ARBETSMILJÖ FONDENS SAMMANFATTNINGAR

1256

# Metoder och system för datorstödd yrkesmedicinsk epidemiologi

*För innehållet i denna sammanfattning svarar Olav Axelson, Yrkesmedicin, Regionsjukhuset, 581 85 Linköping, tel 013-19 41 14.*

*Pnr 78-2150 Lokalt skyddsarbete och företagshälsovård (70)* *Februari 1989*

## Bakgrund

Genom epidemiologiska undersökningar har en rad faktorer i arbetslivet kunnat identifieras som mer eller mindre troliga orsaker till sjukdom. Epidemiologiska projekt är synnerligen arbetskrävande, men modern datorteknik har underlättat vissa faser i sådana undersökningar, närmast vad gäller själva hanteringen av data och främst då i s k kohortstudier. Kohortstudien innebär i korthet att en grupp exponerade personer följs över lång tid och att dödligheten eller sjukdomsfrekvensen observeras i olika avseenden och jämförs antingen med en annan, likartad men oexponerad grupp, eller med förväntningsvärden för dödligheten eller sjukligheten såsom framräknade från officiell dödsstatistik eller cancerstatistik. Detta innebär således att andra icke-dödliga sjukdomsyttringar än cancer kräver jämförelsekohort för att slutsatser skall kunna dras. Också andelen dödsfall i en viss sjukdom av alla dödsfall kan jämföras hos exponerade och oexponerade så att en uppfattning kan erhållas beträffande exponeringens inverkan på dödlighetsmönstret. Undersökningar av detta slag brukar betecknas som proportionella mortalitetsstudier.

Statistiska beräkningar ingår som en ofrånkomlig del i epidemiologiska undersökningar. Dessa beräkningar är oftast måttligt resurskrävande och kan i viss utsträckning klaras redan med hjälp av kvalificerat programmerbara räknedosor.

Flera dataprogram för kohort-studier och proportionella mortalitetsundersökningar har introducerats på olika håll i världen under de senaste åren. Dessa program är dock oftast relativt begränsade och kan inte anses utgöra några mer generella lösningar på datorstöd för arbetsmedicinsk epidemiologi. Det aktuella projektet har syftat till att ta fram ett mer omfattande paket av datorprogram som stöd för epidemiologiska

undersökningar inom företrädesvis yrkes- och miljömedicin, men dessa program torde också mycket väl kunna utnyttjas för epidemiologiska undersökningar inom andra områden.

## Principiella krav och lösningar

Populationer kan med hänsyn till sin omsättning av personer betecknas som antingen öppna eller slutna. Öppna populationer, t ex en kommuns befolkning, har en omsättning av personer och populationen kommer att karakteriseras av dessa individers status under en viss period, t ex mantalskrivningen i en viss kommun. Den slutna populationen saknar omsättning och utgörs av individer vilka kan sägas ha blivit berörda av någon speciell händelse. Detta senare är typiskt för en yrkesmedicinsk kohort och innebär att ett antal personer kännetecknas av händelsen av att ha utsatts för viss industriell exponering, t ex karakteriserad av viss duration och/eller intensitet.

Undersökningar i öppna populationer aktualiseras emellertid också inom arbetsmedicinen, oftast i form av fall-kontroll-studier. I en sådan studie identifieras ett antal individer med en viss sjukdom och undersöks beträffande sin tidigare industriella exponering i något speciellt avseende, varefter en jämförelse görs med en grupp kontroller (även kallad referenter) såsom dragna ur och representativa för den befolkning ur vilken fallen härrör. Eventuell skillnad i exponering mellan fall och kontroller i berört avseende vittnar om exponeringens inverkan. I den mån man arbetar med döda fall och kontroller erhålles en studiekaraktär som nära ansluter till den proportionella mortalitetsstudien men som metodologiskt oftast är att föredra framför den senare. Fall-kontroll-studier kan ibland förläggas inom en kohort, vilket inte sällan aktualiseras i yrkesmedicinska sammanhang, t ex när man vill göra mer detaljerade utvärderingar beträffande samverkan mellan olika riskfaktorer.

Konsekvensen av de studietekniker och befolkningsstrukturer som här skisserats, blir att ett epidemiologiskt programpaket bör tillåta hanteringen av såväl öppna som slutna befolkningar och fungera antingen man har tillgång till kompletta data eller får nöja sig med ett urval av kontroller såsom i fall-kontroll-studier. För de fullständigt kända populationerna, där den yrkesmedicinska kohorten är typexempel, finns möjligheter att som sjukledighetsmått antingen studera den kumulativa incidensraten, dvs antalet under en viss tid ackumulerade fall per ursprungligt antal individer i studien, alternativt antalet fall per ackumulerade personår under observation eller incidensraten. I det senare fallet bidrar varje individ under sin framlevnad med personår under observation i olika åldersgrupper alltefter åldrandet och under de olika kalenderår som studien omspänner. Utvärderingar av sjukligheten mot bakgrund av personår under observation kan vara att föredra såsom medförande en större känslighet, eftersom hänsyn därvid kommer att tas till vad som händer i kohorten i termer av god eller dålig överlevnad.

Ett flexibelt epidemiologiskt programpaket bör också väl tillgodose möjligheterna att enkelt kunna splittra upp ett material i olika undergrupper, t ex subkohorter eller exponeringskategorier. Vidare måste hänsyn kunna tas till multipla exponeringar och villkoren lätt ändras med hänsyn till krav på latenstid, dvs den tid som kan vara rimlig att låta löpa mellan påbörjad exponering och den tidpunkt då en individ tas under observation (dvs tillåts bidra med ackumulerade personår) för risk att insjukna eller avlida som följd av exponeringen.

Ett epidemiologiskt programpaket bör, förutom att tillhandahålla direkt applicerbara epidemiologiska program, också ge användaren möjlighet att bearbeta data med åtminstone något av de mer omfattande statistikprogram som finns kommersiellt tillgängliga samt även medge en enkel överföring av uppgifter från exempelvis frågeformulär, personlistor etc till datormedium. Denna process kan ibland vara ganska arbetskrävande och därmed kostsam. Senare tids utveckling på mikrodatorområdet gör det även intressant att åtminstone i viss utsträckning tillhandahålla epidemiologisk

progamvara för mikrodator främst vad avser epidemiologisk-statistik analys av data.

## Resultat i projektet

Mot bakgrund av de krav som ovan i korthet skisserats har ett epidemiologiskt programpaket utvecklats för stordator och givits beteckningen EPILIN med programdelarna FORMA, RATEMAN, COHORT och FALKON. Dessutom finns framtaget EPILIN-MICRO i två mikrodatorversioner, en avsedd för ABC 800 och en för IBM-kompatibla datorer.

FORMA är ett datainmatningsprogram för EPILIN-systemet. Via den gemensamma databasen överföres uppgifterna till beräkningsprogrammen COHORT respektive FALKON avsedda för cohort- respepektive fall-kontrollanalyser. Data till beräkningsprogrammen kan också införas via vanliga textfiler. RATEMAN-programmet används för uppbyggnad av incidensrater baserat på existerande mortalitetsstatistik vid Statistiska Centralbyrån respektive Cancerregistret och förser således COHORT-programmet med erforderliga bakgrundsdata.

FORMA-programmet medger att ett frågeformulär motsvarande det som praktiskt använts genereras på skärmen. Formulärsvaren kan sedan direkt ifyllas och data omhändertas därmed för den vidare bearbetningen.

RATEMAN-programmet medger uppbyggnad av sjukdomsspecifika eller sjukdomsgruppsaggregerade rater från basdata såsom dessa föreligger i den offentliga statistiken. ICD-kod (International Classification of Disease) används och vad gäller förändringar mellan olika upplagor av ICD-kodningen medger programmet visst hänsynstagande i dessa ofta besvärliga övergångar från en klassificering till en annnan.

COHORT-programmet beräknar antalet personår under observation inom en godtyckligt definierad studieperiod och fördelar de observerade personåren på kön, åldersgrupp och kalenderår, varefter multiplikation sker med de ålders-, köns-, och kalenderårsspecifika raterna. Därefter summeras de fraktionella bidragen i varje speciell kategori till ett förväntningsvärde för kohorten under den aktuella tidsperioden och för den speciella sjukdomen eller gruppen av sjukdomar. Kvoten mellan observerat antal fall och detta förväntningsvärde beräknas därefter (SMR). P-värden för skillnaden mellan observerat och förväntat antal fall liksom konfidensintervallet räknas sedan fram med förutsättningen att Poissonfördelningen gäller.

FALKON-programmet är i huvudsak ett statistiskt programpaket av standardkaraktär, men har fördelen att man direkt från EPILIN-programmets databas kan hämta fram relevanta data, göra bearbetningar med hjälp av SPSSX-paketet och därefter direkt företa de olika analyser som är aktuella beträffande fall-kontroll-studier, främst vad avser testning med Mantel-Haenzel-testet eller varianter härav.

EPILIN-MICRO omfattar epidemiologiska statistikprogram som tidigare funnits tillgängliga för programmerbara räknedosor. Den fördel som vunnits med användningen av mikrodator i sammanhanget hänför sig främst till ökad överskådlighet vid analysen av data samt till bättre utskriftsrutiner. Programvaran baserar sig eljest i huvudsak på en sammanställning av Rothman och Boice. (Rothman KJ, Boice JD. *Epidemiologic analysis with a programmable calculator.* NIH Publication No 79-1649, Washington.)

## Diskussion

Det nu framtagna programpaketet med beteckningen EPILIN har under sin utvecklingstid använts i ett betydande antal studier, uppskattningsvis ett 30-tal, utförda företrädesvis i Stockholm (arbetarskyddsstyrelsen/arbetsmiljöinstitutet), Örebro (yrkesmedicinska kliniken) och Linköping (yrkesmedicinska kliniken och barnkliniken). Datorprogram tenderar emellertid att ha egenskapen att inte bli helt färdiga, då förbättringar alltid är möjliga och EPILIN-systemet utgör inget undantag i detta hänseende. Programutvecklingen har emellertid i princip avslutats och det finns nu dels en längre uppsats (på engelska) som presenterar systemet, dels en huvudmanual (förutom att det till de olika delprogrammen

finns manualer från tidigare utvecklingsstadier och dessa program kan fortfarande också användas om så önskas). Erfarenheterna har hittills varit väsentligen goda av programutvecklingen, men av och till har problem påträffats som krävt vissa modifieringar. Fortsättningsvis är det därför önskvärt att användare rapporterar problem och önskemål så att ytterligare justeringar kan vidtas om så är erforderligt.

Systemet är höggradigt interaktivt innebärande att utförliga anvisningar om tillvägagångssätt vid användningen av programvaran kan erhållas på operatörens skärm med möjligheter till speciell hjälpinstruktion. Den rutinmässiga datahanteringen bör därför i stor utsträckning kunna genomföras av personal utan speciella datortekniska kvalifikationer. Det kan dock vara väsentligt att påpeka behovet av epidemiologisk kunskap för den som avser att med hjälp av EPILIN genomföra ett epidemiologiskt projekt. I själva verket kanske man kan påstå att faran för missvisande studier har ökat såtillvida att många av de tekniska svårigheter som eljest tenderar att begränsa den epidemiologiska verksamheten i och med programpaketet är övervunna. I princip gäller således att någon form av resultat alltid kommer ut ur systemet oavsett om primärt inmatade data hänför sig till en adekvat upplagd studie eller inte.

Dessa senare synpunkter kanske gäller i särskilt hög grad för EPILIN-MICRO som blir lättillgängligt genom att en diskett och erforderliga instruktioner kan erhållas av den som så önskar för egenhändigt utnyttjande i en mikrodator. Tills vidare är EPILIN-paketet endast tillgängligt i stordator vid Linköping Datacentral (LIDAC), men modifiering och överföring till andra datorer är i och för sig möjlig. Vissa modifieringar har sålunda efterhand vidtagits för att underlätta överföringen till VAX-dator.

## Rapportering

EPILIN-programmet är tillgängligt i form av manual med en introducerande uppsats om systemets uppbyggnad (på engelska, kostnad 300:- kr). EPILIN-paketet kan sedan utnyttjas via fasta förbindelser, eventuellt genom andra datacentraler i landet, eller via telefonförbindelser med LIDAC. Taxan är härvid tänkt att bli den som gäller för utnyttjande av datorservice vid LIDAC + 50 % programspecifik ersättning som är avsedd att täcka kostnaderna för programunderhållet.

EPILIN-MICRO tillhandahålles i form av en instruktionspärm (på svenska) med en 5 1/4" diskett. Versioner finns för användning på såväl ABC 800 som IBM-kompatibla mikrodatorer och kan köpas för en kostnad av 1 500:- kr per styck.

Beställning kan göras hos Yrkesmedicinska kliniken, Regionsjukhuset, 581 85 Linköping alt 701 85 Örebro, tel 013-19 14 41 eller 019-15 24 71.

Arbetsmiljöfonden

Box 1122, 111 81 Stockholm
Tel 08-796 47 00 (vx)

ISSN 0183-6718 Norstedts Tryckeri, Stockholm 1989