

Hur man mäter effekten av prevention

Den statistiska aspekten

Anders Ahlbom

I denna artikel beskriver Anders Ahlbom, forskarassistent och statistiker vid soc.med.-institut., Huddinge sjukhus, statistiskt metodologiska synpunkter på problemet om hur effekten av prevention skall mätas. Dels diskuteras några principiellt olika sätt på vilka undersökningar av detta slag kan läggas upp och dels diskuteras en del faktorer av betydelse vid sådana preventiva åtgärder, där screening ingår som ett moment.

Artikeln är en bearbetning av ett inlägg vid symposiet om utvärdering av prevention, Svenska läkaresällskapets riksstämma 1975.

Inledning

Syftet med denna artikel är att ur statistisk — metodologisk synpunkt belysa problemet om hur effekten av prevention skall mätas. Inledningsvis kan framhållas att redan vid planeringen av en sådan undersökning bör en statistiker knytas till projektet. Valet av statistisk metod beror till en förhållandevis liten del i valet mellan chi-två test och t-test, utan framför allt i valet mellan olika sätt att lägga upp en undersökning. I planeringen bör en statistiker och en medicinare gemensamt diskutera fram en precis frågeställning och en undersökningsuppläggning som svarar mot denna.

Artikeln är fortsättningsvis disponerad i två avsnitt. I det första diskuteras några principiellt olika sätt på vilka undersökning av detta slag kan läggas upp. I det andra diskuteras en del faktorer av betydelse vid sådana preventiva åtgärder där screening ingår som ett moment.

Olika sätt att lägga upp en undersökning

Med en prospektiv studie menas en studie där en preventionsgrupp och en kontrollgrupp följes framåt i tiden och jämföres. Med en retrospektiv studie menas en studie där individer med en viss sjukdom eller ett visst symtom jämföres med en kontrollgrupp med avseende på om de tidigare blivit utsatta för den preventiva åtgärd vars effekt analyseras. Terminologin på den här punkten är oklar. Ibland kallas det som här kallas prospektiv studie för kohortstudie och det som här kallas retrospektiv studie för case-control studie.

Fördelarna med en prospektiv studie är egentligen nackdelarna med en retrospektiv. Tre sådana skall nämnas här. Vid retrospektiva studier kan kausalitetens riktning vara svår att avgöra. Om exempelvis personer som slutat motionera är sjukare än de var medan de motionerade kan man argumentera för en kausalitet i båda riktningarna.

Med tolkningsbara risktal menas tal, som exempelvis säger hur stor andel i en viss grupp som får en viss sjukdom under en bestämd period. Sådana tal erhålles inte vid retrospektiva studier, även om man där kan erhålla förhållandet mellan risktalen i den grupp som utsättes för en preventiv åtgärd och i kontrollgruppen.

En stor och omvitnad svårighet vid retrospektiva studier består i att erhålla en jämförbar kontrollgrupp. Denna svårighet är sannolikt det vikt-

Med en prospektiv studie menas en studie där en preventionsgrupp och en kontrollgrupp följes framåt i tiden och jämföres.

En stor och omvittnad svårighet vid retrospektiva studier består i att erhålla en jämförbar kontrollgrupp.

gaste skälet för prospektiva studier.

Även för en retrospektiv studie skall tre skäl nämnas. De två första behöver inte kommenteras närmare. De är dels urvalsstorleken, som särskilt vid sällsynta sjukdomar blir mycket stor vid prospektiva studier, dels tidsfaktorn, där fördelen vid retrospektiva studier ligger i att exponeringstiden redan är till ända.

Det tredje skälet för en retrospektiv studie är att det är etiskt omöjligt med en prospektiv studie i situationer där den preventiva åtgärden sannolikt har stor effekt.

I vissa fall kan man undvika att hamna i situationer där man behöver beakta detta genom att redan när en preventiv åtgärd börjar användas, dvs innan det finns några egentliga belägg för att den är effektiv, genomföra en prospektiv studie.

Randomisering

Vid en prospektiv studie jämföres två eller eventuellt flera grupper av individer med avseende på om de får en viss sjukdom, ett visst symptom eller med avseende på någon annan variabel man önskar påverka. Man önskar då att de grupper man jämför skall vara helt lika, innan den preventiva åtgärden insättes. I viss mån kan man erhålla grupper som är lika genom att dela upp efter kön, ålder, socialgrupp, rökvanor etc. Men man kan naturligtvis aldrig erhålla identiskt lika grupper. Den metod som finns för att komma tillrätta med detta kallas randomisering och innebär paradoxalt nog att man låter slumpen sköta fördelningen av individerna på de olika grupperna. Det finns då statistiska metoder som kan användas för jämförelse av grupperna. De metoderna innebär i princip att man beräknar hur stora avvikelser som slumpen kan förväntas orsaka. De avvikelser som går utöver detta bör kunna hänföras till den preventiva åtgärden.

Nu är det så att det finns flera metoder att

göra en randomisering. Det finns exempelvis parvis randomiserade experiment och fullständigt randomiserade experiment och det finns blockvis randomiserade. Det är bland annat valet mellan dem som statistikern och medicinen skall diskutera när undersökningen börjar planeras.

En prospektiv randomiserad undersökning är sannolikt den idealiska studien för att utreda effekten av prevention. Men det är inte alltid möjligt att utföra en prospektiv studie och det är inte heller alltid möjligt att utföra en randomiserad studie. Det är t ex svårt att göra experiment som avser sådana saker som allmän levnadsstandard, kost och bostäder m m. Där kan man kanske komma en bit på väg genom att göra tidsserier eller genom att göra internationella jämförelser. Svårigheten ligger i att isolera olika faktors effekt från varandra.

Valet av variabel på vilken effekten skall mätas är viktigt. Det är framför allt en fråga om vad man vill åstadkomma med den preventiva åtgärden. Det man vill åstadkomma kan ofta vara någonting som är svårt att mäta, någonting som har med livskvalitet att göra. Förmodligen tvingas man här ofta till en kompromiss mellan vad man vill åstadkomma och vad man kan mäta. Till den kompromissen bör man komma i en diskussion mellan medicinare och statistiker, och man bör komma dit redan på planeringsstadiet.

Ett annat viktigt val vid planeringen av en undersökning av detta slag är valet av storlek på preventions- och kontrollgrupp. Nedan redovisas några beräkningar avseende detta. Dessa avser en situation där man slumpmässigt väljer ut en preventionsgrupp och en kontrollgrupp och studerar dem prospektivt.

När man gör en statistisk utvärdering i en sådan studie kan man dra felaktiga slutsatser på två sätt. Det första innebär att man drar slutsatsen att preventionen har effekt, utan att den har det.

En prospektiv randomiserad undersökning är sannolikt den idealiska studien för att utreda effekten av prevention.

Risken för detta fel är testets signifikansnivå. Det andra felet innebär att man drar slutsatsen att preventionen inte har effekt, trots att den har det. Risken för detta slags fel kallas helt enkelt för risken att begå fel av Typ II.

Risken för fel av Typ I beror förstås bl a på hur stor effekten av preventionen är.

Jag har räknat på en del olika alternativ avseende signifikansnivå och risk för fel av Typ II.

Den variabel som effekten är mätt på är andelen i preventionsgruppen resp kontrollgruppen som insjuknar i en viss sjukdom under en viss tidsrymd. Nollhypotesen är att andelen som insjuknar inte påverkas och mothypotesen är att denna andel blir mindre när den preventiva åtgärden insättes. Hur stor andel som insjuknar och hur mycket den andelen sänks enligt mothypotesen påverkar också den erforderliga urvalsstorleken. Även härvidlag har jag räknat på några olika alternativ. Se tabell 1 och tabell 2.

I tabell 1 redovisas den erforderliga urvalsstorleken i var och en av preventionsgrupp och behandlingsgrupp i fem olika situationer, där andelen som insjuknar utan preventiv åtgärd är 20 %,

1 % resp 0,1 %. I tabell 2 specificeras de fem situationerna, med avseende på signifikansnivå, risk för fel av Typ II och andel av insjuknandena som förhindras genom preventionen.

Som synes i tabellerna krävs i en del fall mycket stora, för att inte säga orimligt stora, urval. Särskilt när andelen som insjuknar blir liten erfordras stora grupper. Här kan man vinna en del genom att registrera insjuknandet under en längre period, eftersom andelen sjuka då ökar.

Man kan diskutera vilken signifikansnivå och vilken risk för fel av Typ II man skall välja. Signifikansnivån talar om risken att dra slutsatsen att en åtgärd har preventiv effekt utan att den har det. Den bör säkert vara låg. Risken för fel av Typ II talar om risken att missa en åtgärd som har preventiv effekt. Det är möjligt att det är ett mindre allvarligt fel. Det som händer i praktiken om man gör en undersökning med för små grupper är förstås att den risken blir stor. Det påverkar alltså inte slutsatsen de gånger man får signifikant resultat men däremot de gånger man inte får det.

Preventiva åtgärder där screening ingår som ett moment

I detta avsnitt skall några aspekter på screening-situationen diskuteras. Screening innebär ju att man vid enklare undersökningar identifierar individer med riskfaktorer eller förstadiet till en viss sjukdom. Nu ligger det mer eller mindre i screeningsteknikens natur att man missar en del individer, liksom att man väljer ut en del individer som inte skulle vara med, dvs man får en del falskt positiva och en del falskt negativa.

Screeningsituationen kan enkelt beskrivas som i figur 1. Rektangeln illustrerar den population som screenas. Jag menar att det är den som är målpopulationen för preventionen och inte risk-

Tab. 1. Erforderlig urvalsstorlek under olika förutsättningar.

p*	Urvalsstorlek				
	(1)	(2)	(3)	(4)	(5)
0,200	710	160	60	260	270
0,010	17 100	3 700	1 400	6 000	6 400
0,001	172 100	36 900	13 700	60 200	64 500

Tab. 2. Specifikation av förutsättningarna i tabell 1.

	(1)	(2)	(3)	(4)	(5)
Signifikansnivå	5	5	5	1	5
Risk för fel av Typ II	20	20	20	20	5
r**	25	50	75	50	50

* Andel som löper risk att insjukna utan prevention

** = Andel av insjuknade som förhindras med prevention

Screening innebär att man vid enklare undersökningar identifierar individer med riskfaktorer eller förstadiet till en viss sjukdom.

Falskt negativa
får ej behandling

Falskt positiva
får behandling i onödan

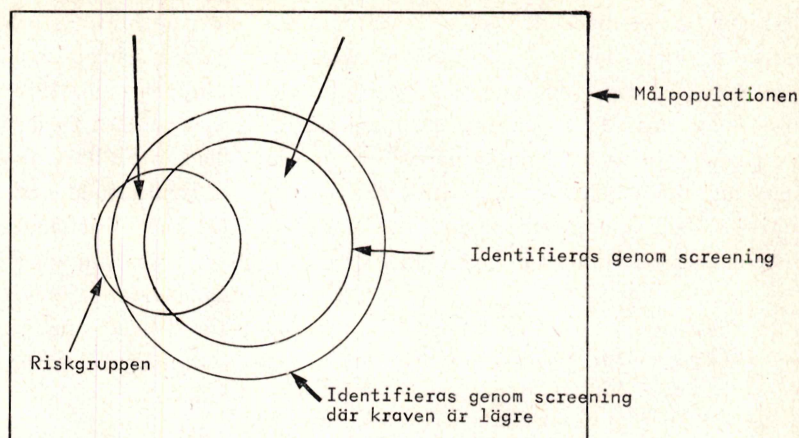


Fig. 1. Modell för screening.

gruppen eller de som identifieras vid screeningen. Den vänstra ringen betecknar de individer, som utan prevention skulle komma att insjukna. Man skulle kunna föra resonemang liknande de som skall föras här om man i stället lät den ringen illustrera den grupp som uppvisar vissa riskfaktorer. Den högra ringen betecknar de individer, som identifieras genom screeningen. De falskt negativa utgör den grupp som kommer att insjukna och av vilka en del skulle kunnat förebyggas. De falskt positiva utgör den grupp som behandlas eller utreds vidare i onödan.

Bilden illustrerar också vad som sker när kraven vid screeningen sänks. Då kommer antalet som identifieras vid screeningen att öka. Men det kommer inte bara att innebära att antalet falskt negativa minskar utan även att antalet falskt positiva ökar.

De här förhållandena har förstås betydelse när man skall diskutera effekten av sådan prevention, som innehåller screening som ett moment. Två sådana skall diskuteras här. Om man vet att en preventiv åtgärd har en viss effekt på individer som kommer att insjukna, eller som löper stor risk att insjukna, så kommer den effekten att delvis bli en annan när den skall tillämpas i kombination med screeningförfarande. Den preventiva åtgärden skall ju vidtas på de personer som identifierades genom screeningen, och bland dessa finns

en del falskt positiva. På dessa har preventionen ingen effekt. Däremot finns de falskt negativa på vilken preventionen till en del skulle haft effekt.

I tabell 3 visas hur stor andel av insjuknandena i hela målpopulationen som kan förhindras, för olika värden på andelen insjuknanden i riskgruppen som identifieras. Om man t ex kan förhindra 75 % av insjuknandena i en riskgrupp så kommer man att förhindra 60 % av insjuknandena i hela målpopulationen om man fångar upp 80 % av riskgruppen vid screeningförfarandet.

Tab. 3. Andelen som förhindras av antalet insjuknade i hela populationen.

P*	Andel, som förhindras, av antalet insjuknade i riskgruppen				
	10	25	50	75	90
99	9,9	24,75	49,50	74,25	89,10
95	9,5	23,75	47,50	71,25	85,50
90	9,0	22,50	45,00	67,50	81,00
80	8,0	20,00	40,00	60,00	72,00

* = Andel av de, som löper risk att insjukna, som identifieras vid screeningen.

Den andra effekten som skall tas upp berör de falskt positiva. Den gruppen innehåller de individer som fångats upp vid screeningen, men som inte löper risk att insjukna. De individerna ut-

sätts ibland för psykiskt och kanske också fysiskt obehag under kommande utredningar, som kanske dessutom är kostnadskrävande, om än inte för individen ifråga. Jag har gjort kalkyler över hur stor andel av dem som fångas upp vid screeningen som tillhör denna kategori. Den andelen beror på hur stor andel av riskgruppen som fångas upp, hur stor andelen av de utanför riskgruppen som inte fångas upp och på hur stor andel av målpopulationen som riskgruppen utgör.

Tab. 4. Andelen av de som identifieras vid screeningen som inte löper risk att insjukna (falskt positiva).

Andelen i hela pop som löper risk att insjukna	Andel av de som löper risk, som identifieras (sensitivitet) = andel som ej löper risk, som ej identifieras (specificitet)			
	99	95	90	80
20	4	17	31	50
10	8	32	50	69
1	50	84	92	96
0,1	91	98	99	99

I tabell 4 redovisas kalkyler för olika värden på dessa andelar. Om t ex andelen i riskgruppen som identifieras och andelen utanför riskgruppen som ej fångas upp båda är 90 % och om riskgruppen utgör 10 % av målpopulationen så kommer andelen av de som identifieras vid screening som inte tillhör riskgruppen att vara 50 %. Vid sällsynta sjukdomar blir den andelen mycket stor. Det här förhållandet kan få förödande konsekvenser i epidemiologiska sammanhang. Betydelsen i samband med prevention beror på vad det betyder för individen och samhället att behandla eller utreda vidare sådana som inte löper risk att insjukna.

LITTERATUR:

- McMahon B, Pugh, TF: Epidemiology — Principle and Methods. Boston, Little, Brown and Company. 1970.
- Armitage, P: Statistical Methods in Medical Research. Blackwell scientific publications. 1971.