

Vårddata för forskning – hinder och möjligheter

Individbaserade vårddata är en viktig komponent i modern epidemiologisk forskning. Ett av syftena med EU:s ambitiösa initiativ European Health Data Space (EHDS) är att underlätta sekundäranvändning av vårddata för forskning [1]. Men hur lätt är det i dag för forskare i Sverige att få tillgång till individdata från vårdens it-system?

PCBase Xtend (Prostate cancer database Sweden – extended treatment and endpoint data) är en utvidgning av ett etablerat epidemiologiskt forskningsprojekt om prostatacancer [2, 3]. I PCBase har Nationella prostatacancerregistret (NPCR) länkats till andra hälsodataregister och demografiska databaser. Sedan 2009 har mer än 250 artiklar om olika aspekter av prostatacancer publicerats baserade på data i PCBase [4–8].

I PCBase Xtend har vi gått ett steg längre. Under närmare 4 år har vi samlat in individdata från sjukvårdens it-system i alla 21 regioner i Sverige, inklusive privata vårdgivare och laboratorier. Vi efterfrågade värden för prostataspecifikt antigen (PSA), PAD-svar från prostatabiopsier (fritext), resultat av magnetisk resonanstomografi (MR) av prostata (fritext) och kemoterapi som använts vid prostatacancer på sjukhus och i sjukhusanknuten öppenvård (rekvisitionsläkemedel som inte finns i Läkemedelsregistret).

Syftet med vårt projekt är att kunna studera asymtomatisk testning, utredning, diagnostik, riskkategorier, behandlingsstrategier och behandlingsresultat vid prostatacancer.

Begäran om utlämning av data

Första begäran om datauttag gjordes i november 2021 till Västra Götalandsregionen, och den sista begäran skickades i maj 2024. Genom att stegvis över tid tillfråga nya regioner och privata vårdgivare kunde vi anpassa processen efter lärdomar vi gjort i arbetet med de första regionerna. Det krävdes 46 separata ansökningar: 35 till regioner/sjukhus och 11 till privata vårdgivare. I vissa fall krävdes flera ansökningar till samma region, exempelvis i Värmland, Västmanland och Västra Götaland.

Ansökan var utformad som ett standardbrev med en noggrann redovisning av syfte, legal grund, dataspecifikationer och it-säkerhet för överföring av data. I bilagor fanns etikansökan, projektplan och etikgodkännande. Begäran riktades till den administrativa funktion för utlämning av forskningsdata som fanns i några regioner. Annars skickades begäran till FoU-enhet eller chefsläkare. I många regioner fanns det särskilda blanketter för datauttag, alla olika utformade.

Juridik och etik

Rättslig grund för personuppgiftsbehandlingen finns i artikel 6 e (uppgift av allmänt intresse) och artikel

Rolf Gedeberg, docent, institutionen för kirurgiska vetenskaper, Uppsala universitet
 ● rolf.gedeberg@uu.se

Lennart Holm, projektkoordinator, Regionalt cancercentrum Mellansverige, Uppsala

Pär Stattin, professor i urologi, institutionen för kirurgiska vetenskaper, Uppsala universitet

9 (allmänt intresse på folkhälsoområdet) i EU:s dataskyddsförordning (GDPR) för behandling av personuppgifter inom ett forskningsprojekt som godkänts av Etikprövningsmyndigheten. Enligt 6 kap 5 § i offentlighets- och sekretesslagen (OSL) ska en myndighet »på begäran av en annan myndighet lämna uppgift som den förfogar över, om inte uppgiften är sekretessbelagd eller det skulle hindra arbetets behöriga gång». Data sekretesskyddas genom att såväl statistiksekretess enligt 24 kap 8 § som hälso- och sjukvårdsekretess enligt 25 kap 1 § överförs med stöd av 11 kap 3 § OSL.

Vår begäran om utlämning ifrågasattes inte på legala eller etiska grunder, med två undantag. I dessa fall avslögs begäran med hänvisning till att det skulle »hindra arbetets behöriga gång» enligt ovanstående lagtext. Efter att respektive regions jurister tagit del av argumentationen i vårt överklagande så omprövades avslagen utan att behöva prövas av domstol och data lämnades även ut från dessa två regioner utan ytterligare svårigheter.

Extraktion av data från vårdens it-system

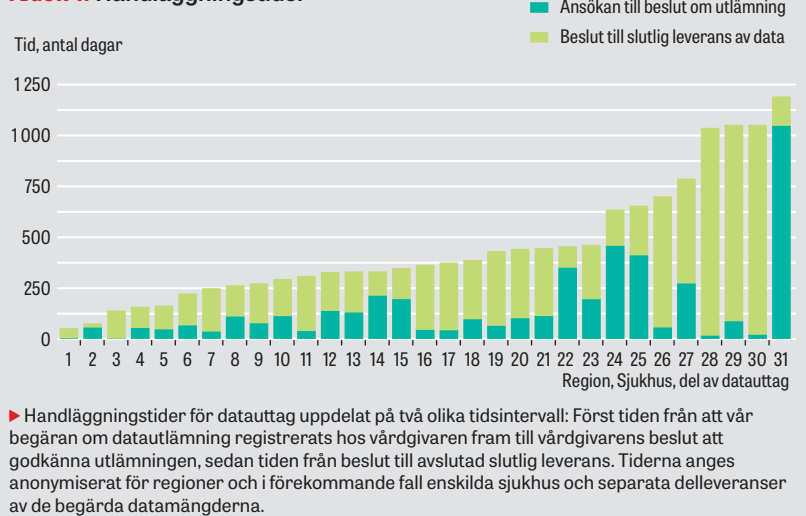
Vi efterfrågade endast information i digital form. Tillgängligheten påverkas av hur väl it-systemen är anpassade för att lämna ut individdata och hur regionen konstruerat sina datalager. Exempelvis var information om läkemedel i journalsystemet Melior betydligt mer lättåtkomlig i Västra Götalandsregionen än i Region Skåne, beroende på att man satt upp sina datalager på olika sätt. Generellt är det svårare att extrahera data från äldre it-system, men viktigare faktorer är i vilken utsträckning it-leverantörer kan stödja sina kunder med att underlätta och genomföra uttag och hur väl förberedda vårdgivarna är för att lösa uppgiften.

Data från klinisk kemi var enklast att hantera,

HUVUDBUDSKAP

- För att göra epidemiologiska studier av prostatacancer samlade vi in individdata om PSA-värden, PAD-svar från prostatabiopsier, resultat av MR prostata samt kemoterapi från alla berörda vårdgivare i Sverige.
- Det var i de flesta fall möjligt att hämta denna information från vårdens it-system, men det var olika svårt och tog olika lång tid beroende på vårdgivare, typ av information och typ av it-system.
- Tidig kontroll av kvaliteten på levererade data är nödvändig för att identifiera felaktiga datauttag.
- Vårdgivare behöver integrera forskningsuttag i sina administrativa funktioner och it-strukturer.
- Det behövs incitament så att vårdgivare migrerar äldre data vid byte av it-system.

FIGUR 1. Handläggningstider



eftersom dessa data har en enkel struktur och informationen sedan länge lagras i strukturerade it-system. Men det fanns fallgrorpar ändå. Exempel på sådana är att det ofta krävdes klargöranden av vilken specifik typ av PSA-prov vi efterfrågade i relation till varje enskilt laboratoriums beteckningar. Det var dessutom nödvändigt att få tillgång till både det uppmätta och det utsvarade värdet för att klargöra analysens detektionsgräns och kunna se hur den förändrats över tid. En annan viktig fråga är täckningen av olika vårdgivare, så att alla sjukhus, primärvård och privata aktörer kommer med i uttaget.

När det gällde PAD-svar var vi huvudsakligen intresserade av mellannålsbiopsier. Det visade sig dock vara svårt att identifiera dessa på ett enkelt sätt eftersom typ av provtagning endast var sökbar i ett fåtal system.

I många fall behövde vi anlita it-leverantörer som Sectra, Agfa, CGM och Tietoevry för att skapa uttagsfunktioner från respektive it-system. Alla leverantö-

rer utom CGM debiterade för sitt arbete, vilket även de flesta regioner gjorde. Det fanns en stor spridning i hur mycket man debiterade.

Ostrukturerad fritext som PAD-svar och resultat av MR kräver naturligtvis en hög grad av bearbetning, inte bara för att extrahera den specifika information som behövs för forskningen, utan också för att tillförlitligt avlägsna personuppgifter som ibland fanns inbäddade i fritexten.

Migrera data från gamla it-system

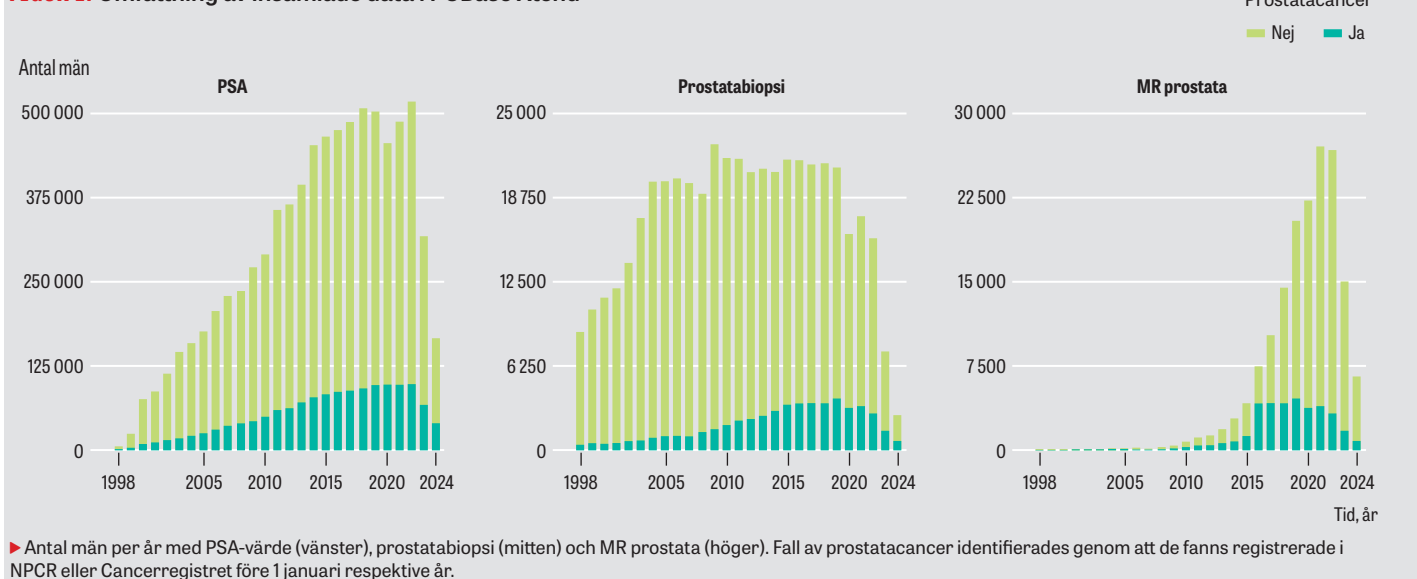
En viktig begränsning uppstår när en region byter it-system och äldre data då inte alltid migreras till det nya it-systemet. Extra komplicerat, och ibland olösligt, blir det när regionen bytt mellan ett privat laboratorieföretag och verksamhet i egen regi. I dag saknar vårdgivare incitament för att bevara data från äldre it-system eftersom historiska data om en enskild patient finns i dennes journal, medan den samlade databasen med till exempel laboratorievärden fortfarande kan vara av stort värde för forskningsändamål.

Levererade data

Efter ett intensivt arbete under närmare 4 år, med en dedikerad projektkoordinator med vänlig attityd, oändligt tålamod och stor noggrannhet och otaliga mejl, telefonsamtal och videomöten, har samtliga regioner och privata aktörer levererat huvuddelen av den begärda datamängden. Insamlingen omfattade alla män över 30 år i Sverige mellan åren 1998 och 2024, oavsett prostatacancerdiagnos – totalt över 5 miljoner män. Täckningen bakåt i tid varierade.

Regionerna hade mycket olika grad av mognad i rutinerna för datautlämning. Vissa regioner hade etablerade processer, medan andra helt saknade struktur. I några mindre/medelstora regioner tog det över 1 år från ansökan om data till beslut. Handläggningstiden från ansökan till färdig dataleverans varierade från 2 dagar till över 3 år (Figur 1). I Region Stockholm tog det 2 år och 10 månader till beslut, och ytterligare drygt 4 månader till leverans.

FIGUR 2. Omfattning av insamlade data i PCBase Xtend



Kontroll av datakvalitet

Tidig kontroll av datakvalitet efter leverans visade sig vara nödvändig. Genom att jämföra med data från Nationella prostatacancerregistret kunde vi identifiera felaktigheter i utlämnade data från flera regioner. På så vis kunde vi reklamera uttaget och få ett korrigerat uttag medan alla parter fortfarande hade ärendet aktuellt.

Lärdomar

Genom denna insamling berikade vi PCBase med 11 miljoner PSA-värden, 550 000 prostatabiopsier, 170 000 MR-undersökningar av prostata och >6 miljoner läkemedelsadministreringar (Figur 2). Resultatet är en mycket värdefull nationell datamängd för studier av vägen från ett förhöjt PSA-värde till prostatacancerdiagnos och det efterföljande sjukdomsförloppet.

Våra erfarenheter från detta projekt är att tidig dialog är viktig och är ett effektivt sätt att underlätta arbetet och minska arbetsinsatsen för alla parter. De allra flesta företrädare för regioner och vårdgivare var mycket tillmötesgående, och det fanns en genuin vilja hos de allra flesta att bistå på bästa sätt.

En kostnad på ungefär 1,5 miljoner kronor och en tidsåtgång på närmare 4 år är dock uppenbara hinder för mer utbredd användning av vårddata för forskningsändamål. Nuvarande strukturer för arkivering och utlämning av vårddata i Sverige är i flera avseenden bristfälliga. I vissa regioner finns en väl strukturerad process på plats med central bedömning av kompetenta projektledare, till exempel i Region Skåne. I andra regioner är processen fragmenterad, oömsogen och omständlig. De varierande och ibland mycket långa handläggningstiderna medförde att dataleveranser blev mycket utsträckta över tid, vilket innebar att den totala täckningsgraden sjönk under den senare delen av studieperioden (Figur 2). Det är en olycklig och onödig följd av administrativa hinder för datatillgång.

Vi vill betona att vi fått ett mycket positivt bemötande av representanter för regioner och vårdgivare. Med få undantag har bemötandet från de chefer, administratörer, sjukvårdspersonal och it-tekniker vi samverkat med i detta projekt varit engagerat och konstruktivt. Tack vare deras insats har vi i Sverige nu en internationellt unik möjlighet till epidemiologiska studier av prostatacancer.

Hur kan tillgången till vårddata bli bättre?

Att för forskningsändamål få tillgång till vårddata från regioner och privata vårdaktörer är mödosamt, kostsamt och tidskrävande. Om European Health Data Space ska få effekt för forskningen och inte bli en skrivbordsprodukt så behövs ett samlat grepp för

att underlätta tillgängliggörandet av data bland vårdgivare. Sverige har i många avseenden redan goda förutsättningar för sekundär användning av sjukvårdsdata för forskning, genom nationella hälsodataregister, kvalitetsregister och regionernas höga ambitionsnivå när det gäller digitalisering. EHDS bör förbättra dessa möjligheter ytterligare genom att vi får ett enhetligt system för ansökan, tillgång och användning av hälsodata för forskning i hela EU. Utmaningarna när det gäller att implementera EHDS är dock sannolikt betydande. Svårigheterna är bland annat kopplade till behov av rättslig harmonisering och mer kraftfull nationell styrning. Vi har en fragmenterad hälso- och sjukvårdsorganisation med varierande digital mognad, och investeringsbehoven är sannolikt betydande både på regional och nationell nivå. Vårt projekt belyser en del av de praktiska problem som finns i dagsläget, men även om implementeringen av EHDS inte är enkel så är det ändå ett steg mot bättre tillgång till sjukvårdsdata för forskning.

Vårdgivare behöver ha anpassade processer och integrera forskningsuttag i sina it-strukturer och administrativa funktioner. Chefer inom vården behöver stöd för att hantera begäranden om datauttag för forskning. Det behövs också incitament så att vårdgivare migrerar äldre data vid byte till nytt it-system, eftersom dessa data kan ha mycket stort värde för forskning. Ett första steg mot en bättre uttagsprocess är att funktionalitet för datauttag finns med när it-system upphandlas och datalager sätts upp. Däremot behöver inte data lagras centralt nationellt för effektiv och säker sekundäranvändning av vårddata i forskning [9].

Vi har i vårt projekt lärt oss den hårda vägen om uttag av vårddata för forskning. Problemet är att de erfarenheter vi gjort i en region inte kommer andra forskare och andra regioner till godo. Vi efterlyser därför en nationell struktur som stödjer regioner och forskare i juridiska, etiska, medicinska och it-tekniska frågor för att underlätta forskning på vårddata. Det finns många goda regionala initiativ, kunniga individer och lärrika exempel att bygga på. Kanske kan fortsatt utveckling av Vetenskapsrådets initiativ med den nya webbplatsen »Dataguiden« (<https://dataguiden.se>) och forum som »Kliniska studier Sverige«, ett nationellt samarbete mellan sjukvårdsregioner, öppna vägen till mer strukturerade och effektiva sätt att praktiskt tillgängliggöra vårddata för forskning. ○

Läs mer!

Engelsk
sammenfattning på
[Läkartidningen.se](http://Lakartidningen.se)

● Potentiella bindningar eller jävsförhållanden: Rolf Gedeberg är även anställd vid Läkemedelsverket. De åsikter som uttrycks i denna artikel behöver inte nödvändigtvis överensstämma med Läkemedelsverkets ståndpunkter.

Citera: *Läkartidningen*. 2026;123:25159. doi: 10.65868/AANT2294

REFERENSER

1. Europeiska kommissionen. Förordningen om det europeiska hälsodataområdet. https://health.ec.europa.eu/health-digital-health-and-care/european-health-data-space-regulation-ehds_sv
2. Westerberg M, Holm L, Garmo H, et al. Cohort profile update: the National Prostate Cancer Register of Sweden and PCBase. *Int J Epidemiol*. 2025;54(5):dyaf172.
3. Beckmann K, Garmo H, Franck Lissbrant I, et al. The value of real-world data in understanding prostate cancer risk and improving clinical care: examples from Swedish registries. *Cancers (Basel)*. 2021;13(4):875.
4. Van Hemelrijck M, Wigertz A, Sandin F, et al; NPCR and PCBase Sweden. Cohort profile: the National Prostate Cancer Register of Sweden and Prostate cancer data base Sweden 2.0. *Int J Epidemiol*. 2013;42(4):956-67.
5. Maljkovic J, Bill-Axelsson A, Hållberg H, et al. Time trends for the use of active surveillance and deferred treatment for localised prostate cancer in Sweden: a nationwide study. *Scand J Urol*. 2024;59:200-6.
6. Stattin P. How to improve cancer care by use of guidelines and quality registers. *Scand J Urol*. 2024;59:190-2.
7. Scilipoti P, Liedberg F, Garmo H, et al. Risk of prostate cancer death in men diagnosed with prostate cancer at cystoprostatectomy. A nationwide population-based study. *Scand J Urol*. 2024;59:98-103.
8. Nationella prostatacancerregistret (NPCR). Publikationer. <https://npcr.se/publikationer/>
9. Gedeberg R, Igl W, Svenblad B, et al. Federated analyses of multiple data sources in drug safety studies. *Pharmacoepidemiol Drug Saf*. 2023;32(3):279-86.

SUMMARY

Access to health data for research

Secondary use of individual-level health data for research is cost-efficient but challenging. In the PCBase Xtend project, we collected PSA values, biopsy reports, MRI results, and chemotherapy data from all 21 regions and private healthcare providers in Sweden. The process required 46 applications over a period of almost 4 years and revealed major variability in IT systems, administrative routines, and costs. The time to data delivery ranged from weeks to years. Early quality checks were essential to correct errors in the data extraction. Despite challenges, >11 million PSA values, 550 000 prostate biopsies, 170 000 MRI results, and >6 million drug administrations could be added to PCBase for use in epidemiological studies of prostate cancer. Better administrative routines, data extraction integrated in IT systems, incentives for data migration, and national support structures are needed to improve access to health data for research.