

# Att strukturera och dokumentera forskningsprojekt

**SANDRA ELORANTA, MSc,**  
biostatistiker

sandra.eloranta@ki.se

**ANNA LV JOHANSSON, MSc,**  
biostatistiker; båda vid institu-  
tionen för medicinsk epidemio-  
logi och biostatistik, Karolinska  
institutet

**SIGURDUR Y KRISTINSSON,**

PhD, leg läkare, Hematologiskt  
centrum, Karolinska universi-  
tetsjukhuset, Solna

**THERESE M-L ANDERSSON,**  
MSc, biostatistiker, institu-  
tionen för medicinsk epidemiologi  
och biostatistik, Karolinska in-  
stitutet

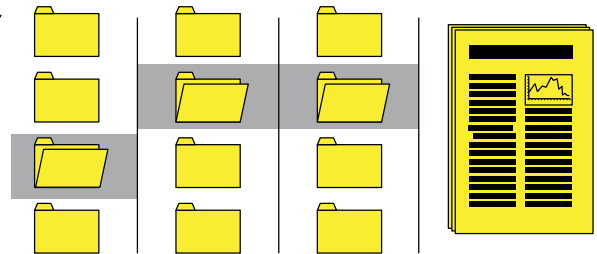


Illustration: Jakob Robertsson/Typoform

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

Att forska är för många läkare liktydigt med sena kvällar hemma med datorn då det gäller att effektivt få ihop något av anständig kvalitet. Då är inte forskningsdokumentation det första man prioriterar. Vi vill visa på några enkla knep för att åstadkomma en enklare och säkrare tillvaro för sina data och forskningsresultat, vilket i längden leder till bättre forskning och ett effektivare arbetssätt. Strukturerad dokumentation bör ses som en naturlig del i forskningsarbetet – inte något man gör i efterhand när röran redan är ett faktum.

Effektivitetsvinster och säkerhetsaspekter borde vara incitament för alla forskare att noggrant dokumentera hur de kom fram till sina forskningsresultat. Utöver detta finns lagligt krav på god hantering av känsliga personuppgifter inom forskning samt särskilda regler kring specifika typer av studiedesigner, till exempel randomiserade kliniska försök. Detta berörs emellertid inte vidare i denna artikel. God datahantering handlar om att kunna återskapa publicerade resultat – från tabellen i artikeln till uppgifter om enskilda patienter i ursprungsfilerna som användes i analysen. När referentkommentarerna med ändringsförslag anländer är det extra viktigt att revisionsarbetet går snabbt och problemfritt. Detta möjliggörs genom ordnad dokumentation. En strukturad forskningsdokumentation och kompetent datahantering går således hand i hand. God forskningsdokumentation minimerar startsträckan varje gång man tar itu med projektet på nytt.

En framgångsrik dokumentation förutsätter en struktur och en genomtänkt plan. Men när får man lära sig hur man lägger upp en god mappstruktur i sin dator eller hur man namnger sina filer? Vi ger exempel på hur man enkelt kan lägga upp sina projekt och hantera sina filer, samt beskriver användbara dokument såsom analysplanen, loggboken och kodböcker. I det långa loppet handlar lämplig datahantering och forskningsdokumentation om god projektledning. En diskussion kring dataflödet i projektet bör finnas med i alla steg i forskningsprocessen – från hypotes till arkivering (Figur 1).

## Var proaktiv – spar tid med en bra mappstruktur

Ett vanligt misstag som forskare ofta begår vid starten av ett projekt är att överskatta den egna förmågan att hantera stora mängder filer av olika slag. Det kan röra sig om rådatamaterial, bearbetade/rensade data, textdokument, programfiler, bilder, tabeller m m. Det kan visa sig svårt att hitta aktuella filversioner, något som stjälar tid varje gång arbetet återupptas.

En effektiv mappstruktur kan förebygga många problem, särskilt i stora och komplicerade projekt. Vi rekommenderar att du vid starten av varje nytt projekt skapar en logisk och lättbegriplig struktur. En lämplig regel är att skapa en huvudmapp för varje publikation. Om en och samma datamängd ska användas till flera publikationer, skapa en generell projektmapp för

det huvudsakliga projektet och därefter undermappar för varje publikation. Bra mappstruktur åskådliggör ditt forskningsprojekt genom att färre delar ger bättre överblick (Figur 2). Huvudmappen för varje publikation delar du in i ett antal undermappar för olika typer av filer.

**Data.** Hit hör rådatamaterial, rensade datafiler samt viktig information relaterad till datainsamlingen, till exempel en kopia av databeställningen eller enkäter/patientformulär som användes vid datainsamlingen, samt formulär för abstraktion av journaldata och information om vilka labbanalysers som gjordes.

**Dokument.** I den här mappen läggs manusversioner, godkänd etikansökan, forskningsplan, forskningsmedelsansökningar, korrespondens mellan forskaren och tidskriften, studieprotokoll för kliniska prövningar, information om biologiska metoder som använts samt dokumentationsfiler som analysplan och kodböcker.

**Program.** I programmappen läggs de statistikprogramfiler som genererade forskningsresultat i den statistiska analysen.

**Resultat.** Hit hör resultatfiler (output) från programfilerna, tabeller, figurer och grafer.

Man bör även skapa en så kallad README-fil, en textfil som ska ge läsaren en snabb överblick av projektet och som därför bör innehålla övergripande information om projektet, tidsperiod, namn på inblandade forskare samt en beskrivning av mappstrukturen. Den är tänkt att fungera ungefär som ett register i en pärm och bör sparas överst i projektmappen. Denna typ av dokument namnges ofta README för att snabbt dra till sig läsarens uppmärksamhet.

## God dokumentation börjar med analysplanen

Analysplanen bör du påbörja innan projektet sätter igång genom att skriva en kort sammanfattning av din frågeställning/

## ■ SAMMANFATTAT

**God forskningsdokumentation** handlar om att säkerställa att forskningsmaterial och resultat på en snabbt och enkelt sätt kan reproduceras, kontrolleras och, vid behov, återanvändas. **Positiva effekter av en** genomtänkt dokumentationsrutin ger tidsvinster samt minskar risken för felhantering av forskningsmaterial.

**Nyckelfaktorer för en lyckad** forskningsdokumentation är en

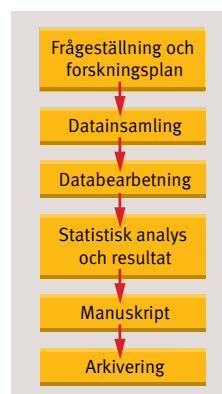
genomtänkt mappstruktur, analysplanen, logg och kodböcker. **Den rutin som rekommenderas** kan hjälpa forskaren att arbeta i enlighet med rådande svensk lagstiftning, effektivisera forskningen och författandet av vetenskapliga artiklar, samt arkiveringen av slutförda studier. **Enkla verktyg syftar till** att ge en tydlig överblick av forskningsprocessen.

hypotes, vilka variabler som är av intresse och hur dessa mäts. Analysplanen är det viktigaste dokumentet. Den knyter ihop det vetenskapliga tankegångarna med filer och resultat och ska innehålla all viktig projektinformation, exempelvis beslut som fattats, analyser som gjorts (även de som slutligen inte omnämns i artikeln), information om hur dina data samlats in, vilka variabler som använts samt kontaktuppgifter till alla medförfattare. Till skillnad från en forskningsplan, som används vid forskningsmedelsansökningar eller etikprövning, beskriver analysplanen det faktiska arbetet som utförs under projektets gång och inte bara vad du planerar göra. Fakta 1 visar vilka rubriker som bör finnas med i analysplanen. Analysplanen ger dig överblick och bör uppdateras kontinuerligt för att fylla sin funktion.

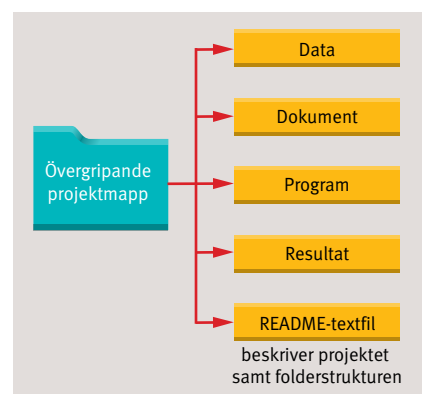
Om man exempelvis vill studera om patienter med multipelt myelom har ökad risk att drabbas av blodproppar [1] är det viktigt att notera inklusions- och exklusionskriterier för att ingå i myelom-kohorten, vilka klassifikationskoder för diagnosen blodpropp som använts samt varifrån dessa data hämtas. I detta exempel är myelom exponering och blodpropp utfall. Här bör även beskrivas vilken mätmetod som används för att analysera sambandet mellan exponering och utfall. Tänk noga kring datainsamling; ibland händer att intressanta frågeställningar bara blir halvt besvarade för att man inte tänkt igenom exakt hur data ska analyseras. Tänk efter före!

### Att bearbeta datamaterial med statistiska program

För många kan det kännas som ett oöverstigligt hinder att lära sig ett statistikprogram, men det är väl investerad tid. I de flesta statistikprogram finns möjligheten att analysera data genom att klicka sig fram i menyer och därför krävs inte förkunskaper i programmering. Det är viktigt att du skriver ner hur du klickar dig fram i analysen, vilka variabler du lägger in i tabeller och modeller, för att senare kunna återskapa din analys. Enklart är att föra en loggbok, där du skriver ner vilken ordning du klickade dig fram och vad resultatet blev. Spara resultatet i en fil och protokollför i loggboken. Det bästa är dock att analysera data genom att skriva programkod för att göra uträkningarna (inte



Figur 1. Dataflöde i ett forskningsprojekt.



Figur 2. Rekommenderad mappstruktur.

bara klicka sig fram i menyerna). Programkoden sparas i en fil, som du enkelt kan köra för att återskapa dina resultat vid uppdateringar och vid referentgranskning. Det är också enklare att leta efter fel i programkoden, jämfört med menydrivna analyser. Programkodfilen i sig är en dokumentation över vilka analyser du gjort, och du behöver därför inte skriva en särskild loggbok för dina analyser. Kommentera programkodfilen väl, till exempel genom att i ord skriva ut vilka beräkningar de olika programfunktionerna utför. Det gör det lätt att läsa och förstå även för personer som inte i detalj känner till hur just din statistiska programvara fungerar.

Ofta har menydrivna programvaror en modul som automatiskt skapar programkod. Denna kod kan man spara för att senare kunna köra om sin analys. På så vis kan man kombinera ett menydrivet arbetssätt med skriven programkod.

En tumregel är att varje programkodfil endast ska ha ett syfte, till exempel ett program för att skapa analysdatasetet och ett program för deskriptiv statistik. Namnge filen så att den beskriver vad programmet gör, exempelvis »skapa\_analysdata« eller »deskriptiv\_statistik«.

En loggfil kan ofta skapas av statistikprogramvaror och ger

### FAKTA 1. Exempel på analysplan: struktur, rekommenderade rubriker och typ av information som bör ingå

| Rubriker                                            | Förklaring                                                                                                                                                                                                              |
|-----------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Frågeställning och hypotes                       | Kort beskrivning av syftet med studien                                                                                                                                                                                  |
| 2. Förkortningar och termer                         |                                                                                                                                                                                                                         |
| 3. Studiepopulation – studiedesign                  | Beskriv hur data har samlats in samt hur stor studiepopulationen är.                                                                                                                                                    |
| Inklusionskriterier                                 | Beskriv hur studiepopulationen valts ut, samt vilka exkluderingar som gjorts och hur många som exkluderades i varje steg.                                                                                               |
| Exklusionskriterier                                 |                                                                                                                                                                                                                         |
| 4. Information om viktiga variabler                 |                                                                                                                                                                                                                         |
| Utfall                                              | Vilken sjukdom studeras?                                                                                                                                                                                                |
| Exponering                                          | Vilken variabel/riskfaktor är av primärt intresse?                                                                                                                                                                      |
| Andra viktiga variabler                             | Vilka andra variabler är av intresse för den statistiska analysen?                                                                                                                                                      |
| 5. Beskrivning av datahantering och mappstruktur    | Var kan man hitta viktig information om projektet? Var sparas originaldata, program osv? Referera t ex till loggbok, kodböcker och forskningsplan.                                                                      |
| 6. Statistiska analyser                             | Ange vilka statistiska analyser som planeras, samt var resultat från analyserna kan hittas.                                                                                                                             |
| 7. Kontaktinformation/Lista över alla medförfattare |                                                                                                                                                                                                                         |
| 8. Referenser                                       | Nyckelreferenser listas med fördel här.                                                                                                                                                                                 |
| Appendix 1 – Mötesanteckningar                      | Inkludera t ex även e-postkorrespondens.                                                                                                                                                                                |
| Appendix 2 – Resultat                               | Huvudsakliga fynd för publikationen, t ex de slutgiltiga tabeller och figurer som valdes ut för publikation. Dessa utgör oftast endast en liten mängd av alla resultat som producerats under ett typiskt projekts gång. |
| Appendix 3 – logg/dagbok                            | Kort avrapportering av exempelvis resultat från känslighetsanalyser och beslut som tagits.                                                                                                                              |

**FAKTA 2. Exempel på loggbok för att studera om patienter med multipelt myelom har ökad risk att drabbas av blodpropp.**

| Input data                                                            | Programkodfil               | Output-data                                | Beskrivning                                                                                                                             |
|-----------------------------------------------------------------------|-----------------------------|--------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------|
| MM_patienter_SCR <sup>1</sup> .data<br>(Antal observationer N=18 627) | MM_patienter_SCR_rensad.txt | MM_patienter_SCR_rensad.data<br>(N=18 627) | Rensa ett dataset som innehåller alla individer med multipel myelom registrerade i Svenska cancerregistret mellan 1958 och 2006.        |
| Tromboser_PR <sup>2</sup> .data<br>(N=1795760)                        | Tromboser_nr1_PR.txt        | Tromboser_nr1_PR.data<br>(N=264897)        | Skapa ett dataset som innehåller alla »första förekommande« fall av tromboser registrerade i patientdataregistret mellan 1964 och 2006. |
| MM_patienter_rensad_SCR.data<br>Tromboser_nr1_PR.data                 | Analysdata.txt              | Analysdata.data<br>(N=18 627)              | Skapa analysdataset med alla patienter och information om eventuella blodproppar (utfall).                                              |

<sup>1</sup> SCR, Svenska cancerregistret; <sup>2</sup> PR, Patientregistret

en fullständig sammanställning av körningen av en viss programfil. Spara gärna loggen/resultatet efter körningen (ge den samma namn som programfilen), så att du inte behöver köra om programmet varje gång du vill kontrollera ett resultat. Undvik att använda å, ä, ö och mellanslag i fil- eller variabelnamn. Alla programvaruföretag ger nätkurser och har nybörjarböcker till försäljning – allt för att du som användare snabbt ska bli bekant med programmet. Använd även inbyggda introduktioner (tutorials), supportservice eller sektionen »vanliga frågor« (FAQ, frequently asked questions) på företagets hemsida. Support ingår alltid i licensen, tveka inte att mejla om du har frågor. Dessutom finns ofta mycket material och tips från andra användare i forum på nätet.

**Att ha koll på dataflödet med kod- och loggböcker**

För att hålla reda på programfilerna och datafilerna bör man använda sig av loggbok och kodböcker, som beskriver hur dessa filer är relaterade och vad de innehåller [2]. I loggboken (som inte ska förväxlas med statistikprogrammets loggfiler) beskriver du alla programfiler du skapat, vilka datafiler de använder, vilka datafiler (uttryckt i ord) de skapar samt vad programmet gör. På detta sätt får man överblick över hur program- och datafilerna hänger ihop och därmed över dataflödet. En helhetssyn över dataflödet minskar dessutom risken att man av misstag skriver över datafiler och därmed riskerar att tidigare analyser blir oreproducerbara. Viktigt är att aldrig ändra informationen i originaldatafilerna då alla resultat bör kunna härledas i dessa. Fakta 2 ger ett exempel på en loggbok, som kan konstrueras både grafiskt och, som här, i tabellform.

Kodböcker innehåller beskrivningar för variablerna i datafilerna, varje datafil ska ha en egen kodbok. Det ska tydligt framgå hur variablerna är kodade (Fakta 3). Använd alltid beskrivande namn då du namnger nya variabler, så att man enkelt förstår vilken information den innehåller.

**Levande dokumentation ger tidsvinster**

Som tidigare nämnts bör analysplanen uppdateras kontinuerligt under hela projektet. Ha gärna en sektion i analysplanen där du för »dagbok« över vad du gjort. Du märker snart hur mycket kortare startsträckan blir när du återupptar arbetet med projektet. Utöka regelbundet analysplanens olika sektioner, t ex den statistiska analysen med detaljer över vilka metoder du använt, varför just de metoderna använts och om några förändringar sker. Skriv även in vilka inklusions- och exklusionskriterier du har för din studiepopulation och hur många som exkluderades i varje datarensningsteg.

Under arbetets gång upptäcker man ofta att någonting måste göras om. Kanske ska en variabel kategoriseras annorlunda eller några ytterligare personer i analysdatafilen exkluderas. Gå då tillbaka till den programfil där den kategoriseringen eller exklusionen ursprungligen utfördes istället för att försöka lösa problemet där det dök upp. Analysen riskerar annars att

**FAKTA 3. Exempel på kodbok**

Exempel på utdrag ur en kodbok som beskriver analysdata-materialet som använts för att studera om patienter med multipelt myelom har en ökad risk att drabbas av blodpropp.

| Variabel | Beskrivning                | Kodning                                                                |
|----------|----------------------------|------------------------------------------------------------------------|
| lopnr    | Löpnummer för MM-patienten | 1–18 627                                                               |
| kon      | Patientens kön             | 1 = man<br>2 = kvinna<br>9 = information saknas                        |
| alder_mm | Ålder vid diagnos          | 1 = <40<br>2 = 40–49<br>3 = 50–59<br>4 = 60–69<br>5 = 70–79<br>6 = >79 |
| M_conc   | M-proteinkoncentration     | 1 = <10 g/l<br>2 = 10 g/l eller mer<br>9 = information saknas          |

bli ett lapptäcke med osystematiska ändringar, istället för den välorganiserade struktur av separata programfiler med distinkta uppgifter som du eftersträfvade. Om du uppdaterar ett program, kom ihåg att även köra om alla programkodfiler som bygger på programmet du uppdaterat. Använd loggboken, där du lätt kan se hur programmen och datafilerna hör ihop. Kontrollera att du alltid gör allt i rätt ordning så att du inte av misstag analyserar en gammal version av datafilen.

Ett system för versionshantering ser till att gamla program och resultat inte skrivs över. Det finns olika program, men ofta räcker ett versionhanteringssystem som bygger på att till exempel alltid inkludera datum när man namnger sina program. Filen med senast datum är alltid den aktuella.

Dessa råd för uppdateringar kommer underlätta skrivandet av själva artikeln avsevärt. Metod- och resultatavsnitten finns redan i punktform i analysplanen! Innan du skickar in din artikel kan det vara bra att du skapar en kopia av ditt manuskript där du, för din egen skull, infogat hänvisningar/kommentarer till de programkodfiler som beräknar de olika resultaten i manuskriptet. På så sätt kommer referentprocessen bli effektivare och enklare att genomföra. Om du har ordning på alla dina filer, har skrivit bra dokumentation och gjort ett annoterat manuskript är det enkelt att veta vilken analys som skapat varje siffra i dina tabeller och dina figurer.

**Säkerhet vid hantering av persondata**

Personuppgiftslagen (Fakta 4) anger att alla som hanterar personuppgifter måste vidta lämpliga tekniska och organisa-

## FAKTA 4. Personuppgiftslagen, PUL

- PUL innehåller bestämmelser om hur och när personuppgifter får hanteras.
- Syftet är att skydda mot att individens personliga integritet kränks vid behandling av personuppgifter, det vill säga all information om en levande person som direkt eller indirekt kan kopplas till personen.
- Till känsliga personuppgifter hör ras eller etniskt ursprung, politiska åsikter, religiös eller filosofisk övertygelse, medlemskap i fackförening och uppgifter som rör hälsa eller sexualliv.
- Känsliga personuppgifter får endast användas för forskningsändamål om hanteringen godkänts av etikprövningsnämnd.

toriska åtgärder för att skydda personuppgifterna från hot eller störningar [3]. Med störningar och hot menas hantering av data som kan leda till att personuppgifterna förstörs/ändras så att de på ett eller annat sätt blir felaktiga eller att de oavsiktligt sprids. Detta kan exempelvis ske genom att obehöriga personer får tillgång till din dator, sabotage/dataintrång, olyckshändelser eller kanske rent av stöld av din dator.

Ofta har myndigheter/arbetsplatser en skriftlig säkerhetspolicy som beskriver organisationens säkerhetsstrategi och ansvarsfördelning. Likaså har Datainspektionen sammanställt allmänna råd kring konkreta säkerhetsåtgärder vid hantering av personuppgifter [4]. Exempel på tekniska åtgärder som man ganska enkelt kan se över själv och främst syftar till att motverka olika typer av dataintrång finns i Fakta 5.

Arbetar du hemifrån bör arbetsgivaren tillhandahålla säkerhetsrutiner, exempelvis kryptering av hårddisken samt ett system för säkerhetskopior (back-up). Man kan även kontakta sin datorleverantör för råd kring rimliga tekniska lösningar som kan minska skadan vid en eventuell förlust av datorn.

## Att tänka på vid samarbeten

Innan man påbörjar ett samarbete med andra forskare bör man tillsammans göra en plan för hur filer ska delas och hur arbetet ska dokumenteras. Ska det finnas en huvudkopia av datamaterialet på ett ställe, hur ska det rensas och vem är ansvarig för dokumentationen? Det ska finnas etiskt godkännande för all användning av datamaterialet, så om datamaterialet ska bearbetas på flera ställen ska det täckas av godkännandet. Ett gott råd är att inte lägga flera orensade kopior av datamaterialet på olika platser, eftersom det lätt kan leda till att olika rensningar görs på olika ställen och att resultaten blir olika. Det är klokare att ha en huvudfil av materialet på ett ställe som sedan delas ut till övriga samarbetspartner. Överföringen av filen måste ske på ett tekniskt säkert sätt, exempelvis med kryptering och en lösenordsskyddad överföring. E-post är ett klart olämpligt sätt att överföra okrypterade filer och ska aldrig användas för filöverföring av känsliga personuppgifter [5]. Olika lagar gäller för överföring av personuppgifter inom Sverige, EU och utanför EU. Mer information finns hos Datainspektionen (<http://www.datainspektionen.se>). Generellt kan man säga att det som är lagligt i Sverige även är det inom EU, medan det utanför EU nästan alltid krävs speciella avtal, och till majoriteten av världens länder är det olagligt att överföra patientdata.

## Arkivering – slutet på projektet

Efter publikation är det dags att arkivera projektet. Till arkivet läggs alla handlingar som är nödvändiga för att reproducera de publicerade resultaten, inklusive etik tillstånd, rådatafiler, forskningsdokumentationen samt den publicerade artikeln. Fullständiga regler kring hur länge materialet bör bevaras och vilka filer som kan gallras innan arkiveringen finns beskrivna i

## FAKTA 5. Tips på tekniska åtgärder

- Lösenord bör innehålla stora och små bokstäver plus siffror.
- Byt lösenord regelbundet.
- Skydda ditt lösenord så att andra inte kan logga in på din dator.
- E-posta aldrig lösenord.
- Logga alltid ut när du lämnar din dator obebaktad.
- Flera personer ska inte ha samma användaridentitet.
- Skriv inte ut känslig information på en skrivare som obehöriga personer har tillgång till.
- Ett bra virusskydd och en brandvägg skyddar dina data från att skadas av skadliga program.
- Spara inte känslig data på USB-minnen eller CD-skivor. De är lätta att glömma och är inte att betrakta som säkra för långtidsförvaring av viktiga data.

arkivlagen [6, 7] och alla myndigheter har en egen arkivbildningsplan. Det finns flera skäl till att bevara forskningshandlingarna. Dina resultat ska kunna kontrolleras av andra och ditt forskningsmaterial kan komma i fråga för återanvändning långt efter att ditt projekt har avslutats. Ditt material kan även komma att efterfrågas vid en eventuell utredning om en anklagelse för vetenskaplig oredlighet [8]. Att tänka på i samband med arkiveringen är således att de format som dina olika typer av filer är sparade som måste vara läsbara många år framöver. Det kan därför vara nödvändigt att konvertera vissa filer som finns sparade i icke-stabila format (till exempel datafiler som endast kan öppnas med en specifik version av en programvara) till mer beständiga format (exempelvis kommassparade textfiler). Det är också viktigt att uppdatera README.txt-filen och kontrollera att den är lätt att hitta i projektmappen.

## Ord på vägen

Det kan kännas övermäktigt att ställas inför dessa råd, det vet vi efter flera års undervisning kring datahantering och dokumentation. Men det handlar om att lägga om sitt arbetssätt, och sådant måste få ta tid. Det viktigaste är inte att dokumentationen är perfekt, utan att du börjar någonstans. Fokusera inte på gamla synder. Lagg kraften på pågående och kommande studier. Gör en analysplan och skriv en kodbok för din datafil. När du sedan sätter dig ner för att skriva artikeln ett år senare kommer du att inse att det var mödan värt. Din forskning har blivit tydligare, och ditt arbetssätt mer effektivt och tidsbesparande.

■ *Potentiella bindningar eller jävsförhållanden: Inga uppgivna.*

## LÄS MER Engelsk sammanfattning Läkartidningen.se

### REFERENSER

1. Kristinsson SY, Pfeiffer RM, Björkholm M, et al. Arterial and venous thrombosis in monoclonal gammopathy of undetermined significance and multiple myeloma: a population-based study. *Blood* 2011;117(14):3938.
2. Juul S, Lauritsen JM, Jørgensen A. Take good care of your Data. Aarhus universitet, Danmark 2004. [citerat 2 januari 2012] <http://www.epidata.dk/downloads/takecare.pdf>
3. Statistiska centralbyrån, Socialstyrelsen och Datainspektionen. Personuppgifter i forskningen – vilka regler gäller? Reviderad januari 2011 [citerat 2 januari 2012]. <http://www.datainspektionen.se/Documents/faktabroschyr-pul-forskning.pdf>
4. Datainspektionen. Datainspektionens allmänna råd – Säkerhet för personuppgifter. Reviderad november 2008 [citerat 2 januari 2012]. <http://www.datainspektionen.se/Documents/faktabroschyr-allmannarad-sakerhet.pdf>
5. Datainspektionen. IT-säkerhet och myndigheters e-tjänster. November 2008 [citerat 2 januari 2012]. <http://www.datainspektionen.se/Documents/faktablad-e-tjanster-sakerhet.pdf>
6. Arkivlag (SFS, 1990:782).
7. Riksarkivet. Riksarkivets föreskrifter och allmänna råd om gallring av handlingar: Statliga myndigheters forskningsverksamhet. RA-FS 1999:1
8. Gustafsson B, Hermerén G, Peterson B. God forskningssed. Stockholm: Vetenskapsrådets rapportserie; 2011:1. ISSN: 1651-7350