

Epi Info

- jeho popis a praktický význam



Petr Kachlík
Brno, 2020

Co je Epi Info

- zdarma získatelný a pro nekomerční účely použitelný softwarový balík typu „public domain“ (<https://www.cdc.gov/epiinfo/support/downloads.html>)
- je určený především pro odborníky i další zájemce v oblasti veřejného zdravotnictví
- Epi Info je obchodní značkou (TM) CDC (Centers for Disease Control and Prevention), Atlanta, Georgia, USA
- software, dokumentace a výukové materiály jsou ale veřejným vlastnictvím a lze je bez omezení kopírovat, distribuovat a překládat

Z čeho se Epi Info skládá a co umí

- sběr, záznam, zpracování a prezentace dat
- modulární systém
- spolupráce s databázemi, tabulkovými procesory a webovými prohlížeči
- program výživy a antropometrie
- příprava formulářů, kontrola dat na vstupu, jejich ochrana, komprese a kódování
- výuková cvičení

Z čeho se Epi Info skládá a co umí

- 5 základních modulů: Make View (tvorba dotazníků), Enter (vkládání dat), Analysis (statistická analýza), Epi Map (prezentace údajů v geografických souvislostech), Epi Report (generování výstupů)
- v současnosti je třeba k instalaci alespoň Win XP SP3, kolem 300 MB volného místa na HDD, alespoň 128 MB RAM, takt CPU minimálně 300 MHz
- starší verze běžely pod MS-DOS, Win 9x, NT, 2k..., v. 6 byla lokalizovaná do Cz vč. helpu

Epi Info Museum 1

- 1976: minipočítače se začínají prosazovat ve veřejném zdravotnictví v US, epidemiologie prasečí chřipky (sw SOCRATES, lg. FORTRAN)
- 1978-84: Epidemiologic Analysis System (EAS), lg. BASIC, Georgia
- 1978-84: sdružení epidemiologů, sestavení počítačové pracovní skupiny
- 1984: počátek systémového vývoje epidemiologického sw EPI-AID pro terénní praxi a výzkum

Epi Info Museum 1



 Centers for Disease Control and Prevention
CDC 24/7: Saving Lives. Protecting People.™



Epi Info Museum 2

- 1986: Epi Info verze 1, běželo pod DOSem
- distribuován 27stránkový manuál k programu a 5 1/4palcová (!) disketa
- Programs for Automatic File Definition, Data Entry and Tabulation = Epi Info
- sw zvládal sestavování dotazníků, sběr a analýzu dat, stal se rychle populárním

Epi Info Museum 2

Epidemiologic Analysis System

Centers for Disease Control
Epidemiology Program Office
Jeffery A. Dean - Author

September 1, 1985

- A. Data Entry and Editing (EnterInfo)
- B. Data Conversion Routines (Convert)
- C. Data Output and Analysis (Crosstab)
- D. Return to Master Menu

Enter option letter desired:

C\EAS>

EAS> TABLES VANILLA ILL

VANILLA	:	N	Y	ILL	
				:	Total
	N	18	3	:	21
	Y	11	43	:	54
Total				:	75

EAS>



Epi Info Museum 3

- 1987: Epi Info verze 2
- vázaný 112stránkový manuál, menu, modulární struktura
- tvorba dotazníku (EPED), práce s daty (ENTER, ANALYSIS), epi kalkulátor (STATCALC), grafika - histogramy (EPICURVE), práce se soubory (MERGE, IMPORT, CONVERT), základní statistické výstupy
- 1988: Epi Info verze 3
- 150stránkový manuál, vylepšená statistika a grafické výstupy

Epi Info Museum 3

```

Epi Info
August 28th, 1986 256K
Centers for Disease Control

Run program Other System Information

Word processor - Make a questionnaire
EnterInfo - Enter data
Analysis - Analyze data

Convert data for use in other programs
StatCalc - Epidemiologic calculator
Merge - Merge EpiInfo datasets
Import files from other programs
Run any other program and return

TO THE
of a
using Programs
Program Office
Centers for Disease Control
Atlanta, GA 30333

Please give copies of the programs and manual to your friends and colleagues.

Use 4-7 key to see menu
Select items by first letter or with 1,2,3,4 keys followed by 4-7

Programs listed in gray are not accessible
Insert disk containing programs and choose "Disk Changed" from Other menu.

Press F1 if you changed a disk. Press F10 to quit.
    
```

```

EPI INFO
Version 3
Epidemiology Program Office
Centers for Disease Control, Atlanta, GA 30333

<F10> - Done
EPED - Editor/word processor
ENTER - enter or edit data
ANALYSIS - Analyze data
CHECK - Set up data entry constraints
STATCALC - Epidemiologic calculator
EPICRAPH - Make graphs from tables
CONVERT - Make files for other programs
IMPORT - Use files from other programs
MERGE - Merge or update files
EDIT - Edit Epi-Info file headers
QUIT

Make a questionnaire or do general word processing

A Public Domain Program
Please give copies to your friends and colleagues.

"EPI M" or "EPI B" may be more visible on monochrome screens.
    
```

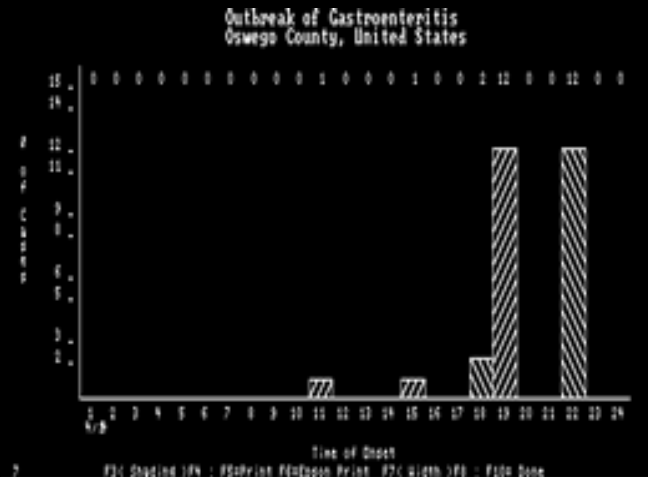
EPI> read c:\epi5\data\palufin

EPI> tables sexe palu

SEXE	PALU		Total
	0	1	
1	354	62	416
2	458	94	552
9	17	4	21
Total	829	160	989

An expected cell is < 5. Chi square is not valid.
 Chi square = -171930.22
 2 degrees of freedom.
 p value = 1.00000000

EPI>



Epi Info Museum 4

- 1988: Epi Info verze 4, mnoho nových funkcí, ale též řada chyb, spolupráce CDC a WHO
- vytvořen základ pro verzi 5

A N A L Y S I S

Epi Info Version 4.1

November 15, 1988

Developed by:

Epidemiology Program Office
Centers for Disease Control
Atlanta, GA 30333

Version 4.1 modifications by:

Global Programme on AIDS
World Health Organization
Geneva, Switzerland

Written by:

Jeffrey Dean

Overlay buffer:

57616 bytes

Epi Info Museum 5

- 1990: Epi Info verze 5
- 384stránkový manuál, kódování proměnných, regresní funkce, inovace grafiky, generátor zpráv, nutriční antropometrie
- první mezinárodní distribuce, dovolovala evropský formát data a národní abecedy
- zdrojové kódy byly přeloženy do francouzštiny, španělštiny, čínštiny, arabštiny a ruštiny
- 1992: Epi Map verze 1 a 2, ovládané myší

Epi Info Museum 5

Dataset: (None) Free memory: 254K
Use READ to choose a dataset

Output = Screen

Choose a command

General	QUIT	RECORD	CLS	IMMEDI
QUIT	Page control	BROWSE	COPY	WRITE
VARIABLES	TITLE	UPDATE	DIR	GENERATE
ROUTE	HEADER	GRAPHICS	ECHO	RELATE
LIST	FOOTER	PIC	SPACE	SET
PRINT	RESPACE	BAR	REMOVE	BREAK
TABLER	Use manip	BAR	TYPE	QUIT
TABLER	SELECT	HISTOGRAM	LOAD	RETURN
ROADS	SORT	LINE	SAVE	
REGRESS	DEFINE	SCATTER	Program ends	
SUMFREQ	LET	DOS commands	RUN	
SUMTABLES	IF	DOS	REPORT	

Commands

EPID
EPID
EPID

F1-Help F2-Commands F3-Variables F4-Browse F5-Printer on F9-DOS F10-Quit

Epi Info Version 5.01 - octubre 1990

Software de Dominio público para Epidemiología y vigilancia

Centers for Disease Control
Epidemiology Program Office
Atlanta, Georgia

World Health Organization
Global Programme on AIDS
Geneva, Switzerland

Programas	Descripción
EPED Editor de textos	Use EPED para: » Crear cuestionarios » Escribir memorias o cartas » Ser ayudado en investigaciones » Escribir o editar programas de Epi Info
ENTER Base de datos	
ANALYSIS Analiza datos	
CHECK Valida entradas	
STATCALC Calculadora Epi	
CONVERT Exporta datos	
IMPORT Importa datos	
MERGE Une ficheros	
VALIDATE Valida ficheros	

EPED es un procesador de textos para crear cuestionarios Epi Info o procesar textos. Contiene el sistema EPIAID para pensar sobre investigaciones de epidemias y producir borradores de informes de investigaciones.

Información de Epi Info
Quit, Sale de Epi Info

Params: EPED <Nombre de fichero o programa EPIAID>

EPED

[J], utilice una letra o espacio F9 Configuración, F10 salir

Epi Info Version 5.01 ENSP F 1910 - Septiembre 1991

LOGICIEL DU DOMAINE PUBLIC POUR L'EPIDEMIOLOGIE
C D C O M S
Centers for Disease Control Organisation Mondiale de la Santé
Epidemiology Program Office Programme Global sur le SIDA
Atlanta, USA Genève, Suisse

TRADUCTION FRANCAISE ENSP

Programas	Description
EDITE un texte	Utilisez EDITE pour : » Créer des questionnaires (Fichiers .QES) » Créer des masques de saisie lors d'enquêtes » Ecrire ou éditer des programmes Epi Info
SAISIE de données	
ANALYSE de données	
CONTROLE à la saisie	
CALCUL manuel	
EXPORTE des données	
IMPORTE des données	
FUSION de 2 fichiers	
COMPARE 2 fichiers	

EDITE est un traitement de texte pour créer des questionnaires Epi Info ou d'autres textes. Il contient le Système EPIAID d'aide à la production d'enquêtes épidémiologiques et à l'édition de rapports d'enquête.

Au sujet de Epi Info
Quitter Epi Info

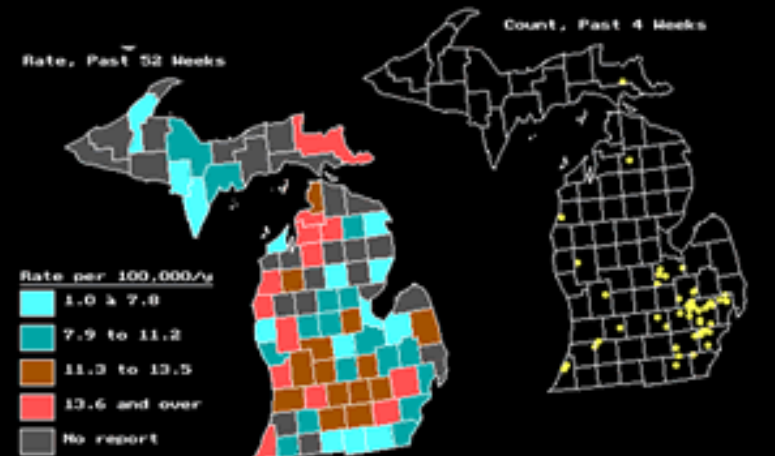
Paramètre : EDITE <Fichier Texte(.QES ,.PGM...)>

EDITE

[J], Lettre, Entrée ou Espace pour Valider F9 ->configurer, F10->Quitter

Salmonellosis Notifications in Michigan

Last Four Week Period Ending with 1994 Week # 12, 1994
Count Past Four Weeks and Rates Past 52 Weeks



Right Arrow, Right Mouse-Next Esc, F10-Quit

Epi Info Museum 6

- 1992: Epi Info verze 6
- 600stránkový manuál, vylepšená statistika, grafika, rozsáhlejší příklady, programovatelné menu, generování epidemiologických hlášení
- verzi 6 užívala i WHO a lokálně výzkumníci mnoha zemí světa, položen základ Win verze
- 1993: konference Microcomputers and the Future of Epidemiology, Atlanta
- 1994: Epi Info distribuováno již on-line, hotline, manuály v řadě jazyků včetně češtiny (!)

Epi Info Museum 6

- 1996: diskusní skupina na netu, další překlady sw a manuálu (celkem do 14 jazyků)
- 1997: ověřování sw a zpětná vazba od uživatelů, 150 000 kopií ve 117 zemích, 1 207 citací, 52 síťových zdrojů Epi Info nabízelo ke stažení
- 1998: příprava kompatibility s „y2k“, a to i zpětně (starší datové soubory) – v. 6.04b, c
- 2001: v 6.04d - rekompilace programu po úpravě zdrojového kódu v lg. TURBO PASCAL, řešila přechod na moderní CPU architekturu PC, typická distribuce na 3 FDD 3,5 palce, 1,44 MB

Epi Info Museum 6



Epi Info Museum 6

```

Dataset: C:\EPI6\06VEG0.REC (75 records)      Free memory: 292K
Criteria: All records selected                Time: 0.00 sec

Output - Screen
Exact 95% Mid-P limits for MLE              5.93 < OR < 109.15
Probability of MLE >= 22.15 if population OR = 1.0  0.00000026

RISK RATIO(OR)(Outcome:ILL++; Exposure:VANILLA++) 5.57
95% confidence limits for RR                1.94 < RR < 16.03

Ignore risk ratio if case control study

          Chi-Squares    P-values
-----
Uncorrected:      27.22    0.00000018  ←
Mantel-Haenszel:  26.86    0.00000022  ←
 Yates corrected:  24.54    0.00000073  ←

Commands
EPI6> tables vanilla ill
EPI6>

F1-Help F2-Commands F3-Variables F4-Browse F5-Printer on F9-DOE F10-Quit
    
```

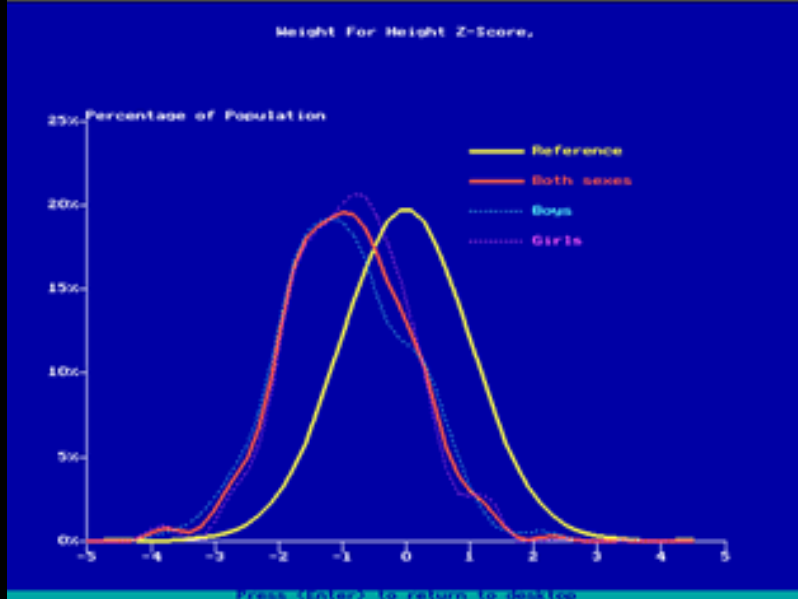
```

EPI Info Cluster Sampling Analysis vers 6.00
Epi Info CSAMPLE

Main
VAR1  LOCATION
VAR2  WEIGHT
VAR3  POPU
Crosstab
PRENATAL  Value 1
          Value 2
Output options
(*) Screen
() Printer
() File
File name

Tables
Menus
Cancel
Quit

F1 Help F10 Quit
    
```



CDC's EPI INFO HOME PAGE - Microsoft Internet Explorer

Epi Info's Home Page
Epidemiology Program Office

Epi Info is a series of microcomputer programs for word processing, data management and epidemiologic analysis, designed for public health professionals. Epi Info is easy to use, but also offers programming language for both data input and analysis so that permanent health information systems can be developed.

 DOWNLOAD SOFTWARE	Technical Support	GENERAL INFORMATION
	Frequently Asked Questions	Discussion Group
Links Outside of CDC		

[Software](#) | [Technical Support](#) | [General Information About Epi Info](#)
[Frequently asked questions](#) | [Epi Info Discussion Group](#)
[Other CDC Links](#)

[Feedback](#)

These pages last modified January 8, 1999

Internet zone




Epi Info Museum 7

- 1998: vznik týmu pro vývoj Epi Info 2000
- 1999: vydána beta verze Epi Info 2000 pro Win 95, 98 a NT, výstup ANALYSIS ve formátu html
- snadnější tvorba formulářů s dotazníky, manipulace s databázemi, statistikou, vylepšené grafické prezentace
- zpětná vazba, sdílení dat po internetu

Epi Info Museum 7



VANILLA by ILL

		ILL		
VANILLA	Yes	No	Total	
Yes	43	11	54	
No	3	18	21	
Total	46	29	75	



Single Table Analysis

	Point Estimate	95% Confidence Interval		
		Lower	Upper	
PARAMETERS: Odds-based				
Odds Ratio (cross product)	23.4545	5.8410	94.1811	(T)
Odds Ratio (MLE)	22.1490	5.9280	109.1500	(M)
		5.2153	138.3900	(F)
PARAMETERS: Risk-based				
Risk Ratio (RR)	5.5741	1.9383	16.0296	(T)
Risk Difference (RD)	65.3439	46.9212	83.7666	(T)

(T=Taylor series; C=Cornfield; M=Mid-P; F=Fisher Exact)

STATISTICAL TESTS	Chi-square	1-tailed p	2-tailed p
Chi square - uncorrected	27.2225		0.0000013505
Chi square - Mantel-Haenszel	26.8596		0.0000013800
Chi square - corrected (Yates)	24.5370		0.0000018982
Mid-p exact			0.0000001300
Fisher exact			0.0000002600

Epi Info: současnost a budoucnost

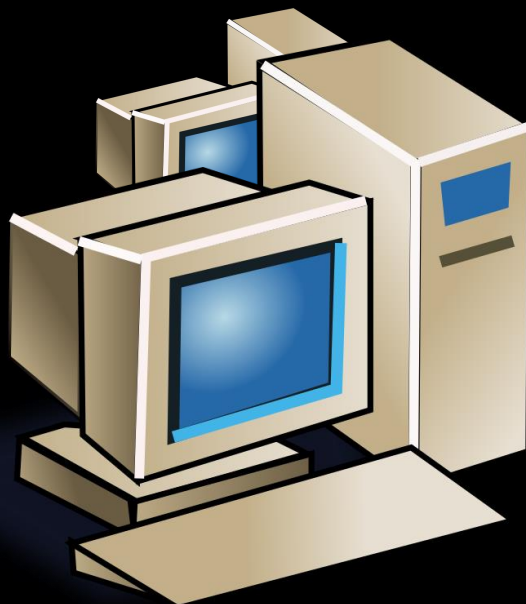
- 2000-2016: nové verze 3.5.x pro OS Windows vč. XP a 7, později W8-10
- v současnosti aktuální Epi Info verze 7.2 (2020)
- vážne lokalizace do Cz, u EI 3.5.x existuje však Sk verze
- rychlé a účinné sdílení informací
- k dispozici portable versions (mobily, tablety, notebooky), web a cloud version
- snadný přístup k odborným zdrojům přeloženým do hlavních světových jazyků

Epi Info: současnost a budoucnost

- vývoj metod pro přístup a analýzu dat pocházejících z klinických zdrojů s výstupy do veřejného zdravotnictví
- komunikace s domácnostmi (uživateli, respondenty) pomocí adresných služeb (videokonference, IT)
- zadávání dat z terénu pomocí hlasového vstupu
- adaptace sw pro nové operační systémy
- ...

Epi Info 3.5.1

Seznámení s programovým balíkem



Epi Info: úvodní obrazovka

- základní přehled a rychlá nabídka podprogramů
- pro verzi 3.5.1 existuje oficiální částečná lokalizace do slovenštiny, čeština bohužel chybí
- položky Zadat data, Náhled, Analýza dat, Výstup, Mapy, Hlášení, Web
- z menu Výživa navíc přístup k nutriční epidemiologii

Epi Info: práce s daty

- data lze: načítat, importovat, propojit, ukládat, exportovat, slučovat, editovat
- proměnné lze: definovat, zobrazit, přiřadit, překódovat
- lze vybrat data dle různých požadavků na jejich třídění
- základní statistika: seznam proměnných, tabulky absolutních a relativních četností, testování (ANOVA, χ^2), grafy, kartogramy

Epi Info: práce s daty 2

- pokročilá statistika: regresní analýza, analýza přežití,...
- výstupy: nastavení záhlaví, typu výstupu, tisku a ukládání souborů
- uživatelské programy a příkazy, makra
- interakce s uživatelem a nastavení prostředí (jazyk, pracovní adresář, typ a verze databáze...)

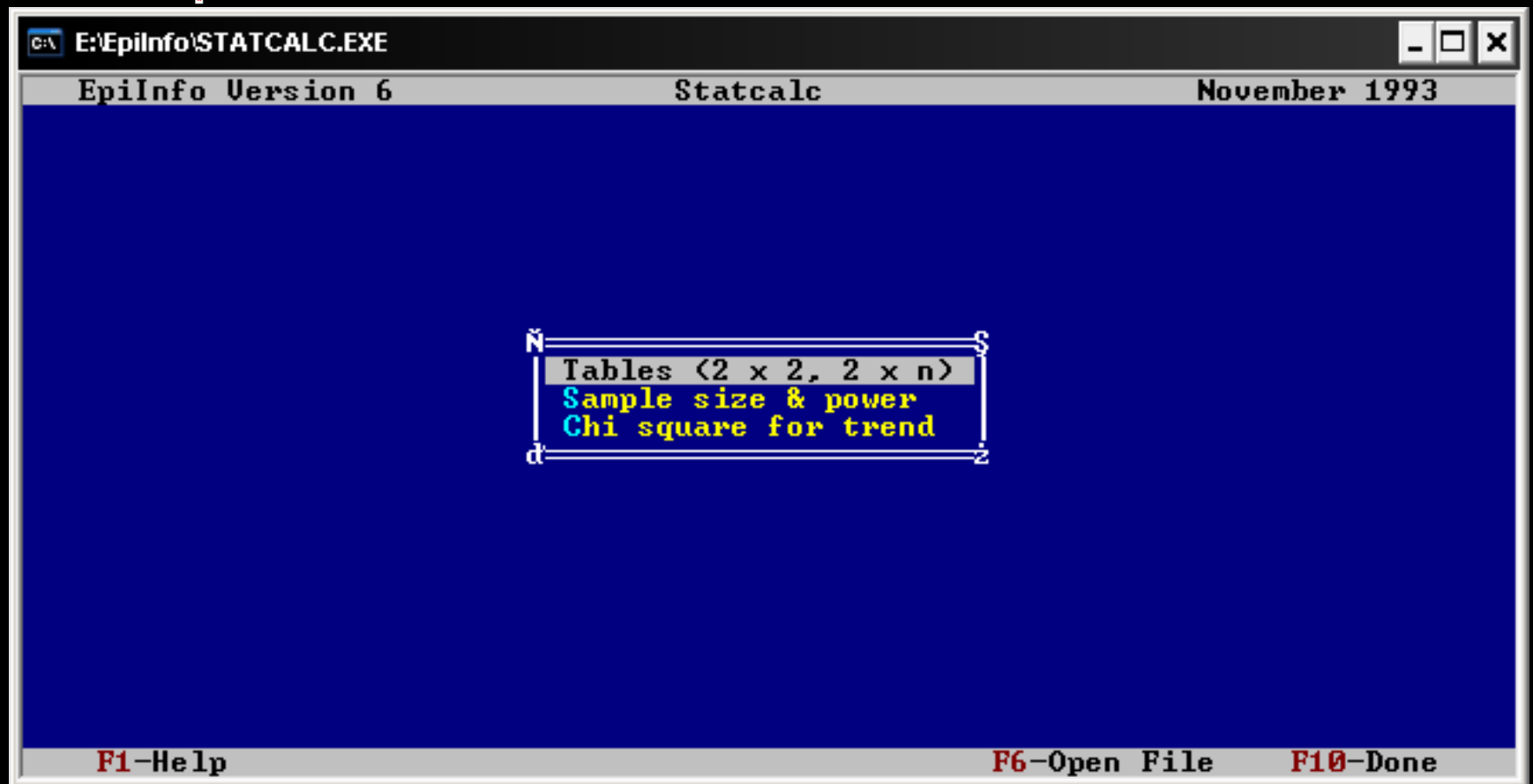
Epi Info: práce s daty 3

- porovnávání dat: výpis odlišných záznamů a odlišných proměnných včetně konkrétního rozdílu
- EpiLock: zaheslování a dekodování dat, archivace
- EpiInfo Help: standardní nápověda včetně rejstříku, vyhledávání klíčových slov a slovníku výrazů

Epi Info: pomocné programy

- StatCalc: čtyř- a vícepolní tabulky po test dobré shody (x² test)
- výpočet velikosti vzorku

Epi Info: StatCalc



Epi Info: StatCalc

E:\EpiInfo\STATCALC.EXE

EpiInfo Version 6 Statcalc November 1993

+ Disease -

+	15	25	40
-	12	48	60
E	27	73	100

x
p
o
s
u
r
e

Analysis of Single Table
Odds ratio = 2.40 (0.89 <OR< 6.51)
Cornfield 95% confidence limits for OR
Relative risk = 1.88 (0.98 <RR< 3.57)
Taylor Series 95% confidence limits for RR
Ignore relative risk if case control study.

	<u>Chi-Squares</u>	<u>P-values</u>
Uncorrected :	3.73	0.0534731
Mantel-Haenszel:	3.69	0.0546812
Yates corrected:	2.89	0.0889073

F2 More Strata; <Enter> No More Strata; F10 Quit

F1-Help F2-Stratum F5-Print F6-Open File F10-Done

Epi Info: StatCalc

E:\EpiInfo\STATCALC.EXE

EpiInfo Version 6 Statcalc November 1993

+ Disease -

+	18	250	268
-	45	310	355
E	63	560	623

x
p
o
s
u
r
e

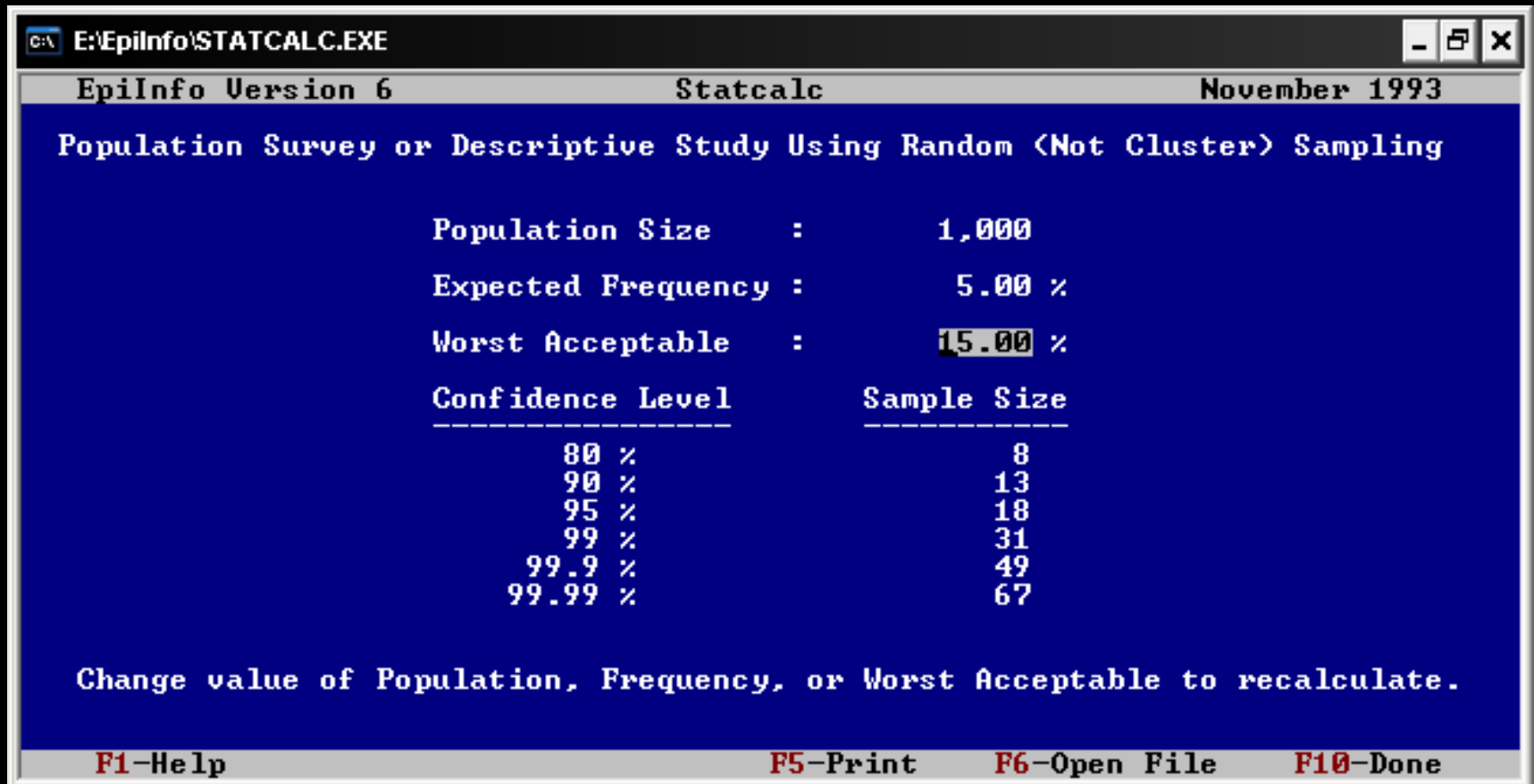
Analysis of Single Table
Odds ratio = 0.50 (0.27 <OR< 0.91)
Cornfield 95% confidence limits for OR
Relative risk = 0.53 (0.31 <RR< 0.89)
Taylor Series 95% confidence limits for RR
Ignore relative risk if case control study.

	Chi-Squares	P-values
Uncorrected :	5.97	0.0145753 ←
Mantel-Haenszel:	5.96	0.0146547 ←
Yates corrected:	5.33	0.0209679 ←

F2 More Strata; <Enter> No More Strata; F10 Quit

F1-Help F2-Stratum F5-Print F6-Open File F10-Done

Epi Info: StatCalc



The screenshot shows a DOS-style window titled "E:\EpiInfo\STATCALC.EXE". The window content is as follows:

EpiInfo Version 6 Statcalc November 1993

Population Survey or Descriptive Study Using Random (Not Cluster) Sampling

Population Size : 1,000

Expected Frequency : 5.00 %

Worst Acceptable : 15.00 %

<u>Confidence Level</u>	<u>Sample Size</u>
80 %	8
90 %	13
95 %	18
99 %	31
99.9 %	49
99.99 %	67

Change value of Population, Frequency, or Worst Acceptable to recalculate.

F1-Help **F5**-Print **F6**-Open File **F10**-Done

Epi Info: co je co (typy dat)

- data nominální: kvalitativní či kategoriální, neexistuje u nich velikost (např. pohlaví, rodinný stav, krevní skupina, DNA sekvence), dichotomická (alternativní): jen možnosti ANO/NE, diskrétní
- data ordinální: lze je uspořádat a určit, která hodnota je „vyšší“ a která „nižší“, nelze ale posoudit vzdálenost jednotlivých hodnot či kategorií (např. nejvyšší dosažené vzdělání), diskrétní
- data intervalová: lze je třídit a určovat i vzdálenosti mezi hodnotami (např. teplota tělesa), nezahrnují však vzájemný poměr hodnot, spojitá
- data poměrová: mimo jiné jsou vymezeny i poměry jednotlivých hodnot (obvykle základní fyzikální veličiny, u nichž je jasně definovaná nula), spojitá
- spojitá data lze převádět na diskrétní pomocí vytvoření kategorií

Epi Info: co je co (poloha)

- aritmetický průměr: součet členů souboru dělený jejich počtem, nevýhodou jsou extrémny (např. průměrné známky, platy)
- vážený průměr: absolutní četnosti vyjádříme jako relativní, jako podíly z celkového počtu; součet relativních četností (vah) v souboru je roven 1
- geometrický průměr: všechny hodnoty (n) se vzájemně vynásobí, ze součinu se vypočte n -tá odmocnina; užívá se zejména v ekonomii
- medián: přesný střed hodnot seřazených podle velikosti
- modus: hodnota s nejvyšší třídni četností

Epi Info: co je co (variabilita)

- zjišťuje se, jak se vzájemně liší jednotlivé prvky statistického souboru
- variační rozpětí: rozdíl mezi maximem a minimem; je citlivé na extrémny
- mezikvartilové rozpětí: rozdíl mezi horním (0,75) a dolním (0,25) kvantilem; není citlivé na extrémny
- rozptyl: disperse hodnot souboru kolem střední hodnoty základního souboru
- směrodatná odchylka: druhá odmocnina z rozptylu, má stejné jednotky jako sledovaná veličina
- variační koeficient (%): používá se při hodnocení znaků majících odlišné jednotky nebo lišících se mírou polohy, jde o relativní variabilitu vztaženou k průměru

Epi Info: StatCalc, co je co (prevalence)

- udává podíl obyvatel, kteří mají (nebo měli) určitý znak v dané lokalitě v daném časovém období
- v medicíně jde typicky o stav, chorobu nebo rizikový faktor (deprese, kouření,...)
- zjišťujeme ji v prevalenčních (průřezových) studiích
- intervalová prevalence: v celém průběhu sledovaného období
- bodová prevalence: vztažená k určitému datu

Epi Info: StatCalc, co je co (incidence)

- poměr nově vzniklých onemocnění (v určitém časovém rozmezí) k celkovému počtu osob ve sledované skupině
- kumulativní incidence: poměr počtu osob, které onemocní ve sledovaném období, k celkovému počtu osob zahrnutých do studie
- udává se přepočtena na určitý počet osob a délku sledování, v úvahu je brána pouze populace riziková pro příslušné onemocnění; nabývá hodnot od 0 do 1
- incidence rate: rychlost, s jakou se ve sledované populaci vyskytují dané zdravotní jevy; uvádí se jako počet případů za jednotku osobočasu (např. 0,20 případů na osoborok znamená, že se daný jev vyskytne u 1 osoby v průměru jednou za 5 let); nabývá hodnot od 0 do nekonečna

Epi Info: StatCalc, co je co (RR)

- relativní riziko (RR): ukazatel vystihující vztah mezi expozicí rizikovému faktoru a zdravotním následkem
- $RR=1$: daný faktor NEMÁ na vznik onemocnění vliv
- $RR>1$: expozice je RIZIKOVÝM faktorem
- $RR<1$: expozice je PROTEKTIVNÍM faktorem
- RR je vyjádřeno poměrem incidencí v exponovaném a neexponovaném souboru

Epi Info: StatCalc, co je co (AR)

- absolutní riziko (AR): pozorovaná nebo vypočítaná pravděpodobnost výskytu daného onemocnění ve studované populaci

Epi Info: StatCalc, co je co (CI)

- interval spolehlivosti, konfidenční interval (CI): intervalový odhad k získání představy o relativní četnosti a průměru základního souboru
- šířkou CI lze posuzovat náhodnou výběrovou chybu; čím užší CI, tím více se blíží hodnotám základního souboru
- hodnoty konfidenční hladiny rovny 0,9; 0,95; 0,99; 0,995
- pro 95% CI odpovídá hodnota 1,96 a pro 99% CI 2,58
- hodnota 95% nebo 99% znamená, že průměr leží uvnitř tohoto intervalu s pravděpodobností 95 % nebo 99 %
- CI pokrývá hodnotu známého průměru s danou spolehlivostí
- 99% CI znamená, že se ve 100 náhodných výběrech objeví daná charakteristika právě 99krát

Epi Info: StatCalc, co je co (ODDS)

- poměr šancí, ODDS ratio, OR: používá se u studie případů a kontrol, kde nelze aplikovat relativní riziko, poněvadž není k dispozici kumulativní incidence, ale můžeme kvantifikovat pravděpodobnost expozice případů a kontrol
- $OR = a*d/b*c$ (políčka čtyřpolní tabulky)
- a =pozitivní výskyt sledované události v experimentální skupině, b =pozitivní výskyt sledované události v kontrolní skupině, c =negativní výskyt sledované události v experimentální skupině, d =negativní výskyt sledované události v kontrolní skupině
- $OR=1$: ŽÁDNÝ korelát mezi rizikovým faktorem a onemocněním
- $OR>1$: pozitivní souvislost, RIZIKOVÝ faktor
- $OR<1$: PROTEKTIVNÍ faktor

Epi Info: StatCalc, co je co (kontingenční tabulka)

- užívá se k vizualizaci vztahu dvou statistických znaků a testování vztahů mezi nimi
- v řádcích tabulky jsou hodnoty prvního znaku, ve sloupcích hodnoty druhého znaku
- nezávislost: znaky se navzájem neovlivňují v tom, jakých konkrétních hodnot nabývají
- homogenita: očekávané četnosti v políčkách každého řádku jsou ve vzájemném poměru bez ohledu na konkrétní volbu řádku
- ověřuje test dobré shody (χ^2 test)

Epi Info: StatCalc, co je co (x2 test)

- nulová hypotéza: rozložení znaků se navzájem neovlivňuje
- alternativní hypotéza: rozložení znaků se vzájemně ovlivňuje
- v políčkách čtyřpolní tabulky četnosti větší než 5, jinak x2 není korektní, použijeme např. Fisher exact test
- chyba prvního druhu (alfa): zamítáme nulovou hypotézu, když platí (hladina významnosti)
- chyba druhého druhu (beta): přijímáme nulovou hypotézu, když neplatí
- při testování si předem zvolíme úroveň chyby (alfa), tj. hladinu významnosti (v medicíně obvykle 0,1 %, 1 % a 5 %)

Epi Info: co je co (korelace)

- vzájemný vztah mezi dvěma procesy či veličinami
- na základě korelace ještě nelze rozhodnout, co je příčinou a co následkem (korelace neimplikuje kauzalitu)
- určuje lineární vztah mezi znaky, její míru vyjadřuje korelační koeficient, který nabývá hodnot od -1 do +1
- -1: nepřímá závislost, +1: přímá závislost, 0: nezávislost, ale pozor: i při koeficientu 0 může existovat závislost, jen ji nelze popsat lineární fci !

Epi Info: co je co (regrese)

- metody, které umožňují odhadovat hodnotu veličiny závisle proměnné (vysvětlované či cílové proměnné, regresanda) na základě znalosti jiných veličin (vysvětlujících či nezávislých proměnných, regresorů či kovariát) – např. odhad počasí přes den na základě znalosti jeho předpovědi a momentálního stavu ráno
- je-li závisle proměnná skalár nebo vektor (např. délka dožití v letech), jde o tzv. lineární regresi x nelineární regrese-neuronové sítě
- je-li závisle proměnná diskrétní, jde o tzv. diskriminační analýzu (kategorie, třídy bez možnosti určit střední hodnotu) – např. předpověď koupě typu nápoje na základě věku a pohlaví zákazníka