



Folkhälsomyndigheten

Metodbeskrivning

Bilaga 3 till Folkhälsans utveckling – Årsrapport 2020



Innehåll

Introduktion.....	3
Datakällor	4
Definitioner	5
Indikatorer.....	5
Redovisningsgrupper	5
Deskriptiv statistik	7
Deskriptiv statistik efter datakälla.....	7
Statistiska analyser	10
Statistiska analyser efter frågeställning.....	11
Statistiska analyser efter datakälla	12
Referenser	13

Introduktion

De indikatorer som ingår i rapporteringen har valts för att spegla relevanta aspekter av folkhälsan. En annan praktisk utgångspunkt har varit att välja indikatorer där det finns tillgång till data, företrädesvis data med kvalitetsdeklaration. Indikatorerna ska kunna redovisas återkommande och under lång tid framåt, enligt rekommendation i underlagsrapport 9 till Kommissionen för jämlik hälsa [1]. De data som använts i årsrapporten kommer dels från register vid exempelvis Statistiska centralbyrån och Socialstyrelsen, dels från urvalsundersökningar i form av enkätstudier.

Indikatorerna redovisas på hela befolkningen och i möjligaste mån uppdelat på redovisningsvariablerna kön, åldersgrupp, utbildningsnivå, födelseland i fyra grupper, samt region i form av län. Indikatorerna redovisas deskriptivt, och har då datatillgången tillåter även analyserats med statistiska metoder.

Mer information om indikatorernas koppling till folkhälsa och grafik som åskådliggör data finns i det webbaserade uppföljningssystemet [Folkhälsans utveckling – siffror, grafik och text](#) [2].

Datakällor

De data som använts i rapporteringen kommer dels från register, och dels från urvalsundersökningar i form av enkätstudier. Huvudsakliga datakällor för de indikatorer som följs är Patientregistret och Dödsorsaksregistret som handhas av Socialstyrelsen (SoS), Registret över totalbefolkningen (RTB), Arbetskraftsundersökningarna (AKU) och Inkomst- och taxeringsregistret (IoT) som handhas av Statistiska centralbyrån (SCB), Skolverkets elevregister och de nationella enkätstudierna ”Hälsa på lika villkor” (HLV) och ”Skolbarns hälsovanor” som handhas av Folkhälsomyndigheten. I HLV ingår vissa bakgrundsdata från register.

I årsrapporten Folkhälsans utveckling 2020 har även data från SmiNet och Naturvårdsverket använts.

Olika datakällor har olika typer och mängd av innehåll, och därmed olika förutsättningar för redovisning av resultat. Indikatorerna redovisas på hela befolkningen och i möjligaste mån uppdelat på ett antal redovisningsgrupper. Huvudsakliga datakällor för den information som används för redovisningsgrupperna är RTB och utbildningsregistret (UREG).

Register är sammanställningar av befolkningen vilket innebär stor precision. Däremot är till exempel patientregistret beroende av att läkare runt om i Sverige rapporterar likartat. Olika register har också olika mycket uppgifter som saknas, till exempel sakas uppgift om utbildning i UREG i högre utsträckning för de äldsta och personer födda utanför Sverige, än för andra personer [3]. Vissa register uppdateras också med korrigeringar, vilket kan innebära små avvikelser i data som rapporteras här från den officiella statistiken. För mer information kring kvaliteten på olika register se respektive registerhållare.

Resultat från urvalsundersökningar är skattningar. Observera att skattningarna innebär en viss osäkerhet, bland annat på grund av bortfall. För att få en uppfattning om osäkerheten, används i vissa fall 95 % konfidensintervall.

Nationella folkhälsoenkäten är en enkät som skickas ut till ett urval av befolkningen. Under 2006 till 2016 har undersökningen genomförts årligen och urvalet har bestått av cirka 20 000 slumpmässigt utvalda personer i åldern 16–84 år. Sedan 2016 skickas enkäten ut vartannat år. Enkäten skickades till 40 000 personer 2018. Svarefrekvensen för Nationella folkhälsoenkäten har sedan 2006 minskat från 60 procent till 42 procent 2018, vilket i viss mån kan påverka resultaten. Kalibreringar utförs för att kompensera för bortfallet [4]. Ett antal regioner har också gjort ett tilläggsurval med extra antal enkäter, som bland annat redovisas på kommunnivå i det webbaserade uppföljningssystemet [Folkhälsans utveckling – siffror, grafik och text](#) [2].

Definitioner

Rapporteringen baseras på ett antal definitioner för indikatorer och redovisningsgrupper. Här följer en kortfattad beskrivning av hur redovisningsgrupperna är definierade samt en hänvisning till den bilaga som beskriver definitioner för de indikatorer som ingår i rapporten.

Indikatorer

Indikatorerna som redovisas är utvalda för att spegla viktiga aspekter av folkhälsan och bakomliggande förutsättningar i form av livsvillkor och levnadsvanor. De utgör eller bidrar till en ansevärd sjukdomsbyrå, till ojämlikhet i hälsa, samt kommer från en tillförlitlig datakälla, företrädesvis med kvalitetsdeklaration. I bilaga 2 redovisas de indikatorerna med namn, datakälla och definition.

Redovisningsgrupper

Indikatorerna redovisas i möjligaste mån för olika redovisningsgrupper för fördjupning av kunskapen. Redovisningsgrupper är indelningar på redovisningsvariablerna kön, åldersgrupp, utbildningsnivå, födelseort i fyra grupper, och region i form av län. Se Tabell 1 för generella redovisningsgrupper. Om data för en indikator varit begränsad så har tillgängliga redovisningsgrupper använts. Även indikatorspecifika undantag med mer detaljerad indelning inom någon redovisningsvariabel förekommer, med exempelvis deskriptiva data på åldersgrupper i 5-årsklasser.

Vid redovisning på utbildningsnivå för personer under 25 år har utbildningsnivån för den förälder som har högst utbildning används. Förälder definieras som biologisk eller adoptivförälder, där adoptivförälder givits företräde framför biologiska föräldrar.

Referensgrupper

Referensgrupp kallas den redovisningsgrupp som övriga redovisningsgrupper inom en redovisningsvariabel jämförs mot. Följande referensgrupper används kön: kvinnor, utbildningsnivå: eftergymnasial, ålder: åldersgrupp 30-44 år eller i vissa fall den yngsta möjliga åldersgruppen, födelseort: Sverige, och för region används genomsnitt för Sveriges samtliga län. Se noteringar om referensgrupp i Tabell 1.

Tabell 1. Redovisningsgrupper, generella

Redovisningsvariabel	Redovisningsgrupp	Beskrivning
Kön	Kvinnor (referensgrupp) Män	Juridiskt kön
Utbildningsnivå	Eftergymnasial (referensgrupp) Gymnasial Förgymnasial Uppgift saknas	Högst uppnådda utbildningsnivå grupperas efter SUN2000. 100-206 Förgymnasial 310-337 Gymnasial 410-640 Eftergymnasial
Åldersgrupp	15(16)-29 år 30-44 år (referensgrupp) 45-64 år 65-84 år 85+ år	Ålder vid händelse alternativt i slutet av aktuellt årtal.
Födelse land	Sverige (referensgrupp) Övriga Norden Övriga Europa Övriga världen	Sovjetunionen räknas här in i gruppen Övriga Europa. I de fall uppgift om födelse land saknas, räknas som regel födelse land till redovisningsgrupp Övriga världen.
Region, län	Stockholm Uppsala Södermanland Östergötland Jönköping Kronoberg Kalmar Gotland Blekinge Skåne Halland Västra Götaland Värmland Örebro Västmanland Dalarna Gävleborg Västernorrland Jämtland Västerbotten Norrbotten (Referensgrupp är snittet av samtliga län)	Län kommer typiskt från Registret över totalbefolkningen (RTB), SCB.

Deskriptiv statistik

Deskriptiva statistik, som beskrivs i text och visas i figurerna på webben och i tabellerna i bilaga 1, presenteras i form av andelar i procent (prevalens) eller antal per 100 000 eller per 1 000 individer (incidens). Undantag är medellivslängd, ekonomisk standard, halt och utsläpp av partiklar och antal anmälda fall av smittsamma sjukdomar.

De deskriptiva resultaten visas för olika åldrar beroende på indikator. Vad gäller utbildningsnivå visas deskriptiva resultat för åldern 25 år och äldre. Vid 25 års ålder bedöms att de allra flesta haft chans att uppnå eftergymnasial utbildningsnivå. Vid redovisning på utbildningsnivå för personer under 25 år har utbildningsnivån för den förälder som har högst utbildning används.

Bland redovisningsvariablerna har kön en särställning eftersom all statistik redovisas per kön också om siffrorna är uppdelade på någon av de övriga redovisningsvariablerna. För många indikatorer redovisas både Ej ålderstandardiserade och Åldersstandardiserade värden. Ålderstandardisering har gjorts för att justera för effekten av att åldersfördelningen kan vara olika mellan befolkningsgrupper och över tid.

Åldersstandardiseringen har utförts med metoden direkt standardisering och gjorts på 5-årsåldersklasser med medelbefolkning 2017 som standardpopulation. Undantag är bröstcancer och prostatacancer, där standardpopulation är medelbefolkningen av kvinnor respektive män. Inom deskriptiva data i innevarande rapportering används 95 år och äldre som gemensam grupp för de äldsta åldersgrupperna.

För deskriptiva data på län används medelvärden över flera år för att få tillräckligt stort dataunderlag.

Deskriptiv statistik efter datakälla

Socialstyrelsen (SoS)

Indikatorer med källa SoS är hämtade från dödsorsaksregistret och patientregistret och redovisas i form av incidens. Indikatorerna är egna bearbetningar av data från SoS.

Resultaten visas för åldrarna 15 år och äldre utom fallolyckor bland äldre, som redovisas från 65 år.

Som nämnare i incidens för ett årtal YYYY används medelbefolkning beräknad som medelvärde mellan befolkningen YYYY-12-31 och befolkningen föregående nyår [YYYY-1]-12-31. Befolkningen i nämnaren avgränsas till folkbokförda vid tillfället. En individ som föds eller immigrerar under ett visst år, alternativt dör eller emigrerar, bidrar så med 0,5 individer till medelfolkmängden.

Befolkningen i täljaren avgränsas till individer som varit folkbokförda någon gång under perioden 2005-2018. Dödsdatum respektive utskrivningsdatum används för att avgöra vilket år ett dödsfall respektive sjukdomsfall ska registreras.

Uppgifter om utbildning för år YYYY hämtas i första hand från närmast föregående nyår [YYYY-1]-12-31, och om uppgiften där saknas så hämtas uppgiften från innevarande år [YYYY]-12-31 om den finns. Detta gäller täljaren. Till nämnaren hämtas utbildning från år samma som befolkningen definieras.

Indikatorerna redovisas som åldersstandardiserade samt ej åldersstandardiserade värden. Indikatorer redovisas som medelvärden över fem år i deskriptiva data för län.

Folkhälsomyndigheten – Hälsa på lika villkor (HLV)

Indikatorer från HLV redovisas som prevalenser med andelar i procent, tillsammans med 95 % konfidensintervall för dessa skattningar. Kalibreringar utförs för att kompensera för bortfallet [4].

Resultaten visas för åldrarna 16–84 år för data från HLV. Vad gäller utbildningsnivå visas 25–84 år. Här bör noteras att information saknas gällande utbildning för dem över 74 år, 2006–2009.

Utbildningsnivå Uppgift saknas ingår inte som redovisningsgrupp i deskriptiv statistik på grund av få fall.

Indikatorerna redovisas ej som åldersstandardiserade värden. Indikatorer från HLV redovisas som fyraårsmedelvärden för län.

Statistiska centralbyrån (SCB)

Från SCB beställs data från ett antal olika register som innebär olika sätt att redovisa data. Indikatorerna ekonomiska utsatthet, gymnasiebehörighet och gymnasieexamen, långtidsarbetslöshet, sysselsättningsgrad och unga som varken arbetar eller studerar, redovisas som prevalens med andelar i procent, medan spädbarnsdödlighet redovisas som incidens. För medellivslängd visas antal år kvar att leva och ekonomiska standard visas som medianvärden och percentiler. Flera av indikatorerna är egna bearbetningar av data från SCB.

Data för indikatorerna från SCB visas för olika åldrar beroende på indikator. Spädbarnsdödlighet, unga som varken arbetar eller studerar, gymnasiebehörighet samt gymnasieexamen visas för respektive åldersgrupp av intresse, medan ekonomisk standard och ekonomisk utsatthet visas för hela populationen. Långtidsarbetslöshet (i arbetskraften) visas för åldrarna 15–74 år och sysselsättningsgrad (i befolkningen) för dem 25–64 år.

Ekonomisk standard och Ekonomisk utsatthet för år YYYY baseras på uppgifter för befolkningen YYYY-12-31

Vad gäller gymnasiebehörighet och gymnasieexamen utgörs nämnaren av elever i årskurs 9 respektive elever som påbörjat gymnasiet. Nämnare för spädbarnsdödlighet är alla levande födda barn ett visst år. Långtidsarbetslöshet, sysselsättningsgrad och unga som varken arbetar eller studerar är tagna från AKU som är en urvalsundersökning. För information om AKU se Kvalitetsdeklaration från SCB [5].

I de av SCB beräknade värdena för Medellivslängd uppdelat på utbildningsnivå är utrikesfödda exkluderade på grund av metodproblem. Föräldrarnas utbildningsnivå och moderns födelseland används som redovisningsgrupper för spädbarnsdödlighet. Även för gymnasiebehörighet och gymnasieexamen samt ekonomisk utsatthet bland barn används föräldrarnas utbildningsnivå.

Indikatorn vuxna i ekonomisk utsatthet redovisas som åldersstandardiserade samt ej åldersstandardiserade värden, övriga indikatorer redovisas ej som åldersstandardiserade värden. Spädbarnsdödlighet redovisas som medelvärden över fem år i deskriptiva data för län. Medellivslängd redovisas som femårsvärden för län respektive födelseland.

Statistiska analyser

De statistiska analyserna syftar till att försöka besvara tre frågor.

- Finns skillnad i nuläge mellan redovisningsgrupper?
- Finns trend (av ökning eller minskning) under undersökningsperioden?
- Finns skillnad i trend mellan redovisningsgrupperna?

I de statistiska analyserna har redovisningsvariablerna använts som oberoende variabler eller justeringsvariabler, och indikatorn som beroende variabel. Två sorters analyser har gjorts, dels ojusterad analys och dels justerad analys.

Vid ojusterad analys av en indikator har endast en oberoende variabel ingått i modellen. I justerade analyser har en i sänder av redovisningsvariablerna valts som oberoende variabel (exponeringsvariabel) och övriga tillgängliga redovisningsvariabler använts som justeringsvariabler (confounders). Vid justerad analys av skillnader mellan redovisningsgrupper inom en redovisningsvariabel (exempelvis utbildningsnivå) har på så sätt hänsyn tagits till övriga redovisningsvariabler. Detta för att säkerställa att de funna skillnaderna i dödlighet, sjuklighet eller levnadsvanor mellan till exempel olika utbildningsgrupper är relaterad till utbildningsnivå och inte en eller flera av de övriga redovisningsvariablerna.

I möjligaste mån har indikatorerna studerats uppdelade på de generella redovisningsgrupperna (Tabell 1) för jämförbarhet. I analyser som justerats för åldersgrupp har 5-års-åldersgrupper använts. Vid statistiska analyser med utbildningsnivå som oberoende variabel har åldrarna 25 år och uppåt använts. Vid statistiska analyser med utbildningsnivå som justeringsvariabel har även de som är yngre än 25 år tagits med i analysen. Föräldrarnas högsta utbildningsnivå har då använts om sådan uppgift funnits i data, annars personens egen utbildningsnivå.

Frågeställningarna har besvarats genom att anpassa regressionsmodeller till data. Relativ risk (RR) med Quasi-Poisson-regression, alternativt oddskvot (OR) med logistisk regression har beräknats som mått på om en redovisningsgrupp har en högre eller lägre risk avseende indikatorn jämfört med referensgruppen. RR beräknades för incidens eller prevalens för indikatorer från registerdata, och OR för prevalens från HLV. P-värden och 95 procents konfidensintervall för respektive estimat har utgjort underlag till bedömning av om eventuella nulägeskillnader, trender eller trendskillnader var signifikanta. Indikatorspecifika bedömningar kan förekomma vid p-värden 0,01–0,05.

I regressionsanalyserna har ingen hänsyn tagits till multipla tester. Observationer i data med partiellt bortfall i den beroende variabeln eller någon av de oberoende variablerna uteslöts från analyserna.

Samtliga analysberäkningar har utförts i statistikprogramvaran SAS.

Statistiska analyser efter frågeställning

Nulägesanalys av skillnaden mellan två grupper

Syftet med nulägesanalysen har varit att ta reda på om det finns skillnad mellan redovisningsgrupper. Detta har mätts som relativ skillnad mellan respektive redovisningsgrupp och referensgrupp vid senaste mättillfället. För utbildningsnivå till exempel, där Eftergymnasial utbildningsnivå är referensgrupp, studeras om en eventuell skillnad är signifikant mellan gruppen med Förgymnasial utbildningsnivå jämfört med referensgrupp, respektive mellan gruppen med Gymnasial utbildningsnivå jämfört med referensgrupp. Frågeställningen har besvarats genom att anpassa en regressionsmodell till data.

För nulägesanalys på län har flerårsmedelvärden använts för att samla tillräckligt stort dataunderlag.

Trendanalys av utvecklingen inom respektive grupp

Syftet med trendanalysen har varit att ta reda på om det finns förändring över tid under undersökningsperioden. Detta har mätts som signifikant linjär trend av relativ ökning eller minskning över undersökningsperioden, beräknad som relativ förändring i snitt mellan två på varandra följande år.

Vid det som benämnts som oförändrad trend konstateras att varken en signifikant ökande eller minskande trend har skett (5 procent signifikansnivå). I vissa fall har trendanalys inte genomförts på grund av att data inte medgivit beräkning.

Undersökningsperioden är 2006–2019 för de två indikatorerna kring gymnasium, 2006–2018 för övriga om det varit möjligt, annars den period som är möjlig inom de åren, beroende på datatillgång för indikatorn.

Frågeställningen har besvarats genom att anpassa en regressionsmodell med uppföljningsperioden i årtal som en kontinuerlig exponeringsvariabel. Baserat på denna modell har den totala minskningen eller ökningen i incidensproportion alternativt prevalens över undersökningsperioden uppskattats.

Trendanalys av skillnaden i utveckling mellan två grupper

Syftet med analysen har varit att studera om det finns en skillnad i förändring över tid för redovisningsgrupp i relation till referensgruppen. Detta har mätts som signifikant linjär relativ skillnad mellan trend hos redovisningsgrupp och trend hos referensgrupp över undersökningsperioden.

Frågeställningen har besvarats genom att anpassa en regressionsmodell med interaktionsterm mellan uppföljningsperioden i årtal som en kontinuerlig variabel och en av de oberoende variablerna i taget. Interaktionstermens koefficient har använts som mått på utveckling för redovisningsgruppen över tidsperioden i relation till referensgruppen.

Statistiska analyser efter datakälla

Socialstyrelsen (SoS)

Indikatorer från patientregistret och dödsorsaksregistret har analyserats med hjälp av Quasi-Poisson-regression. Relativ risk (RR) har beräknats för befolkningstotaler och respektive redovisningsgrupp.

För indikatorer om cancerformer har yngre åldersgrupper uteslutits från analyserna om det varit få fall i grupperna. Referensgrupp för ålder har typiskt varit den generella referensgruppen 30–44. Där det inte varit möjligt har den yngsta åldersgrupp där dataunderlaget medgett statistiska analyser valts istället.

Folkhälsomyndigheten – Hälsa på lika villkor (HLV)

För indikatorerna från HLV har statistiska analyserna på oddskvoter OR med logistisk regression utförts. I modellen har kalibreringsvikt ingått, se tekniska rapporter [4].

De statistiska analyserna är gjorda på åldrarna 16–84 år för alla redovisningsgrupper utom vad gäller utbildningsnivå. Information om utbildningsnivå för personer 75 år och äldre saknas i HLV år 2009 och tidigare. Statistiska analyser med utbildningsnivå som redovisningsgrupp är därför gjorda för dem 25–74 år där den övre gränsen är vald såsom konstant över hela tidsserien. För justeringar i analyser för utbildningsnivå för personer yngre än 25 år används personens egen utbildningsnivå.

Regressionsanalyser för indikatorer med HLV som datakälla har gjorts på ett dataset med nationella data där ytterligare tilläggsurval för län Gotland ingår. De statistiska analyserna som baseras på HLV-data kan därför skilja sig i enstaka fall i förhållande till nationella data som redovisas i statistikdatabasen Folkhälsodata.

Statistiska centralbyrån (SCB)

Analys av RR med Quasi-Poisson-regression har utförts på de indikatorer – ekonomisk utsatthet, gymnasiebehörighet och gymnasieexamen samt spädbarnsdödlighet – som analyserats statistiskt från SCB. RR har beräknats för befolkningstotaler och respektive redovisningsgrupp.

För Spädbarnsdödlighet används tre olika uppföljningstider (de första 0-6 dygnet, 0-27 dygnet, första levnadsåret). Föräldrarnas utbildningsnivå och moderns födelseort används som redovisningsgrupper för spädbarnsdödlighet. Även för gymnasiebehörighet och gymnasieexamen samt ekonomisk utsatthet bland barn används föräldrarnas utbildningsnivå.

Referenser

1. Vågerö D. Löpande bevakning och rapportering av folkhälsa och ojämlikhet i hälsa. Underlagsrapport 9 till Kommissionen för jämlik hälsa. Stockholm: Kommissionen för jämlik hälsa; 2017.
2. Folkhälsomyndigheten. Folkhälsans utveckling - siffror, grafik och text. Solna: Folkhälsomyndigheten; 2019. [uppdaterad 2019-03-14; citerad 2020-03-23] Hämtad från: <https://www.folkhalsomyndigheten.se/folkhalsorapportering-statistik/folkhalsans-utveckling/>
3. Statistiska centralbyrån. Bortfallsanalys av UREG 2014. Örebro: Statistiska centralbyrån, 2017. [citerad 2020-03-23] Hämtad från: https://www.scb.se/contentassets/b1fd5a0fe6c74df7b103f434f6215678/uf0506_2016a01_br_a40br1705-1.pdf
4. Folkhälsomyndigheten. Tekniska rapporter. Solna: Folkhälsomyndigheten; 2018. [citerad 2020-03-23] Hämtad från: <https://www.folkhalsomyndigheten.se/contentassets/9b1b216c596a487ca6c6aa6dc413efb4/teknisk-rapport-med-bilagor-hlv-2018.pdf>
5. Statistiska centralbyrån (SCB). Kvalitetsdeklaration. Arbetskraftsundersökningarna (AKU). Stockholm: SCB, 2019. [citerad 4 november 2019]. Hämtad från: https://www.scb.se/contentassets/c12fd0d28d604529b2b4ffc2eb742fbe/am0401_kd_2019_lj_190214.pdf

Folkhälsomyndigheten är en nationell kunskapsmyndighet som arbetar för en bättre folkhälsa. Det gör myndigheten genom att utveckla och stödja samhällets arbete med att främja hälsa, förebygga ohälsa och skydda mot hälsohot. Vår vision är en folkhälsa som stärker samhällets utveckling.



Folkhälsomyndigheten

Solna Nobels väg 18, 171 82 Solna. **Östersund** Forskarens väg 3. Box 505, 831 26 Östersund.

www.folkhalsomyndigheten.se