Kresge 502 Cart: I said, there's there's 3 potential answers. Right? 1 one answer is, there's no difference in the the crude estimates are gonna look similar to the A. The age adjusted estimates. One estimate is that the black men will look worse, and the other is that the white men will look worse. So so just show hands who, you know, discussions who thought they'd look the same.

1:00:22

* Kresge 502 Cart: and then who thought they would look? The white men would look like worse mortality.

1:00:49

* Kresge 502 Cart: And then who thought the black men would look like they had worse mortality. So what was your logic behind that?

1:00:55

* Kresge 502 Cart: Oh.

1:01:02

* Kresge 502 Cart: yes.

1:01:06

* Kresge 502 Cart: just for

1:01:13

* Kresge 502 Cart: exactly. And it's because there's such a strong association between age and prostate cancer mortality rates. So even if the mortality rate in each of the age groups was sort of similar, you're gonna upweight. What's happening in the black men. And you're gonna downweight. What's happening in the white men? So ultimately, you're gonna see that there are racial differences in mortality from prostate cancer from this population. Exactly.

1:01:16

* Kresge 502 Cart: Okay. Great. So now I'm gonna go into some of like the core epidemiology basics. And I'm gonna you know, I everyone has taken the intro to epidemiology core. So just as a reminder, there's kind of 2 core study designs we use. The the cohort study and the case control study. This is a hypothetical example of looking at the association between infection with HIV

1:01:46

* Kresge 502 Cart: and the risk of Non Hodgson lymphoma. We will talk a little bit about this in the infections in in cancer lecture. So in this case, you'd start with a group of individuals where nobody had cancer at the start of follow up, so you wouldn't want anybody to have the cancer of interest at the start of follow up, and you'll know there, there'll be some people who are were exposed to HIV infection and

1:02:11

* who weren't. And then you can follow

1:02:36

* Kresge 502 Cart: these groups forward and see which of them develop non-hodx and lymphoma, and which of them do not? So that would be sort of the cohort. Approach. And even, you know here I've just made it very simple that you're either exposed or unexposed at the start of the study. You then get a little fancier and say, I'm gonna keep asking about HIV infection over time and and people might become exposed later on.

1:02:39

* Kresge 502 Cart: So we do case control studies where control studies are often done in cancer, particularly when cancers are rare

1:03:03

* Kresge 502 Cart: right? Because even if you had a cohort of 100,000 individuals for some cancers like glioma, or maybe childhood cancers or cancers that are fairly rare. You're gonna have to have such a large cohort and follow people for sufficiently long enough to be able to have enough cancers cases in in your cohort. Just as an example

1:03:14

* Kresge 502 Cart: we work with Ed and I and Michelle and colleen work with a cohort called the Health Professionals follow up study. It was started in the 1,986 50,000 people who identified as men at the start of the study. They've been followed now for over 36 years, and I think there are just over about 150 liver cancer cases.

1:03:39

* Kresge 502 Cart: There's some bilary tract cancers, probably 40 or 50 that have occurred. So you can really see, even though we had this really large study and a really long follow up, there's still some cancers where the the incidence rate is so low in that population that's not really sufficient. So in that case, what you could do instead is identify your cases in some way. Think about

1:04:03

* Kresge 502 Cart: what's the source population. That's not always easy, and we'll give a quick example of that in a minute. And sample people who don't have the cancer. And then assess the exposure. And one of the really important things is that you want the exposure frequency in the controls

1:04:25

* Kresge 502 Cart: to represent the cohort

1:04:43

* Kresge 502 Cart: that gave rise to the cases. You basically want the at baseline that controls should give you an estimate of the exposure distribution in the cohort that gave rise to the cases, and if you do a case control study well, it can be really an efficient way of an alternative to a cohort study. But there are a lot more issues of bias. And we'll talk about those kind of in a moment.

1:04:47

* Kresge 502 Cart: And

1:05:11

* Kresge 502 Cart: you know, just when we think about what what things do we measure in a cohort study versus a case control study? We often will measure. I think the big focus is really on relative measures. So what's the relative risk of cancer in the exposed group versus unexposed group? You might calculate an odds ratio. Maybe it's a hazard ratio. You can also think about absolute measures of a factor of risk differences.

1:05:14

* Kresge 502 Cart: Another measure that you might see in epidemiology studies relate to either attributable fraction or population attributable fraction. So the difference there is with the attributable fraction. You're just saying of the cases.

1:05:39

* Kresge 502 Cart: what number or what proportion of the cases of cancer can be attributed to the exposure, whereas in the population attributable fraction, the question is of

1:06:01

* Kresge 502 Cart: in the whole population, what proportion of cancer can be explained by the exposure. So so they give you kind of slightly different measures. And so the idea is within the triple fraction. If you got rid of the exposure. Let's say it's smoking. Let's say it's an occupational exposure. If you could completely get rid of it

1:06:13

* Kresge 502 Cart: in the cases, what proportion of the cases would be prevented versus in the entire population. You're taking the prevalence, not in the cases, but in the full population.

1:06:37

* Kresge 502 Cart: Can you explain again the difference between the 2. Yeah, sure. So I think the big difference really is that the prevalence of

1:06:49

* Kresge 502 Cart: well, let me talk about a trivial fraction first. So here we had 1,400 cases. Let's say we had. And you can see there's a population of 100,000 people.

1:07:09

* Kresge 502 Cart: You have 1,400 cases in total, 900 of them didn't have the exposure in this situation. So the the attributable like none of the cases there can be attributed to the exposure. So you're just looking at the exposed cases. What per proportion of those

1:07:21

* Kresge 502 Cart: can be attributed to the exposure? And this is really driven by the strength of the association of so something like lung cancer and smoking. You're going to have a much larger proportion of the cases, or is attributed to smoking than you would for something where maybe the relative risk estimate is much smaller. What's different here is in the population. Trivial fraction is

1:07:41

* Kresge 502 Cart: that you're looking at a different prevalence of exposure. You're using the prevalence of exposure from the full population instead of from the population of cases.

1:08:10

* Kresge 502 Cart: Am I kind of

1:08:21

* Kresge 502 Cart: very simplistically, I'll look to my teaching fellows for that. Yeah.

1:08:23

* Kresge 502 Cart: alright. So let's talk about confounding. This is a real example actually of a project that I was working on with a student looking at physical activity and lung cancer. And this was a cohort of 40,000 women. They had at study enrollment, collected detailed information about how much physical activity, what they were doing, what types of physical activity they were doing.

1:08:33

* Kresge 502 Cart: and if, looking at the association between participating in the highest quintile of physical activity. The relative risk of lung cancer was

1:08:57

* Kresge 502 Cart: 0 point 2 5 compared to women who have the lowest levels of activity. So basically saying, that you could lower your risk of of lung cancer by 75% by engaging the most physical activity. So if you looked at this. Can you think of a confounding factor that you might be worried about that? It's not physical activity. That's the risk factor, but that there's there's some additional factor that's correlated with physical activity.

1:09:08

* Kresge 502 Cart: And on its own as a prospector for lung cancer. Okay? Bmi, that's interesting. Yep, smoking. Yeah. Any other awesome.

1:09:38

* Kresge 502 Cart: Yeah. So I think smoking was probably the one that was the strongest cause. The the association between smoking and lung cancer is probably a relative risk of about 10 to 20. Remember from that figure over the lifetime. Yeah, Bmi is interesting, cause it's so strongly related to physical activity. And Bmi is an interesting one. And we do talk about this at all about like Bmi and lung cancer. And what happens like.

1:09:48

* Kresge 502 Cart: you know, some people who have regressive forms of cancer might lose a little bit of weight before the cancer diagnosis because of the cancer is not detectable. But pancreatic cancer is a big example where you actually have weight loss before diagnosis. And then, of course, with smoking you might look like you have a lean body mass because of the smoking. So body mass is a really interesting one

1:10:13

* Kresge 502 Cart: for this. So anyway. So by adjusting for smoking, actually, and what we did was first we adjusted for just ever smoking. So the relative risk went from being really really strongly protective to still looking pretty protective. But then, when you also adjust for the amount of cigarette smoked and the years someone smoked.

1:10:37

* Kresge 502 Cart: Then essentially, what you can see is that the relative risk went almost to one meaning. There was no association between physical activity and lung cancer that the observed dissociation you had was completely due to confounded by smoking.

1:11:00

* Kresge 502 Cart: Okay, so here's here's an example I want you to think about and remember what I talked about with case control studies. So this is a study that was led by 2 full professors at the time. And some other family members from the Harvard School of Public Health. It was actually one of them was my advisor, and he always used this as an example of what can go wrong in studies.

1:11:17

* Kresge 502 Cart: So they published a study in the New England Journal of Medicine. It was a case control study looking at the association between coffee consumption and pancreatic cancer. It's a hospital-based case control study.

1:11:41

* Kresge 502 Cart: and they recruited 369 cases of pancreatic cancer from 11 hospitals in Massachusetts and Rhode Island. And then they said, Remember it was a case control study. So they had to say.

1:11:56

* Kresge 502 Cart: what was the source population that gave rise to these cases, meaning people who might have other diseases, who still would have gone to the hospitals had they developed the cancer. And so what they decided to do was to sample from the same 11 hospitals in gi clinics, from the same physicians where the pancreatic cancer patients were.

1:12:09

* Kresge 502 Cart: and they had either other cancers or gi conditions. When they did this analysis they found that the relative risk of pancreatic cancer, if you drank one to 2 cups of coffee per day was 1.8, meaning an 80% higher risk compared to no drinking coffee, and if you drank 3 cups of coffee per day. It went as high as 2.7 cups per day, and they adjusted for cigarette smoking, which

1:12:33

* Kresge 502 Cart: they thought could be an important confounder. So it used to be that people who drank a lot of coffee were more likely to be smokers, and smoking is a risk factor for pancreatic cancer. So they still found a positive association.

1:12:58

* Kresge 502 Cart: So talk for a couple of minutes and think about given the design. Why do might they have seen this positive Association Association. Do you think it was causal adjusted for cigarette smoking they had just for other confounders, or what might it? What could have happened in the study design that led to this finding. So talk. This is a little bit more complicated, maybe talk for like 3 or 4 min, and then we'll come back together.

1:13:10

* Kresge 502 Cart: Ok, so let's hear your thoughts on on some of the discussions you've had.

1:17:00

* Kresge 502 Cart: So

1:17:15

* Kresge 502 Cart: so what were some of the things you talked about in your groups.

1:17:19

* Kresge 502 Cart: No, but it's actually really nice to call out each other and support each other, too.

1:17:32

* Kresge 502 Cart: like defining the expansion of copy not being shared with

1:17:53

* Kresge 502 Cart: clocking conceptions that

1:18:02

* I think I said

1:18:05

* Kresge 502 Cart: I just so could be. It could be a couple of things. It could be sort of recall bias in a way that, remember, comes to the other points or the pancreatic cancer is really ruminating about their causes. So they're gonna recall things differently. But then the other thing you mentioned, and so go into that a little bit more when you said you don't know if it's

1:18:09

* Kresge 502 Cart: the exposure before or after the diagnosis. So what was your thought there in saying that specifically. or did somebody else want it? I don't want to put you on the spot.

1:18:33

* Kresge 502 Cart: Did somebody else want to kind of follow up on that.

1:18:47

* Kresge 502 Cart: Do you think having pancreatic cancer could impact how much coffee somebody's consuming potentially? Maybe you're feeling sick. Or maybe you're you're having some gi symptoms. And so if you're if you sort of yes.

1:18:52

* Kresge 502 Cart: yes, yeah. So these are both great points. So in the cases you really want to think about what? What is the window of time? That somebody might be reporting on their coffee consumption. So so remember the odds ratio essentially. The way you're calculating is what's the odds of exposure in the of the exposure, in the cases

1:19:14

* Kresge 502 Cart: divided by the odds of exposure and control. That's sort of one way to think about the odds ratio.

1:19:42

* Kresge 502 Cart: right? So you could have recall bias right? That is, these people are ruminating, and they're overreporting. How much coffee they had compared to. And let's say, there was.

1:19:52

* Kresge 502 Cart: There was no issue here so that could inflate the odds ratios potentially. So the point about the what's interesting about in this scenario, but great thinking about how a case controls so they can infect the exposure. So it's probably less likely that you know. Let's say they even let's say you even asked like how much coffee you've been drinking the last year. Maybe they had been drinking less because they were feeling ill

1:20:05

* Kresge 502 Cart: in this case for the cases that would actually then lower the odds. Ratio right? But the point that you just raised, which is a great one is.

1:20:29

* Kresge 502 Cart: look at where they got the controls from. These are people with other Gi conditions. So they're the ones like you want the control to represent the exposure distribution in the population that gave rise to the cases. So this is too low, right?

1:20:39

* Kresge 502 Cart: And so then, that can also inflate the odds ratio as well. And so they actually did a follow up study to this doing a more appropriate selection of the controls. So they weren't able to get deal with recall bias. But we'll talk about, recall, recall bias in a second, but they were able to then show that there was no association. Between coffee and pancreatic cancer when they more appropriately dealt with the controls in that way.

1:20:56

* Kresge 502 Cart: Okay, so let's talk about recall Bios. So this was one of the papers we were hoping that you had a chance to read before class. And so

1:21:24

* Kresge 502 Cart: who wants to give like a really quick synopsis of what? That

1:21:37

* Kresge 502 Cart: study the design of that study, because it was a pretty cool study that Dr. Jovanucci did as a student. Were you a student then? Or you just graduated?

1:21:43

* Kresge 502 Cart: Yeah, I think it was 1991. Was the publication. Yeah. So you might have like, yeah, graduated. And then it took a little while to publish it. Yeah, I think it was 91. Maybe it was a little later. So what was the really cool design that Dr. Jovanuji did

1:21:57

* Kresge 502 Cart: so. I guess. What was the premise? What was the premise?

1:22:20

* Kresge 502 Cart: Why did you do this study

1:22:29

* Kresge 502 Cart: so? He wanted to kind of. So there was a hypothesis out there that consumption of dietary fat

1:22:36

* Kresge 502 Cart: was associated with an increased risk of breast cancer, and most of the studies

1:22:44

* Kresge 502 Cart: that had been done on this topic had had been done in case control studies. And so there was a concern of

1:22:51

* Kresge 502 Cart: recall bias. Potentially. Again, this idea that people might be ruminating. So he did this really cool design, which was, he had access to a cohort of nurses

1:23:00

* Kresge 502 Cart: called the nurses health Study. There was a food frequency questionnaire that was asked and collected all sorts of information about the types of food someone ate, and they were able to calculate different types of of dietary fats from these foods, and then they followed people forward

1:23:14

* Kresge 502 Cart: for breast cancer. So this was the cohort analysis. This is going to be in theory 3 of recall bias.

1:23:32

* Kresge 502 Cart: So they identified 398 breast cancer cases. They also did age-match controls from the population, and then in 1,989. They went back and said, Hey, 3 years ago, tell us about what you ate right. But they actually had what they had reported in 1,986 before any cancer had occurred. So then they could.

1:23:40

* Kresge 502 Cart: He had both. The recall diet after cases were diagnosed, and the actual diet that was recorded before. So does that make sense kind of what he was able to do.

1:24:04

* Kresge 502 Cart: And so what did you find? So this is looking at odds, ratios, and 95% confidence intervals. The first column is looking at the Association of Total Fat Consumption of breast cancer risk, and then the second one specifically saturated fats in the risk of breast cancer. The first line of data is prospective cohort analysis, and then the other. One is

1:24:15

* Kresge 502 Cart: more of the case-controlled techno analysis. So what did they see in those 2 different approaches? And do you think this showed? Recall bias or not?

1:24:38

* Kresge 502 Cart: Yeah. So what does it say to you in looking at the data? If you were writing it up in a paper, what would you

1:24:56

* repeat the models of quality.

1:25:08

* Kresge 502 Cart: Yeah. And it was specific to the cases because they were the ones really, they had the disease. So they were really thinking of it. And that's why you get this inflation? Because it was just in the cases there might be misremembering on both groups. But that would be non-differential. But here it was clearly a differential type of bias. Exactly. Yep.

1:25:23

* Kresge 502 Cart: yeah, right? That's a great point. So these, this is a cohort of nurses who are working in healthcare. They may be more familiar even than a general population with the literature that, hey? There might be this connection between dietary, fat and breast cancer risk. So anything you wanna add, or

1:25:48

* Kresge 502 Cart: hypothesis made of time? It's made of a graph

1:26:10

* Kresge 502 Cart: case control. Did you say.

1:26:15

* Kresge 502 Cart: Teresuba? I said that. Well, this probably only applies to the US.

1:26:42

* Kresge 502 Cart: Okay, right? Right? So that's interesting. So it could be a concern of recall bias in one population. And, as you said, because they're nurses and more engaged in the healthcare system. Maybe it's more of an issue there, that's really interesting. Yeah.

1:26:48

* Kresge 502 Cart: so I think just to kind of wrap up. We've talked a little bit about measurement error, which you know, it's usually a non differential type of bias. You have recall bias you could have selection bias which we talked a lot about with the pancreatic cancer example. We talked a little bit about this, but and I think we'll talk about this in depth with pancreatic cancer. But other cancers as well. There's this sort of preclinical

1:27:09

* Kresge 502 Cart: phase where the cancer exists, but it hasn't been diagnosed as of yet in some cancer patients, when the cancer is more aggressive, can actually lead to weight loss. So there were some studies in a pancreatic cancer looking at body mass index and pancreatic cancer, suggesting that there was an inverse association. So that higher Bma was protective against

1:27:37

* Kresge 502 Cart: pancreatic cancer. But it really was because of this reverse causation that it was the cancer leading to weight loss. Now, this is gonna be interesting because Brian Wolfen is gonna come talk pancreatic cancer, how they're taking reverse causation and turning it into an early detection opportunity, right? Because the weight loss, then, is a signal that you could use to detect pancreatic cancer earlier, some other things as well. So it's interesting to think of a bias actually being an advantage.

1:28:01

* Kresge 502 Cart: and then we're gonna have a great lecture that Michelle and colleen are gonna lead talking about survival analyses and that's gonna be great and talk a little, maybe a little bit about this kind of bias as well. That'll be later. In the course. And then finally, I just wanted to raise in cancer. You know, we talk about confounding one of the strongest confounding factors can actually be screening

1:28:31

* Kresge 502 Cart: right screening when you screen somebody for cancer. The risk is much much elevated right? You know. Psa, a prostate specific antigen for prostate cancer, colonoscopy for colorectal cancer. Is so strongly related to the outcome. And then we know there's sort of like this. Healthy, you know, kind of lifestyle

1:28:59

* Kresge 502 Cart: affect that. People who eat a healthy diet and engage in physical activity and maybe don't smoke are probably the ones more likely to get cancer screening. So diagnostic biases. Really, I think of really a form of confounding due to screening habits. And so that's something that's gonna be highlighted across different lectures and something as you're reading the literature that you might wanna be thinking about. For example, in your

1:29:23

* Kresge 502 Cart: critiques. You might want to think about it if it's an issue. So

1:29:47

* Kresge 502 Cart: just to wrap up in the last.

1:29:54

* Kresge 502 Cart: No, I'm not going to, because, given the time. I'll I'll bring these last few slides into our lecture on Tuesday. Talking about like latency. I think, will be great. So just to wrap up a reminder that the office hour is going to be on

1:29:58

* Kresge 502 Cart: next next Thursday after class, and it will be recorded so if people can't make it. That is great. Remember, at 50'clock today

1:30:18

* Kresge 502 Cart: the Harvard canvas is gonna open. So you can sign up. For which of the cancers you wanna work on and then finally, we're gonna put in the link. The journal of the National Cancer Institute, just came out recently with an announcement. They really wanna they wanna focus on early career investigators. And so we're gonna put a link in that talks about if you're interested in submitting to that journal

1:30:33

* Kresge 502 Cart: articles, or if you want to get experience. Being a peer reviewer for the journal, which is a great thing to have on your Cv. You can sign up, so we'll be sending that link out anything else that we want to have a great weekend, and we'll see you on Tuesday.

1:30:57

* Should I end the meeting for everybody. Yeah.

1:31:24

* Kresge 502 Cart: exactly. So. So again, the trivial traction just refers to. Only among the cases a proportion of the cases could prevent. If you got rid of the exposure of just the prevalence of the exposure. And so I think the key difference with the population attributable

1:31:57

* is, instead, you're using the prevalence of the exposure in the full population, which often can be lower. So, for example, if you were looking at, say, smoking and lung cancer.