



EpiTour – uvod u

EpiData

Unos i dokumentovanje podataka

[Http://www.epidata.dk](http://www.epidata.dk)

Translated to Serbian language by: Marko Kovacevic

Verzija : 11. Januar 2001

© Jens M. Lauritsen, Michael Bruus & Mark Myatt

EpiData

EpiData je program baziran na windows 95/98/NT i namenjen za:

- Definisanje strukture podataka
- Jednostavan unos
- Unos i primenu validacionih principa
- Editovanje / ispravljanje već unetih podataka
- Proveru da li su podaci konzistentni po varijablama
- Štampanje ili izlistavanje podataka radi dokumentovanja grešaka-kontrolu i pronalaženje grešaka
- Upoređivanje dvostruko unetih podataka
- Eksportovanje podataka radi dalje analize u statističkom softveru i programima

EpiData radi sa windows 95/98/NT/Professional/2000 i Machintosh sa RealPc emulator.

Način citiranja programa EpiData :

Lauritsen JM, Bruus M, Myatt M. EpiData - A tool for validated dataentry and documentation of data. County of Funen, Denmark & Brixton Health UK. 2000. (Nađi verziju u okviru O)

Način citiranja EpiTour uvoda:

Lauritsen JM, Bruus M, Myatt M. EpiTour - An introduction to validated dataentry and documentation of data by use of EpiData. County of Funen, Denmark & Brixton Health UK. 2000. <http://www.epidata.dk/downloads/epitour.pdf> (vidi Verziju gore)

Ovaj dokument je dostupan kao EpiTour.hlp datoteka instaliran je sa EpiData i u obliku PDF datoteke EpiTour.pdf koja omogućava štampanje istog kao celine.

Za dalje informacije i download poslednje verzije vidi <http://www.epidata.dk>

Uvod i osnove

Šta je EpiData ?



EpiData program za unos i dokumentovanje podataka.

EpiData koristimo kada smo podatke već prikupili na papiru i želimo da ih pripremimo za statističku analizu. Podaci mogu biti prikupljeni u anketama ili u bilo kojem drugom obliku na papiru. EpiData **nije** napravljen za analize, to rade mnogi drugi programi.

U radu sa EpiData mogu da se primenjuju principi ”**kontrolisanog unosa podataka**”. Kontrolisano znači da će program EpiData korisniku dozvoliti da unese samo podatke koji ispunjavaju zadate kriterijume, npr. specifikovane dozvoljene vrednosti uz koje su priključene tekstualne oznake (1 = Ne 2= Da), proveru raspona (samo npr. starosti 20-100), dozvoljene vrednosti(npr. 1,2,3 i 9) ili dozvoljeni datumi (npr. 29. febr1999 neće prihvatiti).

EpiData je pogodan za jednostavne skupove podataka sadržane u pojedinačnim anketama, ali i za skupove podataka koji se sastoje iz brojnih razgranatih formulara sa podacima. **EpiData** je freeware i može da se preuzme sa [Http://www.epidata.dk](http://www.epidata.dk). Na istom vebu je istorijat razvoja programa.

Principi EpiData baziraju na jednostavnosti DOS programa EpiInfo, koji ima brojne korisnike širom sveta. Osnovna ideja je u tome da se pisanjem redova ankete isti delovanjem programa pretvaraju u obrazac za unos podataka. Kada je obrazac oblikovan ujedno je spreman i za prihvatanje podataka koji se unose u različita polja.

Ukoliko želite da probate rad sa EpiData preuzmite ga sa veba i instalirajte ga na svoj računar.

Jedna od važnih karakteristika EpiData je da ne ometa rad računara. EpiData se sastoji od programske datoteke i izvesnog broja help datoteka. Ni jedna druga datoteka se ne instalira. (U tehničkom smislu to znači da se ne instaliraju DLL datoteke ili sistemske datoteke – opcije se memorišu u registru.)

Registracija

Mi želimo da se korisnici registruju preko obrasca koji se nalazi na www.epidata.dk. Registracijom ćete dobijati informacije o inovacijama a pomoćićete nam da se odlučimo u kojem pravcu da nastavimo sa razvojem – a i u ubeđivanju drugih da ulože sredstva u razvoj. Registraciju možete da poništite slanjem pisma na adresu info@epidata.dk

Dijagram načina rada možete da preuzmete sa www.epidata.dk

Principi unosa podataka:

EpiData je sa tehničkog stanovišta usmeren na empirijske podatke. To znači da nema opcija koje bi se bavile biasom uzorkovanja, sigurnosti registra, dizajna studije i sl. Složenost struktura podataka u studijama je različita, od jednostavnog upitnika-ankete do kombinovanja podataka poreklom iz većeg broja izvora.

Vidi: Proces unosa podataka – principi

Ključne reči u procesu su: “reproduktibilnost”, “ispisivanje odluka”, “memorisanje svih datoteka”. “Potrebno je praćenje svake vrednosti u svakoj varijabli svake opservacije od konačne datoteke unazad do trenutka kada je opservirana, izmerena ili zabeležena u anketu”.

Vrlo često postoji potreba da se podaci učine anonimnim. Podaci se identifikuju na osnovu identifikacione varijable umesto identifikovanja na osnovu imena ili ličnog registarskog broja. Ovakva varijabla se naziva "id". U pojedinim državama ona se zasniva na "soundex" šifri koja predstavlja numerički prevod slova u prezimenu. Stoga će deo procesa biti posao oblikovanja specijalne datoteke koja sadrži informacije o identitetu, memoriše se posebno i koristi se samo za specijalne namene, npr. praćenje pacijenata jedne kohorte. Ili se koristi višekratno kao što je to primer pri ekstrahovanju informacija iz registra radi praćenja ishoda bolesti. Posebni aspekt je način raspoređivanja podataka po temama, vidi Proces unosa podataka u praksi

Neke korisne adrese na internetu sa sadržajima Biostatistike, Epidemiologije, Javnog zdravlja, EpiInfo itd.:

Data types and analysis: <http://www.sjsu.edu/faculty/gerstman/EpiInfo>
 Simple manuals: http://mkn.co.uk/help/extra/people/Brixton_Books
 EpiInfo home page: <http://www.cdc.gov/epo/epi/epiinfo.htm>
 Statistical routines: <http://www.oac.ucla.edu/training/stata/>
 Epidemiology Sources: <http://www.epibiostat.ucsf.edu/epidem/epidem.html>
 Epidemiology lectures: <http://www.pitt.edu/~super1/>

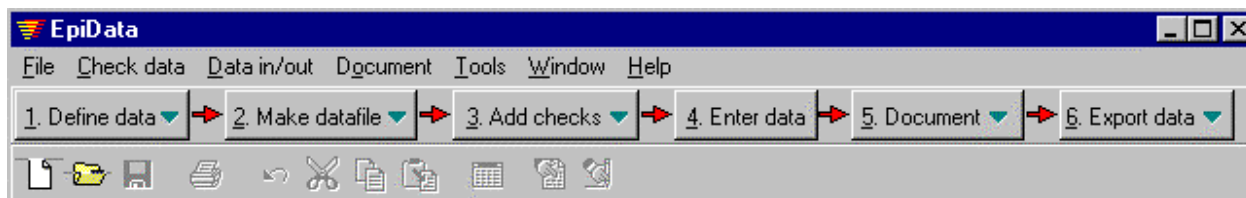
Freeware za unos podataka, izračunavanja i dijagrame:

EpiData (current program) for dataentry is available at www.epidata.dk
Epicalc 2000 Epidemiological oriented calculator. <http://www.myatt.demon.co.uk/>
EpiGram for drawing flowcharts and diagrams <http://www.myatt.demon.co.uk/>

Sada ste spremni da počnete sa poglavljem Način rada sa EpiData

Način rada sa EpiData

Skrin programa EpiData ima standardni windows oblik sa jednom linijom menija i sa dva toolbara (koji mogu i da se isključe). U zavisnosti od vrste posla meni se menja.

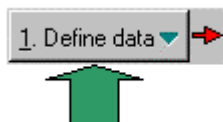


“Tulbar radnog procesa ” korisnika vodi od "1. Definisanje podataka " do “6. Eksport podataka”.

Drugi tulbar pomaže pri otvaranju datoteka, štampanju i drugim poslovima koji će biti kasnije objašnjeni.

- A. Ukoliko želite, možete da isključite tulbarove u Window Meniju, ali u EpiTour ćemo raditi pomoću tulbarova.
- B. Sada startujte EpiData.
- C. Nastavite da radite sa EpiData i čitajte ovo uputstvo (EpiTour).
- D. U meniju Help naćićete informacije o tome kako da se registrujete kao korisnik EpiData.
Registrovani korisnici će dobijati informacije o inovacijama.
- E. Nastavite sa zadatkom 1. Definisanje i testiranje Obrasca za unos podataka

Definisanje i testiranje unosa



- 1.. Kliknite na deo “Definiši podatke” i na “nova qes datoteka”. U Epi-Editoru se prikaže prazna datoteka sa nazivom “untitled”. U vašoj studiji qes datoteka definiše varijable. “Qes” je skraćenica od “questionnaire-anketa, upitnik”, i u nju se mogu uneti sve vrste informacija pomoću programa EpiData. Upitnik je zajednički naziv za takve datoteke.
- 2.. Memorišite praznu datoteku i dajte joj naziv **prvi.qes**.
datoteke se memorišu kroz “file meni” ili pritiskom na kombinaciju tipki “Ctrl+S”.
Zapazite da je Epi-Editor naziv “untitled” izmenio u “prvi.qes”.

Sada u Epi-Editor ukucajte redove koji slede.

Objašnjenje: Svaki red ima tri elementa:

- A.. Naziv varijable (npr. v1 ili ekspozicija).
- B.. Tekst koji opisuje varijablu. (npr. pol ili "datum rođenja ")
- C.. Definiciju unosa, npr. ## za dvocifreni broj.

```
Moja prva anketa
id          <idnum >
V1 pol      #
V2 Visina (metri)    #.##
v3 Datum rođenja <dd/mm/yyyy>
s1 Država u kojoj živi
s2 Mesto (adresa)    <a          >
t1 Današnji datum <Today-dmy>
```

Vrste polja koja ser koriste u EpiData (kojima se definišu varijable) su:

<u>Tip</u>	<u>Primer</u>
Tekst	
ID-broj	<IDNUM>
Numeričko	### ##.##
Velika slova	<A> <A >
Soundex	<S >
Boolean	<Y>
Datum	<dd/mm/gggg> <mm/dd/gggg>
Današnji datum	<today-dmy> <today-mdy>
Tabulator	@<a> @##

Tabulator polje se jedino koristi za poravnavanje drugih polja. Korisno je ukoliko se koriste proporcionalni fontovi.

3.. Ponovo memorišite datoteku. Pritiskom na kombinaciju tipki Ctrl+s

4.. Sada ćemo pregledati obrazac.



Pritisnite na tulbar okvir 2 i odaberite "Pregled obrasca " ili pritisnite Ctrl+T.

Na ekranu se prikazuje pregled ili provera obrasca **prvi.qes**. Probajte da premeštate kursor po poljima koristeći tipke sa strelicama. Uočite da se vrsta varijable prikazuje na statusnoj liniji. U ovaj obrazac ne možemo da unosimo podatke, jer je namenjen samo za proveru broja varijabli.

Ukoliko želite da dodate varijable preklopite na Epi-Editor i dodajte šta želite.



Pritisnite na drugi tulbar za prikaz helpa za oblikovanje polja ili pritiskom na Ctrl+Q možete da isključite help ili pritiskom na "Esc".



Pritiskom na Ctrl+W ili klikom na ikonu aktivira se drugi alat. To je alat za "autokompletiranje" određenih delova upitnika. Npr. ukoliko dodamo jedan #,_, <d, itd. pogledajte šta će se dogoditi. Probajte i pokažite 1.1 ili 8.

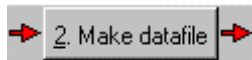
U polju za datum će **EpiData** pomoći u popunjavanju datuma: Ukoliko unesemo "040599" u polje za datum dmy, EpiData će to oblikovati u "04/05/1999". Ukoliko su unesene dvocifrene vrednosti za godine, vek će biti 1900 za godine od 50 do 99 i 2000 za godine od 00 do 49. Ukoliko unesemo "0405", program će upotrebiti tekuću godinu, tj. u godini 2001 memorisaće kao "04/05/2001". Cifru "040503" memoriše kao "04/03/2003" Posle unosa program sve datume proverava. Tako npr. datum 29021999 neće prihvatiti.

Tabulatorsku oznaku @ koristimo tako da je postavimo ispred polja, time je polje poravnano na levo.

Kada ste obrazac završili, zatvorite obrazac i opciju Epi-Editor (Vidi meni Datoteka).

Nastavite sa sledećim delom. Oblikovanje datoteke sa podacima

2. Oblikovanje datoteke za podatke



Pritisnite **2** (Alt+2) i prihvatite **"prvi.qes"** i **"prvi.rec"** kao nazive za "napravi datoteku za podatke". Time će se i fizički oblikovati datoteka na disku. Pregled obrasca postoji jedino u memoriji računara.

Pritienemo **OK** i oblikovali smo datoteku za podatke na osnovu definicija koje smo ranije postavili.

Definisali smo:

A datoteku sa definicijom obrasca i memorisali je pod nazivom **prvi.qes**

B Realnu datoteku za podatke koja će sadržati podatke i memorisali je kao **prvi.rec**.

Rec je skraćenica za "record-zapis" ili opservacija.

(Ukoliko pogledate na datoteke pomoću windows eksplorera možda ćete videti samo dve datoteke sa nazivom prvi, ali ne i ekstenzije "qes" i "rec". Ukoliko je tako, podesite ekplorer da prikazuje i ekstenzije svih datoteka, u preferences)

Nastavite sa sledećim delom. Dodaj Kontrole

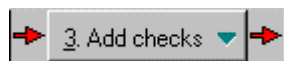
Izbor boja i opcija

U meniju Datoteka kliknemo na "opcije". Ukoliko to aktiviramo omogućeno nam je da zadamo brojne opcije. Između ostalog način menjanja boje skriona, boja polja za unos podataka. Neke od opcija se lakše razumeju, za neke je potrebno malo iskustva. Menjanjem opcija ne možemo da uništimo podatke.

Nazivi varijabli

Nazivi varijabli se mogu zadati tako da se oblikuju na dva načina: npr. v1pol (8 prvih znakova u redu) ili v1 (prva reč u redu). Ukoliko vam više odgovara da su nazivi tipa v1 v2 v3 t1, onda treba da odaberete opciju "Upotrebi prvu reč kao naziv polja " u "Oblikuj datoteku za podatke " (vidi meni datoteka - "opcije"). Pre pristupanja izmenama opcija zatvorite sve datoteke.

3 Dodaj kontrole unosa podataka



Kada "dodajete kontrole" time **specifikujete pravila** za unos podataka. **Dodavanje kontrole je opciono**. Ukoliko za sada niste zainteresovani za zadavanje kontrola pređite na poglavlje Unos podataka.

Range / Legal	
Jumps	1>WRITE
Must enter	Yes
Repeat	No
Value label	sex

Save Edit Exit

Opcija Dodavanje kontrola ima **pet osnovnih** i jedan "napredni" način:

- 1.. **Oznaka**: Dodavanje opisnog teksta numeričkim vrednostima (Oznaka vrednosti)
- 2.. Restrikcija unosa samo na određene vrednosti (**raspon, dozvoljeno**) (i **oznaka**)
- 3.. Specifikovanje redosleda unosa podataka npr. odgovori samo na određena pitanja samo za muškarce, (**jumps-skokovi**)
- 4.. **Obavezan unos** određivanje da je unos obavezan
- 5.. **Repeat-ponovi** – kopira vrednost iz prethodnog zapisa. Npr. ukoliko upisujemo podatke za više škola, razred se mnogo puta ponavlja.

Napredno: Help poruke i druge proširene definicije za izračunavanja, strukture if .. then ...endif itd. Ovo se dodaje pomoću dugmeta "Edit" koje je pored dugmeta Izlaz. Ovaj način se neće detaljnije objašnjavati u ovom dokumentu, vidi help datoteku.

Osnovna pravila:

Neka pravila budu jednostavna

Kada imate kontinuirane podatke koristite **raspon/dozvoljeno** (npr. od 1 do 100 ili 4.2 do 13.1)

Kada imate kategorijske podatke ili samo mali broj podataka koristite **oznake** (npr. 1,2 ili 1,2,3,4,5,9)

Primeri:

range 10-80 plus jedna vrednost 99

Jumps: Kod vrednosti 1 goto s2: 1>s2 . Kod vrednosti 9 goto s3: 3>s3

Range/legal	10-80,99
Jumps	1>§2,9>§3

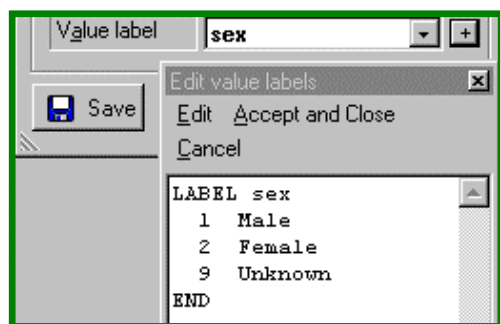
Sada dodajte oznake vrednosti varijablama.

Najpre proverite da slučajno niste u ovom polju specifikovali neki raspon/dozvoljeno .

Smisao oznake je u tome da ona poveže broj koji se unese sa tekstualnim opisom vrednosti u numeričkom polju. Ili za polje sa kratkim nizom da da opis takvog polja. Time se statistički izlazi čine bolje čitljivim i obezbeđuje se da je unesena vrednost pravilno shvaćena.

EpiData eksportuje informacije o oznakama vrednosti kada iste eksportujemo u Stata format. Komercijalni programi koji su na tržištu (npr. StatTransfer) mogu da konvertuju ove datoteke i u druge statističke programe npr. SPSS ili SAS.

Svaka oznaka ima svoj naziv. Ukoliko se naziv oznake prikaže u malom prozoru udesno od "Oznaka vrednosti" i pritisnemo na tipku "+" koja je udesno, prikazaće se vrednosti. Npr. pol kao što se vidi na slici.



Moguće je definisanje novih oznaka, npr. "veličina" koja sadrži 1 malo, 2 normalno, 3 veliko. Da bi definisali novu oznaku pritisnemo na "+" koji je udesno od oznake vrednosti polja. Otvoriće se mali prozor. Posle definisanja oznake možemo joj prikazati veliki broj varijabli.

Sada smo dodali kontrole za unos podataka u datoteku **prvi.rec** na osnovu strukture koju smo definisali u datoteci **prvi.qes**. Definicije ovih kontrola memorisane su u datoteci sa nazivom **prvi.chk**.

Nastavite sa odeljkom Unos podataka

Specijalne i primedbe.

- a- Pritiskom na kombinaciju tipki Ctrl+C aktivira se memorisanje postojećih definicija u buffer. Ukoliko se premestimo na sledeće polje u Obrascu i pritisnemo Ctrl+V definicije koje su kopirane će se primeniti i na to polje.
- b- Nazivi polja u koja treba da se izvrši skok mogu da se dodaju automatski. Najpre upišemo vrednost u delu za skok, (npr. 1) zatim dodamo znak > i kliknemo mišem na polje u koje želimo da kursor skoči (po ekranu se krećemo pomoću tipki PageUp/Page Down). Naziv polja će se automatski dodati u definiciju komande za skok.
- c- Dugme Edit omogućava pristup do svih do sada definisanih kontrola koje su polju dodate. Moguće je dodati u mnogo komplikovanije strukture (npr. if.. then ...endif), vidi help datoteku.

Moguće je definisanje help poruke koja će se prikazati u toku unosa podataka, npr:

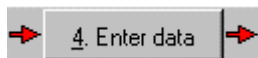
help "ovo je tekst koji se prikazuje licu koje unosi podatke"



- d- Za korisnike EpiInfo! Većina mogućnosti check jezika u verziji 6 EpiInfo je implementirano u EpiData.

Nastavite sa sledećim odeljkom . Unos podataka

4. Unos podataka



Jednostavno aktiviramo dugme **Unos podataka** na tulbaru i prihvatimo datoteku sa nazivom **prvi.rec** za unos.

Unećemo 5-10 zapisa. Zapis je jedna opservacija, jedna jedinica u podacima.

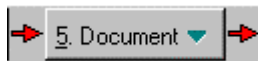
Kao što se vidi na **meniju** (Goto) moguće je prelaženje sa zapisa na zapis, kretanje ka prvoj ili poslednjoj varijabli, traženje vrednosti itd. Ove opcije mogu da se upotrebe ukoliko je potrebno da revidiramo podatke.

Pritiskom na kombinaciju tipki Shift+Del označavamo zapis kao da je izbrisan. Ukoliko je nešto uneseno netačno ili pomešano, zapis možemo da označimo kao da je za brisanje i ostavimo ga tako do daljnjeg. U narednim verzijama programa EpiData razvija se opcija za permanentno brisanje zapisa.

u izvesnim pričikama je korisno da se unesu komentari u toku unosa podataka. Sada dok je još otvorena datoteka prvi.rec, pritisnite na tipku F8. Time se aktivira mali editor dokumenata koji dodaje u tekst i datum i vreme. Sada možemo da unesemo neke komentare. Ova datoteka ima naziv **prvi.not** i treba je kopirati zajedno sa datotekama **prvi.rec**, **prvi.chk** i **prvi.qes**. EpiData to može da uradi i sam, vidi kasnije.

Nastavite sa sledećim delom . [Dokumentovanje podataka](#)

5 Dokumentovanje podataka



Ovaj deo programa nam omogućava da pripremimo informacije o strukturi, datumima, oznakama i sadržaju datoteka i varijabli. Ovde se prikazuje primer datoteke prvi.rec, koju smo oblikovali i toku rada prema ranijim poglavljima **EpiTour**.

DATOTEKA: C:\data\prvi.rec
Oznaka datoteke: Moja prva datoteka kao primer

Veličina: 612 bytes
Poslednja ravizija: 28. okt 2000 22:14
Broj polja: 7
Broj zapisa: 0
Upotrebljene kontrole: Yes (Last revision 28. okt 2000 22:32)

Polja u datoteci :

No.	Naziv	Oznaka	Varijable	Vrsta polja	Širina	Kontrole	Oznake vrednosti
1	id			ID-number	6		
2	v1	pol		Integer	1		pol 1: Musko 2: Žensko 9: Nepoznato
3	v2	Visina (metri)		Fixed number	4:2	Legal: 0.0-2.30,9	
4	v3	Datum rođenja	Datum (dmy)		10		
5	s1	Država		Text	28		
6	s2	Mesto (Current adresa)		Upper-case text	12		string blt: Baltimore cph: Copenhagen Denm rey: Reykjavik Icela sid: Sidney ndh: New Delhi mom: Mombassa bue: Buenos Aires
7	t1	Današnji datum-dmy			10		

Moguć je prikaz vrednosti u obliku:

Observation 1

id	1	v1	Male	v2	1.92
v3	12/12/1945	s1	denmark	s2	Copenhagen
t1	28/10/2000				

Pomoću šifarnika mogu da se oblikuju kondenzovane tabele frekvencije:

Example of this:

```

v1 ----- Age
      type: Integer
      range/legal: 0-100

      missing: 0/25
      range: [4,82]
      unique values: 22

v2 ----- Sex
      type: Integer
      value labels: sex
      range/legal: 1-2,2

      missing: 0/25
      range: [1,2]
      unique values: 2

      tabulation:   Freq.   Pct.   Value   Label
                   11     44.0     1     Male
                   14     56.0     2     Female

v3 ----- Temp
      type: Floating point
      range/legal: 36.00-40.00

      missing: 0/25
      range: [36.00,37.50]
      unique values: 12

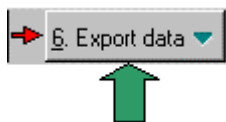
      mean: 36,84
      std. dev: 0,37

```

i sa "Primedbama unosa podataka " mogu da se edituju datoteke koje su aktivirane sa F8 u toku unosa podataka.

Nastavite sledećim delom . **Eksport Podataka**

6 Eksport, Analiza i opcije.



Najjednostavniji eksport je backup podataka. Za slučaj požara, poplave, kvara, krađe, provale u računare morate imati kopije datoteka koje ćete čuvati na raznim mestima.

Stoga svaki radni dan treba da se završi time da se naprave kopije - "**backup**". Aktivirajte opciju "Eksport podataka" i probajte da izvedete ovu operaciju.

Ukoliko želite da analizirate podatke pomoću programa Analysis sistema EpiInfo, tome možete pristupiti odmah. Datoteke koje su napisane pomoću programa EpiData je moguće odmah analizirati u EpiInfo, izuzetak su samo neka specijalna polja kao što su soundex i <today-dmy> .

Ukoliko koristite za analizu neki drugi program podatke je moguće eksportovati u sledeće vrste datoteka:

Jednostavni eksport (podaci i nazivi varijabli)

- 1..jednostavna (comma separated) ascii datoteka
- 2..dbaseIII
- 3..Excel

Kompletni eksport (podaci i nazivi varijabli, oznake)

- 4..Stata

Da bi se podaci koristili u drugim komercijalnim programima moguće ih je konvertovati programima kao što je Stat/Transfer (<http://www.circlesys.com/>) ili DBMS/Copy. Ukoliko npr. koristite SAS ili SPSS, eksport u Stata i konverzija.

Ovo je bila vaša prva zbirka podataka napravaljena pomoću EpiData. Nastavite sa [Support](#) i [About EpiData](#)

Koraci u toku procesa unosa podataka - principi

1 Postavljeni su cilj i namena istraživanja

- Opisana je hipoteza, određena veličina istraživanja, vremenska skala, Snaga izračunata...
- Obezbeđeno je finansiranje, Etički komitetitd.

2 Obezbeđivanje tehničkog kvaliteta podataka pri unosu podataka

Prikupite podatke i proverite njihov kvalitet sa čisto tehničkog stanovišta. Dokumentujte proces u datotekama i spiskovima grešaka.

- Postiže se primenom dozvoljenih vrednosti, kontrole raspona itd.
- Unos svih ili dela podataka dva puta da bi se otkrile greške u kucanju.
- Pronalaženje grešaka i popravke

3 Dostizanje konzistentnosti podataka i logike.

Istraživač treba da pregleda sve podatke. Treba da utvrdi:

- da li imaju logički smisao sa sadržajem predmeta (nema bake koja je mlađa od xx, npr. 35)
- Količinu nedostajućih podataka. Neke varijable će biti isključene ili će deo analize biti pod pitanjem zbog uticaja nedostajućih vrednosti na izračune.
- Odluke o broju odazvanih lica na istraživanje (N).

Opisujemo odluke u dokumentu zajedno sa opisima skupa podataka, varijabli itd.

4 Prečišćavanje podataka, izvedene varijable i konverzija u skup podataka koji će se analizirati.

U većini studija je neophodno da se podaci dobro pročiste i da se izvedu izračunavanja u izvedenim varijablama. Npr. u studiji praćenja treba utvrditi kada su bili periodi ekspozicije, povezati ankete sa podacima koji su poreklom iz registara, izračunavanje skala. Istovremeno je potrebno odlučiti prečišćavanju pojedinih varijabli, opservacijama u vezi sa nedostajućim podacima. Sve ove odluke treba da se dokumentuju.

5 Arhivirati kopiju podataka u arhiv podataka ili u sigurni sef. Uključiti i kopije svih planova projekta, obrazaca, anketa, spiskova grešaka, drugu dokumentaciju. Time se omogućava praćenje svake vrednosti u varijablama od konačnog skupa podataka do originalne opservacije.

Arhivirajte originale anketa i drugog materijala koji je u papirima, kao dokaz postojanja u skladu sa "Uputstvom za dobru kliničku praksu", "Etičkim komitetom za istraživanja" itd. (npr. u trajanju od 10 godina)

6 Aktuelna analiza i izračunavanja su izvršena. Sve analize su izvedene po principima reproducibilnosti. Odgovarajuća dokumentacija o tome će se čuvati kao dokumentacija istraživanja.

Proces dokumentovanja podataka

Postavljeni su ciljevi i namena istraživanja

- Opis hipoteze, veličina istraživanja, vremenska skala, snaga izračunavanja
- Obezbeđenje novca, etički komitet....itd.

Obezbeđenje tehničkog kvaliteta podataka prilikom unosa
Prikupljanje podataka i obezbeđivanje kvaliteta istih sa čisto tehničkog stanovišta
odrediti dozvoljene vrednosti, kontrole raspona itd
unos svih ili dela podataka dva puta, da bi se utvrdile greške kucanja.
traženje grešaka i ispravke
Dokumentovanje procesa u datotekama i spiskovima grešaka.

Konzistentnost i logičnost podataka.

Istraživač detaljno pregleda podatke. da bi utvrdio da li podaci baziraju na:
usklađenosti sadržaja (nema baka ispod starosti od xx, npr. 35)

Količina nedostajućih podataka. Neke varijable mogu otpasti iz analize ili su pod pitanjem
koliko na ocene utiču nedostajući podaci.

Odluka o boju ispitanih (N).

Opisati odluke u dokumentu zajedno sa opisima podataka, sastava varijabli itd.

Podaci očišćeni, izvedene su varijable i konverzija podataka za dalje analize.

U većini studija su neophodna čišćenja podataka i izračunavanja izvedenih varijabli. Npr. u studiji praćenja se moraju odrediti periodi ekspozicije, povezati ankete sa informacijama iz registra, izračunavanja skala itd.

Zajedno sa čišćenjem, odlučuje se o pojedinim varijablama utvrđuje se stanje nedostajućih podataka. To se mora dokumentovati.

Arhivirati kopiju podataka u arhiv ili u sef. Isto važi za kopije svih planova projekta, obrasce, ankete, spiskove grešaka, dr. dokumentaciju. Cilj je da se omogući praćenje svake vrednosti svake varijable od finalnog skupa podataka do originalnih opservacija.

- Arhivirati originale anketa i dr. materijala na papiru kao dokaz postojanja u skladu "Good Clinical Practice Guidelines", "Research Ethical Committees" etc (e.g. for 10 years)

Izvršena je aktuelna analiza i izračunavanja.

Sve analize se rade po principima reproducibilnosti. Odgovarajuća dokumentacija o tome čuva se zajedno sa dokumentacijom istraživanja

Proces unosa podataka u praksi.

U zavisnosti od konkretne studije, datelji procesa koji je gore prikazan biće nešto drugačiji. Stoga zahtevi za dokumentovanim unosom podataka i za proces čišćenja podataka mogu biti različiti. Pogledaćemo ovaj proces malo detaljnije.

a. Koji su izvori podataka na osnovu plana studije. Odluka o tome koji će izvori podataka činiti celokupni skup podataka studije. Npr. anketa, intervju obrazac i nešto uzoraka krvi. Uzorkuju /identifikuju se objekti studije (pacijenti). Generisanje anonimne ID varijablu.

b. Memorisati datoteku ID-KEY koja sadrži dve varijable: **id** i **Broj socijalnog osiguranja, Matični broj građana** ili drugi identifikator učesnika studije

c. Sakupiti podatke:

anketa (zajednička id varijabla): Unos podataka sa kontrolom na nivou varijabli za:

- dozvoljene vrednosti, rangove, filter pitanja (jumps), itd.

intervju obrazac (zajednička id varijabla): Unos podataka i kontrola na nivou varijabli za:

- dozvoljene vrednosti, rangove, filter pitanja (jumps), itd. .

uzorci krvi (zajednička id varijabla):

- Dobijanje podataka kao automatski prikupljene ili unostiti odgovore uz primenu odgovarajućih kontrole.

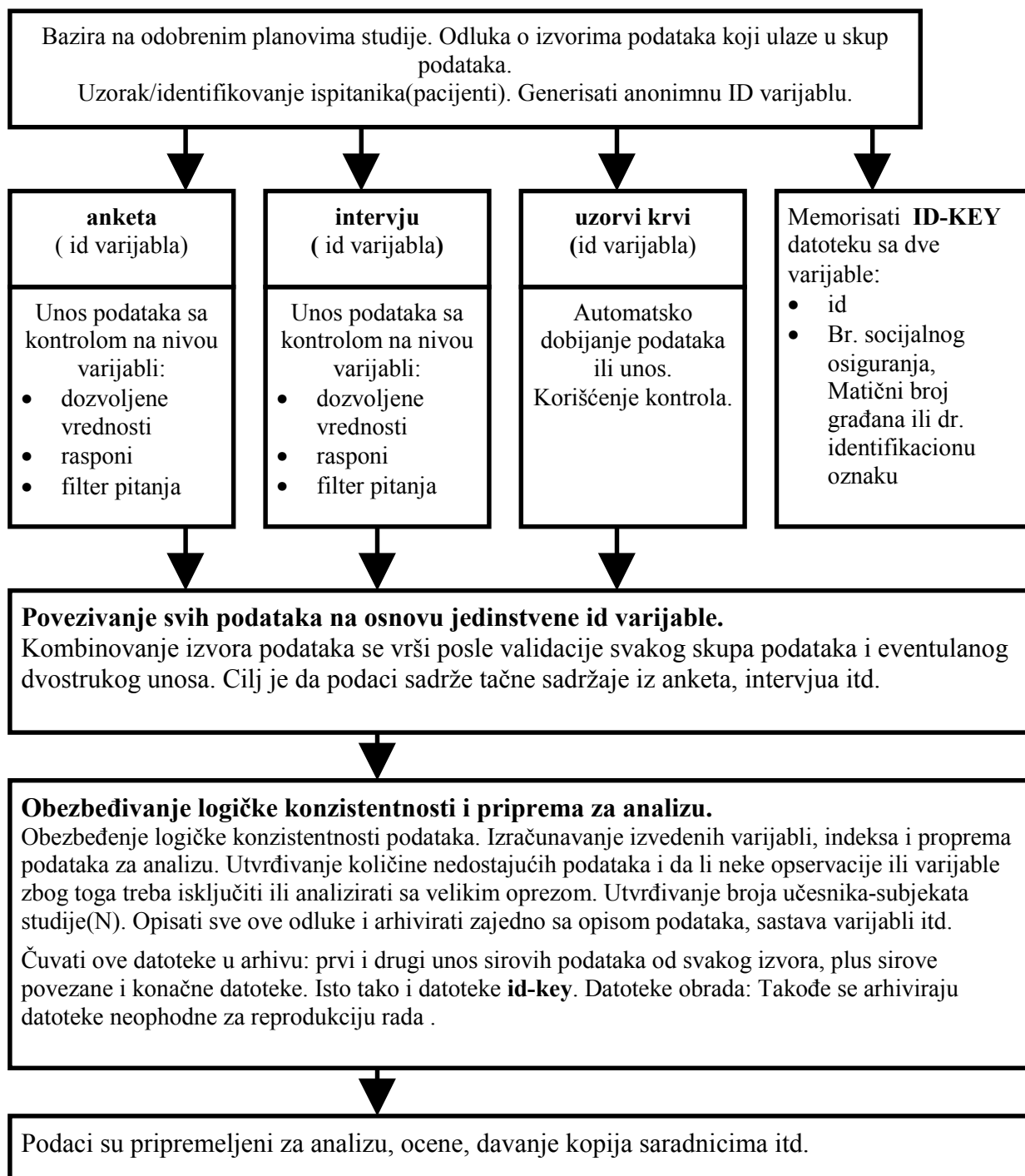
d. Povezivanje svih datoteka sa podacima na osnovu jedinstvene varijable.

Posle validacije i eventualnog dvostrukog unosa podataka vrši se povezivanje, kombinovanje izvora podataka. Cilj je da skup podataka sadrži tačnu repliku informacija koje su sadržane u anketama, obrascima intervjuja itd.

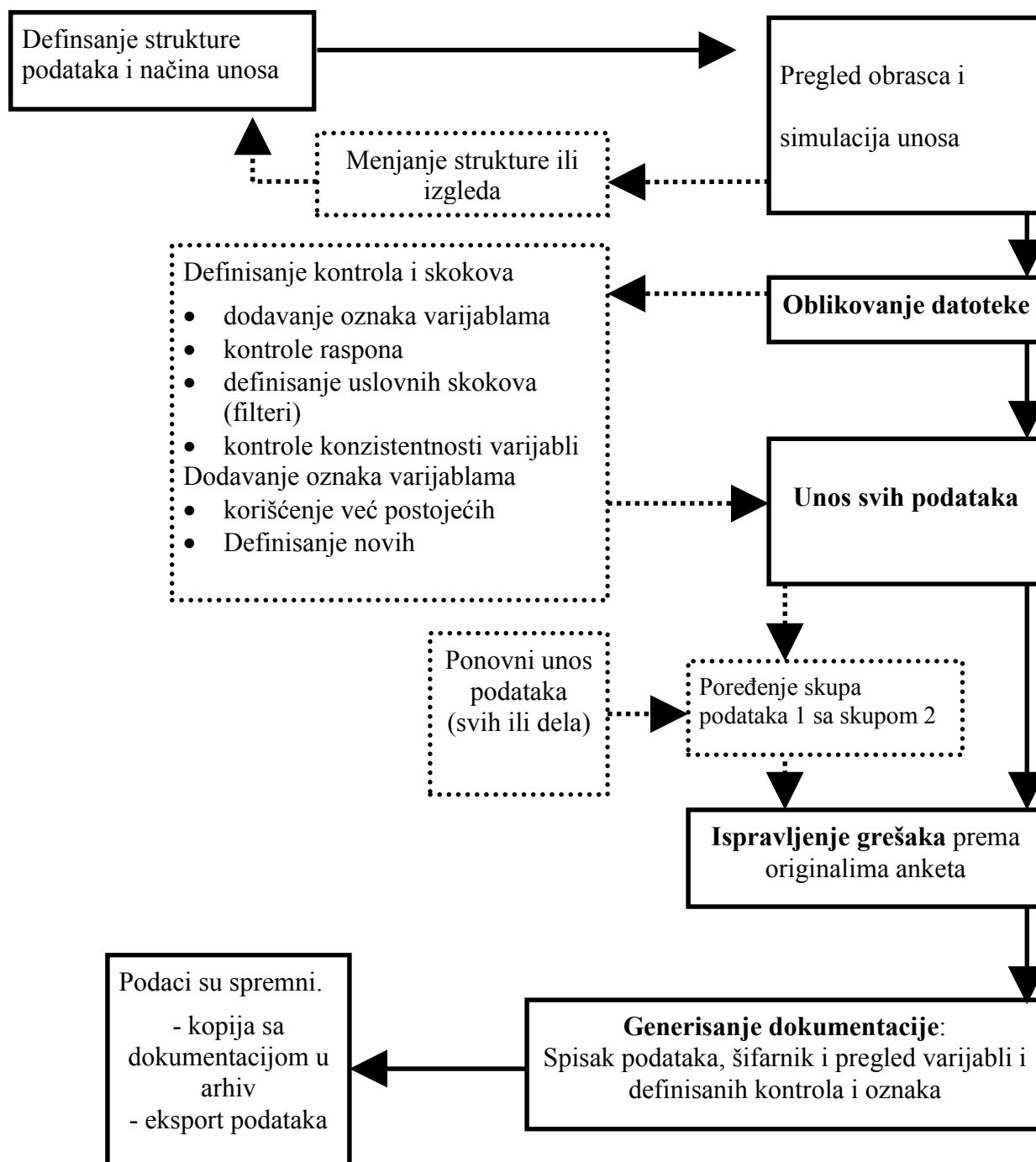
e. Obezbeđivanje logičke konzistentnosti i priprema za analizu.

Obezbediti logičku konzistentnost podataka. Treba izračunati izvedene varijable, indekse i pripremiti skup podataka za anлізу. Treba utvrditi kolika je količina nedostajućih podataka i da li je toliko da je neophodno da se izvesne opservacija ili varijable moraju isključiti ili se u analizi moraju tretirati sa posebnim oprezom. Potrebno je utvrditi broj odazvanih-učesnika (N). Ove se odluke moraju opisati i arhivirati zajedno sa opisima skupa podataka, sastava varijabli i dr.

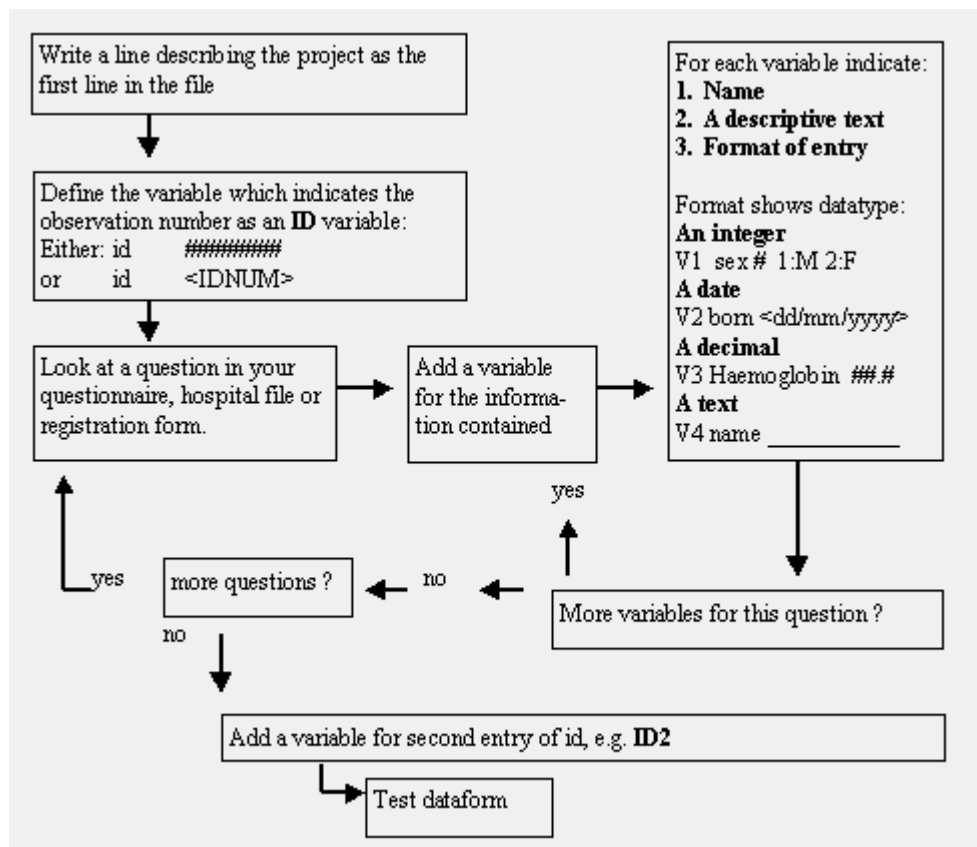
Datoteke se čuvaju u arhivi: prvi unos i drugi unos sirovih podataka iz svakog izvora, plus sirovi spojeni podaci i konačna datoteka . Isto tako se mora čuvati i **id-key** datoteka. Datoteke koje su rezultat obrada: takođe se arhiviraju jer su neophodne za reprodukciju-ponavljanje obrade.



Dijagram načina rada sa programom EpiData



Dijagram toka – Oblikovanje definicija podataka ("qes" datoteka)

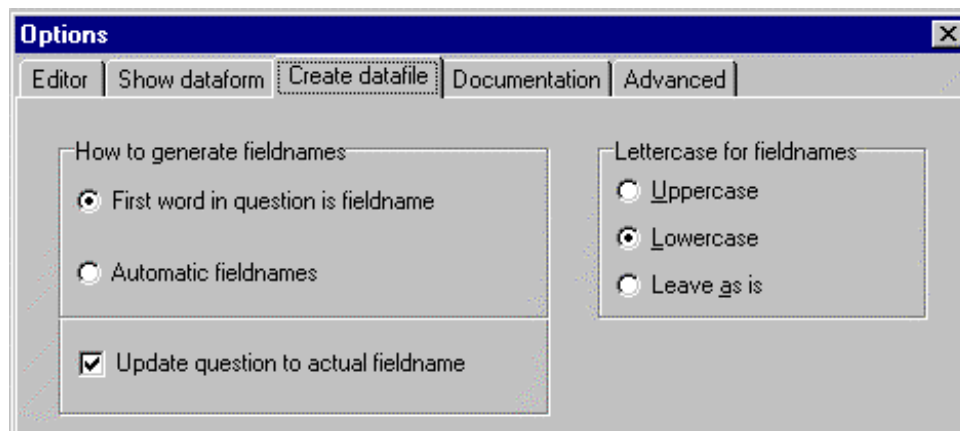


Ukoliko dodamo opisni tekst pre znaka koji definiše vrstu polja (npr. #) onda će tekst biti deo oznake varijable. Ukoliko ga dodamo posle neće biti deo oznake varijable. Zavisno od zadatih opcija varijablama možemo dati nazive v1, v2v8 ili v1starost v2pol ... ili

```

id <idnum>
V1   Starost ##
V2   Pol   #
V3   Temp  ##.##
V3a  Temp  ##.##
V4   LE    ##
V5   AB    #
V6   Cult  #
V7   Serv  #
V8   Trajanje  ##
  
```

v8 Trajanje u ovom primeru:



Ukoliko odaberemo "first word-prva reč" kao što je prikazano u opcijama (meni datoteka) dobićemo u gornjem primeru v1, v2.....v8.

Podrška

Ukoliko otkrijete greške ili bagove prilikom korišćenja programa ili imate sugestije za njegovo poboljšanje koristite formular koji se nalazi www.epidata.dk

Izvori podrške:

- 1..Pročitajte help datoteku za EpiData.
- 2..Pročitajte ovaj dokument- EpiTour
- 3..Skinite - Download sa <http://www.epidata.dk> EpiData help datoteku i datoteku EpiTour koje su u formatu "pdf", koji omogućava štampanje istih
4. Osnovni principi EpiData su isti kao i EpiInfo verzija 6. Dostupan je na sajtu epiinfo: <http://www.cdc.gov/epiinfo/>

Na žalost mi nemamo sredstava za pružanje opšte podrške u vezi sa problemima unosa podataka. Predlažemo da se obratite na EpiInfo internet diskusionu listu. Vidi <http://www.cdc.gov/epo/epi/epiinfo.htm>

Nastavite sa [O EpiData](#)

O EpiData

EpiData je program za Windows 95/98/NT/2000 namenjen unosu podataka (32 bit).

Program razvili

Jens M. Lauritsen, County of Funen, Denmark.

Michael Bruus, Denmark

Mark Myatt, Brixton Health, Wales and The Institute of Ophthalmology, London, UK.

Programirao

Michael Bruus, Denmark.

Registracija

Nemojte da oklevate da se registrujete kao korisnik ovog programa, upotrebite obrazac sa [Http://www.epidata.dk](http://www.epidata.dk) . Kao registrovani korisnik ćete dobijati informacije o inovacijama programa a ujedno ćete nam pomoći pri odlučivanju o daljem smeru razvijanja programa.

Finansiranje

Zahvaljujemo se donatorima koji su omogućili razvoj **EpiData**:

County of Funen, Denmark.

<http://www.fyns-amt.dk>

Brixton Health, UK

<http://www.brixtonbooks.demon.co.uk>

Danish Data Archives/ERAS, Denmark

<http://www.dda.dk>

University of Southern Denmark, Faculty of Health <http://www.sdu.dk/indexE.html>

Valid International. London UK

<http://www.validinternational.org/>

London School of Hygiene & Tropical Medicine, UK

<http://www.lshtm.ac.uk/>

International Centre for Eye Health, UK

<http://www.ucl.ac.uk/ico/ircpb.htm>

Način citiranja:

Lauritsen JM, Bruus M, Myatt M. EpiData, A tool for validated dataentry and documentation of data. County of Funen Denmark & Brixton Health UK. 2000. (version x.y)

EpiData je napravljen za vođenje qes-chk-rec strukture datoteka kao što je to u EpiInfo v6.

EpiInfo 2000 ima drugačije usmerenje. O tome više možete da pročitate na [epiinfo](http://www.epiinfo.org) homepage.

Zahvaljujemo se Andrew Dean iz CDC u Atlanta USA koji nam je dao na upotrebu izvorni kod EpiInfo v. 6. EpiData je u potpunosti nanovo napisan korišćenjem Delphi 3 for windows. Ujedno se zahvaljujemo 150 testera beta verzija programa sa Stata-liste, EpiInfo-liste koji su utrošili svoje vreme na testiranje.

Upozorenje

Program EpiData je razvijen a njegov rad testiran da bi se obezbedio siguran unos podataka i dokumentacije istih. Dali smo sve od sebe da napravimo siguran program, ali ne možemo preuzeti odgovornost za eventualne greške, gubitak podataka, radnog vremena ili druge gubitke prouzrokovane u vezi sa programom.