

Biostatistik har en central roll i epidemiologi

Introduktion till begrepp och metoder utifrån kliniska observationella studier

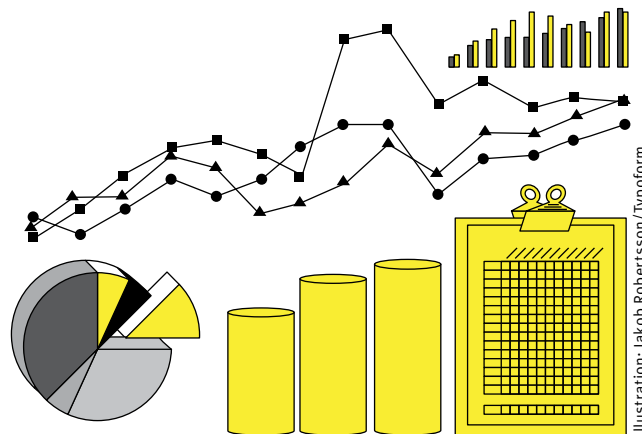
ANDREAS PETTERSSON, med dr, leg läkare, enheten för klinisk epidemiologi, Karolinska institutet, Stockholm; Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA
andreas.h.pettersson@ki.se
MAX GORDON, doktorand, specialistläkare i ortopedi, institutionen för kliniska vetenskaper, Danderyds sjukhus, Karolinska institutet, Stockholm

GUSTAF EDGREN, docent, AT-läkare, Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA
PAUL W DICKMAN, PhD, universitetslektor i biostatistik; de båda sistnämnda institutionen för medicinsk epidemiologi och biostatistik, Karolinska institutet, Stockholm

Biostatistik är en integrerad och central del av epidemiologin, och biostatistiska metoder används vid såväl planering som analys av epidemiologiska studier.

Innan en studie genomförs används sk styrkeberäkningar (power-beräkningar) för att uppskatta hur stor studien måste vara för att kunna upptäcka ett förväntat resultat, t ex en viss skillnad i sjukdomsrisk hos individer med (exponerade) och utan (oexponerade) en viss egenskap eller levnadsvana. När studiematerialet väl är insamlat används andra metoder för att beskriva materialet och beräkna relativa risker och andra riskmått samt för att uppskatta inverkan av slumpmässiga fel i studien. Ytterligare metoder finns för att undersöka om ett samband mellan en exponering och ett utfall i själva verket förklaras av en tredje faktor – en sk förväxlingsfaktor (confounder) (Fakta 1). Därutöver finns fler mer avancerade biostatistiska metoder.

Artikeln ger en introduktion till grundläggande biostatistiska metoder och begrepp som flitigt används i epidemiolo-



»Verktøy for klinisk forskning« är en artikelserie som omfattar 14 artiklar om grundläggande principer för hur man planerar och genomför kliniska forskningsstudier. Serien startade i nr 3/2013.

giska studier. Texten fokuserar på observationella kliniska studier, men flera av koncepten är relevanta även för randomiserade kliniska studier.

Deskriptiv statistik

Deskriptiv statistik används för att sammanställa studiematerial med hjälp av tabeller, figurer och olika former av sammanfattande mått, såsom medelvärden och procenttal. En koncis och informativ sammanställning kan ge god överblick över ett stort studiematerial, peka på trender i materialet och ge ledtrådar om hur man bör gå vidare med mer avancerade analyser.

Vilka former av sammanfattande mått som används i den deskriptiva sammanställningen beror på vilka typer av variabler som ingår i materialet. Det finns två huvudgrupper av variabler: kategoriska variabler (som etnicitet eller kön) och numeriska variabler (som ålder eller kroppsvikt). Kategoriska variabler beskrivs oftast med andelen eller antalet individer i respektive kategori. För numeriska variabler beräknas i stället ofta ett centralmått och ett spridningsmått; centralmättet beskriver variabelns »tyngdpunkt«, medan spridningsmättet

FAKTA 1. Förväxling

Förväxling (confounding) innebär att sambandet mellan en exponering och ett utfall i själva verket helt eller delvis förklaras av att exponeringen samvarierar med en annan exponering (förväxlingsfaktor) som direkt eller indirekt orsakar utfallet man är intresserad av.

Till exempel skulle ett samband mellan rökning (exponering) och koloncancer (utfall) helt eller delvis kunna förklaras av (vara »förväxlat« med) att rökare äter mer rött kött än icke-rökare och att intag av rött kött (förväxlingsfaktor) i sin tur orsakar koloncancer.

En förväxlingsfaktor har tre

egenskaper, den

- samvarierar med exponeringen
- orsakar utfallet direkt eller indirekt
- är inte en del av orsakskedjan mellan exponeringen och utfallet.

En fördel med randomiserade kliniska studier jämfört med observationella kliniska studier är att randomiseringen i sig minimerar risken för förväxling, eftersom deltagarna slumpmässigt randomiseras till att exponeras eller inte. Exponeringen samvarierar därmed inte med någon förväxlingsfaktor, utan är slumpmässigt fördelad.

SAMMANFATTAT

Biostatistik är en central del av epidemiologisk forskning.

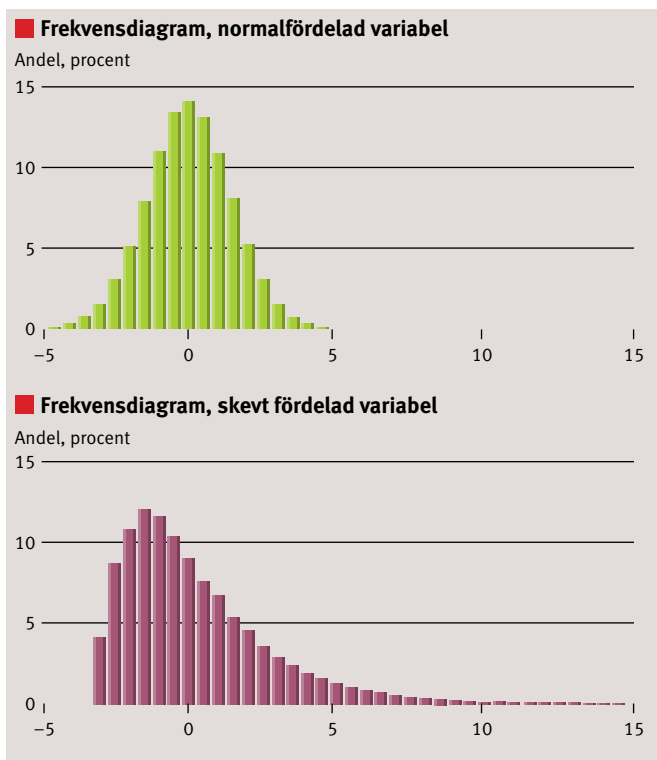
Deskriptiv statistik används för att sammanställa studiematerial med hjälp av tabeller, figurer och olika sammanfattande mått, såsom medelvärden och procenttal.

Regressionsanalys används ofta för att beräkna absoluta och relativa risker; riskmåtten speglar styrkan av sambandet mellan en exponering och förekomsten av ett visst utfall.

Inverkan av slumpmässiga fel i en studie kan kvantifieras med hjälp av P-värden och konfidensintervall.

Vid statistisk slutledning tolkas huvudresultatet (riskmättet) i ljuset av inverkan av slumpmässiga fel i studien.

Statistik slutledning är bara en del av tolkningen av ett studieresultat.



Figur 1. Frekvensdiagram av en normalfördelad (övre bilden) och en skevt fördelad variabel (undre bilden). Övre bilden: Normalfördelade variabler är symmetriskt fördelade kring sitt medelvärde. Det betyder att medelvärdet och medianvärdet (det värde som delar in samtliga mätvärden i två lika stora delar) är detsamma. Normalfördelade variabler kan till fullo beskrivas med hjälp av medelvärdet och standardavvikelsen (standardavvikelsen anger hur mycket varje mätvärde i medeltal avviker från medelvärdet). I det här exemplet är medelvärdet 0, medianvärdet 0 och standardavvikelsen 1,4. Undre bilden: Skevt fördelade variabler är asymmetriskt fördelade kring sitt medelvärde. Det betyder att medelvärdet och medianvärdet inte är detsamma. En skevt fördelad variabel kan inte till fullo beskrivas med hjälp av medelvärdet och standardavvikelsen. I det här exemplet är medelvärdet 0, medianvärdet -0,6 och standardavvikelsen 2,5.

beskriver mätvärdenas spridning kring centralmättet. Variabelns fördelning avgör i de flesta fall vilket centralmått och spridningsmått som används.

Om variabeln är normalfördelad, redovisas i typfallet medelvärdet med dess standardavvikelse. Om variabeln i stället är skevt fördelad eller om det finns extremvärden som påverkar medelvärdet, är medianvärdet tillsammans med interkvartilavståndet ofta att föredra. Centralmått och spridningsmått kan med fördel redovisas grafiskt, exempelvis med hjälp av frekvensdiagram (Figur 1).

En koncis och informativ sammanställning av ett studiematerial är värdefull både för den som analyserar materialet och för den som endast tar del av slutresultatet, t ex läsaren av en vetenskaplig artikel. En god sammanställning kan ge en bild av styrkor och svagheter i materialet. I många fall är det också möjligt att redan vid en anblick av den deskriptiva sammanställningen få en uppfattning om de samband som forskaren använt mer avancerade metoder för att karakterisera.

Riskmått

Målet med många epidemiologiska studier är att bestämma sambandet mellan en exponering och förekomsten av ett visst utfall, exempelvis sambandet mellan högt blodtryck och risken för hjärtinfarkt. För detta används olika riskmått, där de två övergripande är absolut risk och relativ risk. Om risken för

utfallet är 20 procent hos exponerade och 10 procent hos oexponerade är den absoluta riskskillnaden 10 procentenheter och den relativa risken 2,0. I analogi kan sägas att risken är 10 procentenheter högre – eller dubbelt så hög – bland exponerade. Storleken på riskmättet beskriver följaktligen sambandets styrka. Är riskskillnaden 0 procentenheter eller den relativa risken 1,0, finns inget samband mellan exponeringen och utfallet.

I epidemiologiska studier beräknas oftast relativa risker. Det blir dock allt vanligare att komplettera med beräkningar av absoluta risker, eftersom de ofta bättre uttrycker exponeringens faktiska betydelse för såväl klinikern som patienten. De vanligaste riskmåten i epidemiologiska studier beskrivs mer i detalj i andra texter i den här artikelserien.

Statistisk slutledning (inferens)

Statistisk slutledning innefattar olika teorier och metoder som används för att dra slutsatser om en hel population baserat på data från endast ett urval, ett stickprov. I en observationell klinisk studie går den statistiska slutledningen i typfallet ut på att

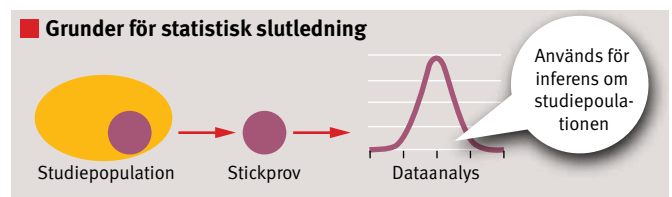
- skatta sambandet mellan exponeringen och utfallet genom att beräkna ett eller flera riskmått
- skatta inverkan av slumpmässiga fel i studien genom att beräkna konfidensintervall och P-värden
- tolka huvudresultatet (riskmättet) i ljuset av inverkan av slumpmässiga fel i studien (konfidensintervall och P-värden).

Vi illustrerar nedan några grundläggande koncept för statistisk slutledning (se även Figur 2).

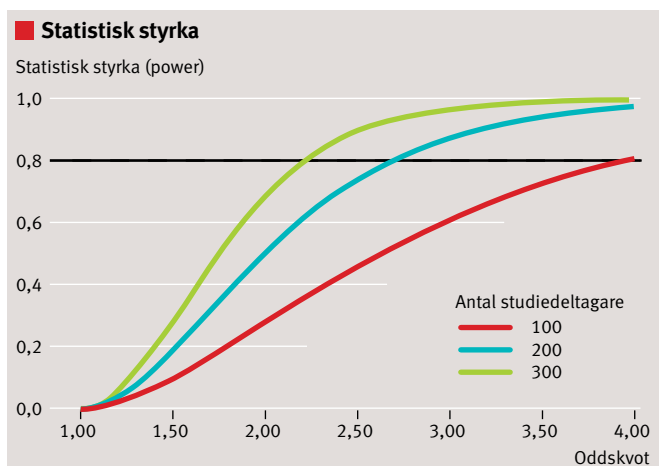
Antag att vi vill skatta förekomsten av hypertoni bland alla medelålders svenskar baserat på ett slumpmässigt urval av 500 medelålders svenskar. Andelen hypertoniker bland alla medelålders svenskar är en sk populationsparameter. Det är just populationsparametern vi är intresserade av att skatta med hjälp av data från vårt stickprov. Det enskilda värde från stickprovet som bäst skattar populationsparametern kallas punktskattning. Om vi hypotetiskt skulle ta om stickprovet, skulle andelen individer med hypertoni variera slumpmässigt från stickprov till stickprov. Andelen skulle kanske vara 23 procent i ett första stickprov, 27 procent i ett andra stickprov, 25 procent i ett tredje osv.

Punktskattningen från vårt stickprov kan därför ses som ett av många tänkbara värden av en slumpvariabel. Denna slumpvariabel följer i sin tur en sannolikhetsfördelning, som kallas samplingfördelningen. Det är i första hand egenskaper hos samplingfördelningen som används i den statistiska analysen, t ex för att beräkna konfidensintervall och P-värden.

För riskmått gäller i praktiken motsvarande resonemang som ovan. Om vi exempelvis studerar sambandet mellan behandling med lågdos acetylsalicylsyra och risken för koloncancer, kan den absoluta eller relativa risken i studien (stickprovet) ses som en punktskattning av det »sanna« sambandet



Figur 2. Grunderna för statistisk slutledning. Från en tänkt studiepopulation (t ex Sveriges befolkning) dras ett stickprov. Bland individerna i stickprovet mäts exempelvis andelen individer med hypertoni, och med olika statistiska metoder skattas sedan andelen i hela studiepopulationen.



Figur 3. Den statistiska styrkan i en studie påverkas dels av den förväntade effektstorleken, dels av antalet studiedeltagare. I detta fingerade exempel har vi antagit en signifikansnivå på 0,05 och en exponeringsprevalens på 30 procent. Med 300 deltagare skulle studien ha 80 procents styrka (power) att finna en oddsquot på drygt 2,0.

(populationsparametern) mellan acetylsalicylsyrabehandling och koloncancer.

Nedan diskuterar vi i generella termer begreppen P-värde, statistisk signifikans och konfidensintervall.

P-värde och statistisk signifikans. P-värdet anger sannolikheten för att observera (om studien upprepades under samma förhållanden) ett samband minst så starkt som det som observerats, under förutsättning att det inte finns något samband (nollhypotesen). I praktiken anger alltså P-värdet sannolikheten för att det samband man observerat, eller ett ännu starkare samband, uppkommit av slumpen. I en studie där man exempelvis funnit att den relativa risken att dö i prostatacancer är 2,0 hos överviktiga jämfört med normalviktiga män med prostatacancer, anger P-värdet sannolikheten för att observera en relativ risk på 2,0 eller mer givet att övervikt i själva verket inte påverkar risken för död i prostatacancer – dvs att den relativa risken i själva verket är 1,0.

Konceptet statistisk signifikans utvecklades i början av 1900-talet och kan beskrivas som att en gräns (signifikansnivå) dras där man anser att inverkan av slumpmässiga fel är så liten att resultatet inte rimligen kan förklaras av slumpen. Inom medicinsk forskning sätts signifikansnivån traditionellt till 5 procent. Det betyder att om sannolikheten, P-värdet, är 0,05 (motsvarar 5 procent) eller lägre anses resultatet vara statistiskt signifikant.

Begreppet statistisk signifikans är omtvistat. Nivån 0,05 är godtyckligt satt och kan med fog ifrågasättas. Den garanterar inte att vi drar korrekta slutsatser utan begränsar endast risken för falskt positiva resultat till 0,05. (Formellt finns dock inget hinder för att sätta signifikansnivån till exempelvis 0,01 om man önskar minska risken för falskt positiva resultat.) Ett ytterligare problem är att begreppet inte gör skillnad på ett P-värde på 0,06 och ett på 0,79 – båda resultaten bedöms vara statistiskt icke-signifikanta. Formellt görs inte heller skillnad mellan P-värden på 0,04 och 0,00004, trots den stora skillnaden i sannolikhet. Som många kliniker vet kan dessutom ett statistiskt signifikant resultat vara betydelselöst i praktisk bemärkelse. Är studien tillräckligt stor kan även små och kanske kliniskt irrelevanta samband vara statistiskt signifikanta.

Konfidensintervall. Ett kompletterande mått på inverkan av slumpmässiga fel i en studie är konfidensintervallet. Det

kan definieras på följande sätt: ett 95-procentigt konfidensintervall innebär att om studien (stickprovet) görs om 100 gånger i samma population, skulle det »sanna« värdet (populationsparametern) rymmas inom intervallet i 95 fall av 100. För att konstruera ett konfidensintervall måste konfidensgraden bestämmas. Traditionellt används konfidensgraden 95 procent, vilket motsvarar signifikansnivån 5 procent. Utöver konfidensgraden bestäms vidden av konfidensintervallet i första hand av studiens storlek och mätvärdets spridning. I studier där punktskattningen och konfidensintervallet motsvarar förekomsten av ett visst utfall, som ofta är fallet i exempelvis kohortstudier, är dock inte studiens storlek avgörande utan snarare hur många som drabbats av utfallet.

Vidden av konfidensintervallet speglar studiens precision; ju snävare intervall, desto högre precision och vice versa. Förutsätt två studier där båda finner en relativ risk på 1,5 – men där konfidensintervallet i den första studien är 1,4–1,6, och i den andra 0,3–4,6. Den första studien har högre precision, och vi är mer säkra på att det »sanna« värdet ligger nära 1,5. I den andra studien, där konfidensintervallet inkluderar 1,0, föreligger inte något statistiskt signifikant samband. Om det finns en relativ risk på 1,5 med konfidensintervallet 0,3–4,6 (eller något annat statistiskt icke-signifikant samband), är det viktigt att begrunda om studiens statistiska styrka var för låg för att kunna påvisa ett statistiskt signifikant samband.

Statistisk slutledning bara en del av tolkningen. Pondera att vi genomfört en studie och funnit ett starkt samband mellan en viss exponering och ett utfall. Förenklat skulle sambandet kunna förklaras av

- systematiska fel
- slumpen
- ett »sant« samband.

Ett lågt P-värde säger oss att sambandet förmodligen inte förklaras av slumpen, men det säger ingenting om huruvida det kan förklaras av systematiska fel. För att ha någon egentlig nytta av statistisk slutledning måste alltså förekomsten av systematiska fel redan ha minimerats genom god studiedesign och kontroll av förväxlingsfaktorer. Stor vikt läggs ofta vid just den statistiska slutledningen när ett studieresultat tolkas – sannolikt för att det är enklare att kvantifiera inverkan av slumpmässiga än systematiska fel i en studie.

Statistisk styrka (power)

Statistisk styrka kan definieras som sannolikheten att en studie ska kunna påvisa ett förväntat resultat. Traditionellt anses 80 procent vara en rimlig statistisk styrka. Säg att vi planerar en fall-kontrollstudie för att undersöka sambandet mellan rökning och lungcancer. Baserat på tidigare studier förväntar vi oss att rökare har omkring 10 gånger högre risk att drabbas av lungcancer än icke-rökare. Vi kan då göra en styrkeberäkning (förutsatt att vi gör antaganden om hur vanligt förekommande rökning är) och uppskatta hur många fall och kontroller vi behöver i studien för att med 80 procents sannolikhet kunna påvisa ett statistiskt signifikant samband mellan rökning och lungcancer.

Om vi i stället var intresserade av att påvisa ett statistiskt signifikant samband mellan rökning och risken att drabbas av koloncancer skulle vi behöva inkludera fler fall och kontroller. Det beror på att det förväntade sambandet mellan rökning

»Begreppet statistisk signifikans är omtvistat. Nivån 0,05 är godtyckligt satt och kan med fog ifrågasättas.«

»Regressionsanalyser används för att studera sambandet mellan olika variabler. De är centrala inom epidemiologisk forskning...«

och koloncancer är betydligt svagare än sambandet mellan rökning och lungcancer.

Något förenklat bestäms den statistiska styrkan av tre faktorer:

- studiens storlek (primärt är det antalet utfall i studien som bestämmer den statistiska styrkan)
- sambandets styrka (effektstorlek)
- signifikansnivå (traditionellt 5 procent).

Ibland vet man inte vilken effektstorlek man kan förvänta sig att observera i en studie, t ex om det saknas tidigare studier som undersökt styrkan av det aktuella sambandet. För att genomföra en styrkeberäkning kan man då välja en effektstorlek som skulle vara kliniskt relevant att påvisa. Ett exempel på hur den statistiska styrkan påverkas av studiens storlek och effektstorleken visas i Figur 3.

Att genomföra en underdimensionerad studie med otillräcklig statistisk styrka är slöseri med resurser och kan leda till den felaktiga slutsatsen att det inte finns ett »sant« samband mellan en exponering och ett utfall.

En statistisk styrka på 80 procent innebär att sannolikheten för ett falskt negativt resultat är 20 procent. En signifikansnivå på 5 procent innebär att sannolikheten för ett falskt positivt resultat är 5 procent. När gängse nivåer för statistisk signifikans och statistisk styrka används är sannolikheten för ett falskt negativt resultat därmed högre än sannolikheten för ett falskt positivt resultat (20 procent jämfört med 5 procent). Om följderna av ett falskt positivt resultat är större än följderna av ett falskt negativt studieresultat, är detta rimligt.

Ibland är dock följderna av ett falskt negativt resultat lika stora som eller större än följderna av ett falskt positivt resultat, och då kan detta system vara ett problem. Samtidigt finns det inget som hindrar att man räknar med 90, 95 eller 99 procent i stället för 80 procent statistisk styrka om man vill minska sannolikheten för falskt negativa resultat, om målet exempelvis är att verkligen utesluta att det föreligger ett relevant samband mellan en viss exponering och ett utfall.

Regressionsanalys

Regressionsanalyser används för att studera sambandet mellan olika variabler. De är centrala inom epidemiologisk forskning, eftersom multipel regressionsanalys kan användas för att kontrollera för förväxling och för att skapa prediktionsmodeller för t ex sjukdomsprognos.

Vid regressionsanalys används studiedata för att bestämma en ekvation eller matematisk modell som beskriver hur en utfallsvariabel varierar som en funktion av en eller flera förklarande variabler. Utifrån modellen kan sedan sambanden mellan variablerna kvantifieras. I t ex en kohortstudie kan man använda regressionsanalys för att beräkna den relativa risken för ett visst utfall hos exponerade jämfört med oexponerade studiedeltagare, justerat för en eller flera förväxlingsfaktorer. Nedan beskrivs schematiskt, med linjär regression som exempel, några av de grundläggande koncepten som används för att bestämma en regressionsmodell.

Vid enkel (univariabel) linjär regression beskrivs sambandet mellan en utfallsvariabel (y) och en förklarande variabel (x) utifrån räta linjens ekvation: $y = a + bx$ där a och b är konstanter; a anger var linjen korsar y-axeln (interceptet) och b anger lutningen på linjen (lutningskoefficienten). Lutnings-

koefficienten anger hur stor förändringen av y blir när x ökar eller minskar med 1 enhet, och den beräknas alltså med hjälp av studiedata. I sin enklaste form kan linjär regression användas för att beskriva hur exempelvis blodtrycket (utfallsvariabel: y) förändras (b) med åldern (förklarande variabel: x).

Flera förklarande variabler kan adderas till ekvationen, varvid modellen benämns multipel (multivariabel) linjär regression: $y = a + bx_1 + cx_2 + dx_3 \dots$. Exemplet ovan skulle kunna utökas med variablerna ålder (x_1), kön (x_2) och vikt (x_3). Konstanten b anger i det här fallet hur mycket blodtrycket varierar med åldern när kön och vikt hålls oförändrade, konstanten c anger hur mycket blodtrycket varierar mellan män och kvinnor när ålder och vikt hålls oförändrade osv. Ålder, kön och vikt är därmed ömsesidigt justerade för varandra.

I många epidemiologiska studier är utfallsvariabeln binär (t ex sjuk eller frisk) och inte kontinuerlig som i exemplet ovan (blodtryck). Om utfallsvariabeln är binär, kan sambandet mellan utfallsvariabeln och en eller flera förklarande variabler inte beskrivas väl utifrån räta linjens ekvation. Det finns flera olika former av regressionsanalys som kan användas i dessa lägen. Ofta bygger dessa analysmetoder på de grundläggande koncept vi beskrivit för linjär regression ovan, även om det oftast krävs mer avancerade matematiska formler och antaganden. I exempelvis fall-kontrollstudier används ofta en form av regressionsanalys, sk logistisk regression. Med logistisk regression beräknas oddskvoten (en form av relativ risk) för utfallet hos exponerade jämfört med oexponerade individer.

Kontroll av förväxling. För att undersöka om det finns ett oberoende samband mellan en exponering och ett utfall är det ofta viktigt att ha information om potentiella förväxlingsfaktorer. Om vi exempelvis studerar sambandet mellan rökning och risken för esofaguscancer, är det värdefullt att ha tillgång till information om vikt, alkoholintag och andra faktorer som vi misstänker helt eller delvis kan förklara ett samband mellan rökning och esofaguscancer.

Förväxling kan kontrolleras på flera olika sätt. Att göra en stratifierad analys, där deltagare delas in i grupper beroende på om de exponerats för förväxlingsfaktorn eller inte, är ett möjligt tillvägagångssätt. I exemplet ovan skulle vi t ex kunna undersöka om det finns ett samband mellan rökning och esofaguscancer separat hos normalviktiga respektive överviktiga män.

Vid en stratifierad analys finns också möjligheten att undersöka förekomsten av effektmodifiering (Fakta 2). I det här exemplet skulle vi kunna undersöka om effekten av rökning på risken att drabbas av esofaguscancer är olika hos normalviktiga och överviktiga män, dvs om övervikt modifierar effekten av rökning på risken att drabbas av esofaguscancer.

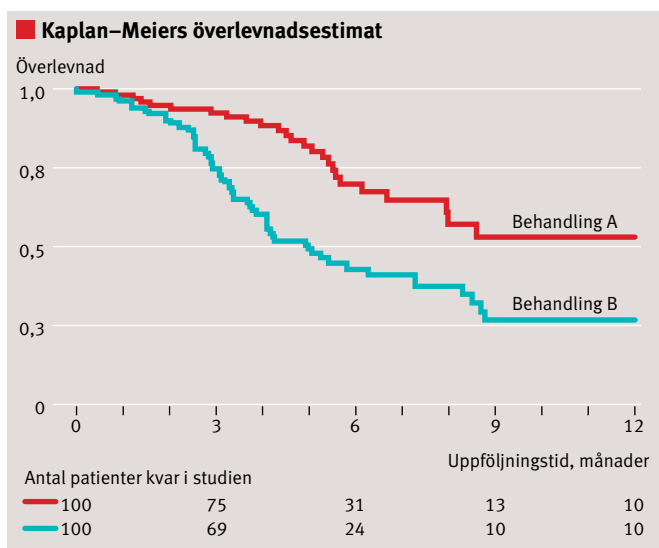
Ofta vill man ta hänsyn till flera förväxlingsfaktorer samtidigt, vilket är betydligt enklare med en multipel regressionsmodell än med stratifierad analys. En regressionsmodell tillåter också kontroll för kontinuerliga variabler. Nackdelen med

FAKTA 2. Effektmodifiering

Effektmodifiering innebär att sambandet mellan en exponering och ett utfall varierar beroende på värdet av en annan faktor (effektmodifieraren).

Till exempel är kolorektala tumörer med mutationer i genen KRAS ofta resistent

mot anti-EGFR-behandling (EGFR = epidermal growth factor receptor). Effekten av anti-EGFR-behandling (exponering) på utfallet (behandlingssvar) modifieras alltså av KRAS-mutationsstatus (effektmodifierare).



Figur 4. Kaplan-Meiers överlevnadsestimat i en fingerad studie med 200 patienter, där två behandlingar jämförs. Överlevnaden utvecklas över tid i de två behandlingsgrupperna. Det framgår tydligt (såväl av tabellen i nedkant som av hur kurvorna successivt blir »hoppigare«) att antalet patienter som är kvar i studien minskar över tid. Minskningen i det här fallet beror både på att patienterna avlider och på censurering. Genom att iakttä hur lång tid det tar innan överlevnaden understiger 0,5 kan man även grafiskt beräkna medianöverlevnaden. I detta exempel är detta endast möjligt för behandling B, där medianöverlevnaden är 5 månader.

regressionsanalys är att den är mer komplex och kräver mer kunskap för att genomföras på ett korrekt sätt.

När man kontrollerar för förväxling är målet att bygga en regressionsmodell som innehåller alla tänkbara relevanta förväxlingsfaktorer; i exemplet ovan är det de faktorer som påverkar sambandet mellan just rökning och esofagus cancer. Vissa av de ursprungliga potentiella förväxlingsfaktorerna kan vid analysen visa sig öka risken för esofagus cancer men inte påverka sambandet mellan rökning och esofagus cancer. Sådana faktorer är riskfaktorer för esofagus cancer, men inte relevanta förväxlingsfaktorer för sambandet mellan just rökning och esofagus cancer. Att identifiera alla relevanta förväxlingsfaktorer är en iterativ och ibland tidskrävande process.

I vetenskapliga artiklar redovisas ofta resultaten från både enkla och multipla analyser. Det är värdefullt, eftersom det ger läsaren möjlighet att själv se förväxlingsfaktorernas inverkan på sambandet som studerats.

Prediktionsmodeller. Regressionsanalys kan användas för att konstruera prediktionsmodeller. För kliniska syften är det vanligt att skapa prediktionsmodeller för att förutsäga prognos eller risken för att en patient har en viss sjukdom eller kommer att drabbas av den.

Samma regressionsanalyser används för prediktion och för kontroll av förväxling. Men hur regressionsmodellerna konstrueras skiljer sig åt. När prediktionsmodeller konstrueras är målet att inkludera de variabler i modellen som bäst predicerar utfallet, utan att ta hänsyn till om dessa representerar orsaks-samband eller kan förklaras av förväxling. Många vanligt förekommande kliniska behandlingsalgoritmer, såsom CURB-65, är utvecklade med hjälp av prediktionsmodeller [2].

Överlevnadsanalys

Vid överlevnadsstudier, exempelvis en studie av riskfaktorer för död hos bröstcancerpatienter, är det inte bara andelen patienter som dör som är av intresse. I den statistiska analysen måste hänsyn tas även till tiden fram till utfallet och till de

bortfall som uppstår när uppföljningen av en patient avslutas utan att hon eller han drabbats av utfallet. De patienter där uppföljningen avslutats i förtid, av andra orsaker än utfallet ifråga, benämns censurerade.

I överlevnadsstudier måste man därmed i praktiken beakta två utfall: dels huruvida patienten censurerats eller drabbats av utfallet, dels hur lång tid patienten följts. En vanlig metod för att studera patientöverlevnad är Kaplan-Meier-metoden. Den används för att konstruera överlevnadskurvor, där andelen exponerade respektive oexponerade patienter som fortfarande följs för utfallet beräknas vid varje given tidpunkt (Figur 4). Eftersom Kaplan-Meier-metoden i sin grundform inte tillåter justering för förväxling används ofta regressionsmodeller som tar hänsyn till både utfallet och tiden fram till utfallet. Exempel på sådana modeller är Cox- och Poissonregression [1].

Biostatistik är en central del av klinisk forskning

Det är vår förhoppning att läsare utan erfarenhet av att genomföra epidemiologiska studier ska ha fått en idé om några av de biostatistiska metoder och begrepp som används inom detta fält och intresse av att fördjupa sina kunskaper inom biostatistik. Vi hoppas också att det står klart att biostatistiken är en integrerad och central del av epidemiologin och den kliniska forskningen. Ett välfungerande samarbete mellan kliniker, epidemiologer och biostatistiker har en stor potential att kunna angripa och besvara kliniskt relevanta frågeställningar.

För en gedigen genomgång av de metoder och begrepp som diskuteras i artikeln rekommenderas »Fundamentals of biostatistics« [1] och »Modern epidemiology« [3]. Teorierna och metoderna som används för att beräkna punktskattning, konfidensintervall och P-värden (sk punkt- och intervallskattning respektive statistisk hypotesprövning) finns utförligt beskrivna i »Fundamentals of biostatistics« [1].

■ *Potentiella bindningar eller jävsförhållanden: Inga uppgivna.*

REFERENSER

1. Rosner B. Fundamentals of biostatistics. 7th ed. Boston: Brooks/Cole, Cengage Learning; 2011.
2. Lim WS, van der Eerden MM, Laing R, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax*. 2003;58:377-82.
3. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. 3rd ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008.

■ SERIE VERKTYG FÖR KLINISK FORSKNING

Med artikelserien »Verktyg för klinisk forskning« vill Läkartidningen ge inspiration till goda forskningsidéer, tips och råd om studieupplägg och hur man söker anslag, gör analyser och skriver en vetenskaplig artikel. Serien består av 14 artiklar.

Hittills publicerade artiklar:

»Verktyg för klinisk forskning« – ny artikelserie i Läkartidningen
Karin Ekström Smedby, Ylva Trolle Lagerros Nr 3/2013 (1/14)

Observationella kohortstudier Översikt och några exempel från området läkemedels säkerhet
Marie Holmqvist, Pauline Raaschou, Martin Neovius, Johan Askling Nr 3/2013 (2/14)

Att skatta behandlingseffekt
Jonatan Lind Nr 4/2013 (3/14)

Etiska aspekter att tänka på i observationsstudier
Stefan Eriksson, Erik Sundström Nr 5/2013 (4/14)

Biobankerna utvecklas för att möta forskningens krav

Sonja Eaker, Anna Beskow, Loreana Norlin Nr 5/2013 (5/14)

Bra prognosstudier kan ge bättre kliniska beslut. Validerade risk-/prognosfaktorer hjälper läkaren
Katja Fall, Lars Holmberg, Johan Sundström Nr 6/2013 (6/14)

Att studera ovanliga utfall
Johan Reutfors, Ola Olén, Cecilia Magnusson, Andreas Pettersson Nr 7/2013 (7/14)

Att strukturera och dokumentera forskningsprojekt
Sandra Eloranta, Anna LV Johansson, Sigurdur Y Kristinsson, Therese M-L Andersson Nr 8/2013 (8/14)