Naar aanleiding van wat feedback nog 1 poging om het probleem in de analyse opzet van BJM (Bakker Jacobs Meester) in hoofdstuk 6 eenvoudiger weer te geven!  
  
Wanneer we cohort/survival analyses opbouwen, willen dit zo opzetten zodat eerlijk vooruit wordt gekeken. We willen een goede prospecto (vooruit) scoop (kijker). Welke bekende fout gaat er met hun prospectoscoop?

Laten we hiervoor het hele politieke/wetenschappelijke context van coronavaccinaties lostlaten voor wat het is, en een gedachtenexperiment doen. Eigenlijk denken we na over het ontwerp van zo’n prospectoscoop, en in welke situaties die zal voldoen.

Laten we ons voorstellen dat we een homogene populatie gaan volgen in jaar 2021, en dat die populatie een homogeen sterfte patroon heeft, met en overzichtelijke maar hoge sterfte: van de overgebleven populatie overlijdt elke week 1%.

Laten we ook zeggen dat vanaf week 0 van dat jaar, een willekeurige 10% van de beginpopulatie elke week wordt blootgesteld aan een prik met zoutwater oplossing gedurende 3 weken en daarna niet meer.

Of ze terminaal zijn etc, maakt niet uit- er is dus geen confounding. We zijn in eerste instantie met name benieuwd naar het sterfte risico binnen een week:

We verwachten van een prikje zoutoplossing niet zoveel, dus een goede prospectoscoop zou geen verband moeten laten zien tussen de prik en sterfte. Vereenvoudigd zou dat er, als er elke week apart wordt gekeken en alleen in die week sterfte wordt bekeken het er zo uit moeten zien:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Week | w1 | w1 doden | w1 % | W2 | W2  doden | W2 % | W3 | W3 doden | W3% |
| Nog niet gevaccineerd | 9000 | 90 | 1% | 7910 | 79 | ~1% | 6831 | 68 | ~1% |
| gevaccineerd | 1000 | 10 | 1% | 1000 | 10 | 1% | 1000 | 10 | 1% |

De tabel hierboven kijkt naar de groepen afzonderlijk. Als we geintereseerd waren in de algemene sterfte tussen de groepen dan zou dat er zo uitzien:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Week | w1 | w1 doden | w1 % | W2 | W2  doden | W2 % | W3 | W3 doden | W3% |
| Niet gevaccineerd | 9000 | 90 | 1% | 7910 | 79 | ~1% | 6831 | 68 | ~1% |
| gevaccineerd | 1000 | 10 | 1% | 1990 | 20 | 1% | 2971 | 29 | ~1% |

Maar wat BJM doen, is een prospectoscoop bouwen waarbij ze de groepen kunnen indelen in:  
- deze week gevaccineerd  
- toekomstig gevaccineerden  
- nooit gevaccineerd.  
  
en gooien dan de toekomstig gevaccineerden weg!

Als we alleen in de eerste week kijken, zijn de delers voor die groepen:

|  |  |
| --- | --- |
| Week | w1 |
| Nooit gevaccineerd | 7000 |
| Toekomstig gevaccineerd | 2000 |
| gevaccineerd | 1000 |

En nu komt het belangrijk punt. Wat is de sterfte? Die teller van voor toekomstige gevaccineerden voor week 1 **moet** 0 zijn in die specifieke week! en dan komen alle 90 doden van de eerste tabel in de nooit gevaccineerde groep, terwijl de deler wel wordt verkleind. De geobserveerde sterfte in deze greep wordt dan:

|  |  |  |  |
| --- | --- | --- | --- |
| Week | w1 | Doden | W1% |
| Niet gevaccineerd | 7000 | 90 | **1.3%** |
| Toekomstig gevaccineerd | 2000 | 0 | 0% |
| gevaccineerd | 1000 | 10 | **1%** |

Dat is geen goede prospectoscoop dus. Ik heb in het voorbeeld maar 30% vaccinatiegraad aangegeven. Maar als uiteindelijke vaccinatiegraad hoger is, wordt in het ergste geval wordt het voor deze steekproef (85% vaccinatiegraad uiteindelijke vaccinatiegraad):

|  |  |  |  |
| --- | --- | --- | --- |
| Week | W1 | Doden | W1% |
| Niet gevaccineerd | 1500 | 90 | **6%** |
| Toekomstig gevaccineerd | 7500 | 0 | 0% |
| gevaccineerd | 1000 | 10 | **1%** |

Nogmaals, dit is alleen in de eerste week. Maar we kunnen dit ook uitbreiden in de tijd en dan kruipt het wekelijkse % dichterbij. Nogmaals de teller moeten we van de eerste tabel nemen.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Week | w1 | w1 doden | w1 % | W2 | W2  doden | W2 % | W3 | W3 doden | W3% |
| Niet gevaccineerd | 1500 | 90 | **6%** | 1410 | 79 | **5,6%** | 1331 | 68 | **5,1%** |

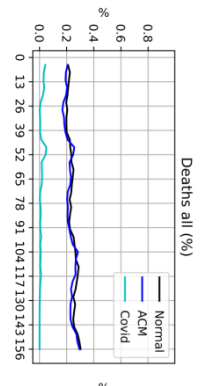
Dat is logisch, zoals hierboven hebben laten zien is het verschil gedreven door kunstmatige verkleining van de deler door toekomstig gevaccineerden uit te sluiten. Hoe meer je in de tijd komt, hoe meer de ‘nooit gevaccineerde’ deler dichterbij de toekomstig gevaccineerd+nooit gevaccineerd deler gaat lijken, totdat de campagne stopt.

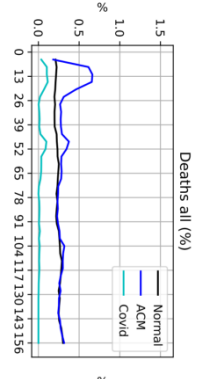
|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Week | W1 | W2 | W3 | W4 | W5 | W6 | W7 | W8 | W9 |
| Nooit gevaccineerd | 1000 | 1000 | 1000 | 1000 | 1000 | 1000 | 1000 | 1000 | 1000 |
| Toekomstig gevaccineerd | 8000 | 7000 | 6000 | 5000 | 4000 | 3000 | 2000 | 1000 | 0 |

Omdat de vertekening in de tijd terugloopt, komen we gedurende de campagne op een vertekening die gemiddeld misschien 4% zal zijn. Omdat het basis risico 1% is, lijkt het dus een factor 4 vertekening te geven. Dit lijkt ongeveer overeen te komen met het relatieve factor 4 verschil tussen hun analyse en de RIVM analyse.

[quote eerder draadje]

Omdat we steeds met een overzichtelijke sterfte risico van 1 hebben gewerkt, kunnen we verder denken hoe resultaten elders vertekend worden door deze greep, zoals in het tweede gedeelte van het hoofdstuk:





De belangrijkste observatie dat wordt gedaan daar is dat de sterfte omhoog schiet als de vaccinatiecampagne in de gematchte tijdsseries (onderste plaatje). We kunnen omrekenen met ons gedachte experiment ook hoeveel de geobserveerde sterfte omhoogsciet.

Laten weer de figuur pakken, die laat zien dat voor cohort 1940-1950 de uiteindelijke vaccinatie graad 90% is.

In de eerste week is de deler in ons gedachtenexperiment 1000, met wederom geen wijziging in de teller (90). Tot nu toe hebben we de 3:1 verhouding er niet bijgepakt, omdat het ging om het directe verschil tussen gevaccineerde en ongevaccineerden.

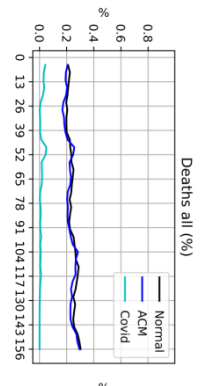
Maar voor de vertekening van de algehele sterfte is die verhouding wel belangrijk. De oorspronkelijke deler gaat van 1000:1000 naar 1000:333. De % sterfte blijf hetzelfde, en dan vullen we een nieuwe teller in:

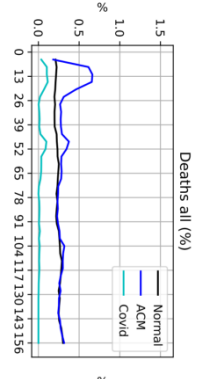
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Week | Deler alles | Teller alles | % alles | Deler 3:1 | Teller 3:1 | % 3:1 |
| Nooit Gevaccineerd | 1000 | 90 | **9%** | 333 | 30 | **9%** |
| Gevaccineerd | 1000 | 10 | **1%** | 1000 | 10 | **1%** |
| Totaal | 2000 | 100 | **5%** | 1333 | 40 | **3%** |

De totaal waargenomen sterfte ten opzichte van wat hij zou moeten zijn (1%) is 3%. Drie keer hoger dus.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Week | Deler alles | Teller alles | % alles | Deler 3:1 | Teller 3:1 | % 3:1 |
| Nooit gevaccineerd | 1000 | 90 | **9%** | 333 | 30 | **9%** |
| Gevaccineerd | 1000 | 10 | **1%** | 1000 | 10 | **1%** |
| Totaal | 2000 | 100 | **5%** | 1333 | 40 | **3%** |

Hieronder de verschuiving in de totaal waargenomen sterfte in % is in de grafieken van 0.2% gesprongen naar ergens tussen de 0.6-7%.





Ik laat iedereen zijn eigen conclusies trekken. Maar tag toch even betrokkenen @jonawalk & anonieme medestander @bramkoers\_tweets Ik denk dat @bonneklok dit ook wel interessant vindt.

Op aanvraag kan ik ook simulaties laten zien met kaplan meier grafieken.