

**REPORT PRODUCTION FOR A  
DATA AND SAFETY MONITORING BOARD.  
A UNIX-BASED SYSTEM FOR REPORT GENERATION**

**Thomas Cook, Kevin Buhr**

*Statistical Data Analysis Center*

*Department of Biostatistics and Medical Informatics*

*University of Wisconsin, Madison, Wisconsin*

<http://www.biostat.wisc.edu/biostat/sdac.html>

[cook@biostat.wisc.edu](mailto:cook@biostat.wisc.edu)

## **Abstract**

Preparation of reports for a Data and Safety Monitoring Board (DSMB) requires the use of data analysis, graphical production and documentation preparation software. Because the report must be based on the most current, up-to-date data, the report generation process needs to be standardized and automated to the greatest extent possible. We have developed a UNIX-based report generation system which uses SAS, Splus and  $\LaTeX$ , capitalizing on the strengths of each. SAS is used to perform many simple analyses on large data sets with many variables. The results are stored as datasets containing summary statistics for each variable and treatment group, p-values for treatment comparisons and labels and other formatting information. These datasets are imported into Splus to generate high quality, annotated graphics and  $\LaTeX$  files containing corresponding supporting tables.  $\LaTeX$  is used to combine the graphics and supporting tables with introductory and explanatory text and possibly material from other sources into a coherent report. The report is sequentially page-numbered with cross-references linking graphics to supporting tables. A table of contents and list of figures is automatically generated. We illustrate the report generation system using data from the Beta-Blocker Heart Attack Trial.

# Outline:

- Introduction to DSMB Reports
- Overview of report generation system
- Sample report pages
- Sample programs

# DSMB Reports

- Must be quick to produce, accurate and user friendly.
- Designed to present data, not interpretation.

## Desirable Features

- Annotated graphical displays
- Supporting tables
- Cross references between graphics and supporting tables.
- Open/closed session versions
- Sequentially numbered pages
- Table of Contents

## Process Overview

Our system primarily makes use of

- SAS
- S-PLUS
- L<sup>A</sup>T<sub>E</sub>X

to capitalize on the strengths of each product.

# SAS

- Facilitate data transfers
- Handle large datasets
- Restructure data
- Store information about variables
- Basic statistical analysis
- Create datasets for use in S-PLUS

# S-PLUS

- Create annotated graphical displays
  - Easily build complex graphs from simple elements
  - Good graphical quality
- Generate  $\text{\LaTeX}$  commands to produce corresponding summary tables

# L<sup>A</sup>T<sub>E</sub>X

- Powerful text formatting capabilities
- Compile the report
  - Import material from a variety of sources including S-PLUS graphics and tables.
  - Automatically generate Table of Contents, List of Figures, Index, and cross references.

---

# **SDAC Sample Report**

## **Beta-Blocker Heart Attack Trial**

### **(BHAT)**

April 19, 1999

---

See Poster P40

*A Sample Report for a Data and Safety Monitoring Board*  
for additional examples.

Statistical Data Analysis Center

---

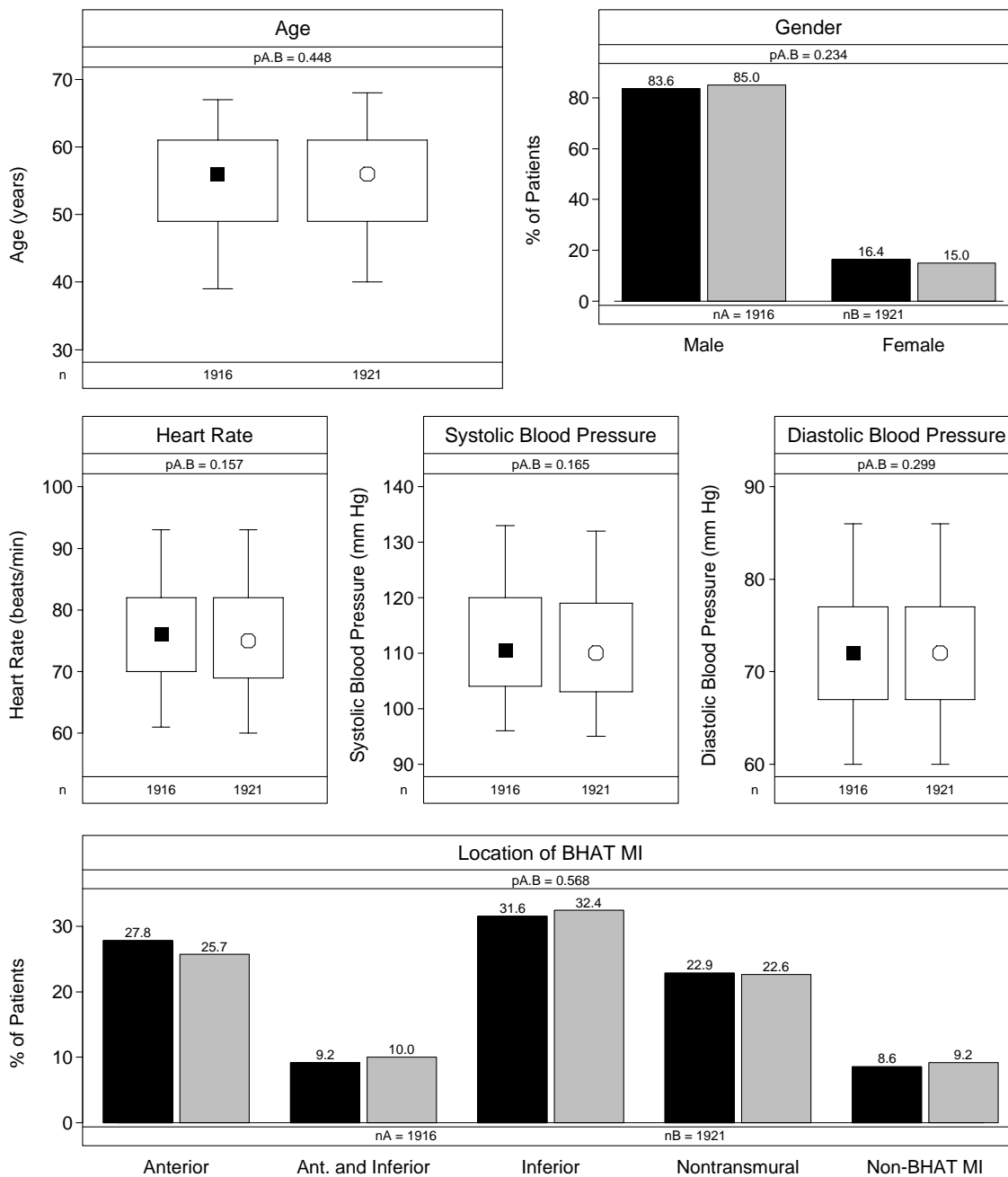
Department of Biostatistics and Medical Informatics  
University of Wisconsin–Madison

# Contents

<b>I</b>	<b>Introduction</b>	<b>1</b>
1	BHAT Protocol and Procedures . . . . .	2
2	Overview of Report . . . . .	4
3	Report Structure . . . . .	6
4	Notes on Chapter Contents . . . . .	7
5	Contact Information . . . . .	11
<b>II</b>	<b>Main Material</b>	<b>12</b>
<b>1</b>	<b>Accrual</b>	<b>13</b>
1.1	Patient Accrual over Time . . . . .	14
1.2	Patient Accrual by Center and Treatment . . . . .	15
1.3	Availability of Follow-up Data . . . . .	16
<b>2</b>	<b>Baseline Characteristics</b>	<b>17</b>
2.1	Demographics . . . . .	18
2.2	Medical History . . . . .	19
<b>3</b>	<b>Adverse Events</b>	<b>20</b>
3.1	Summary of Adverse Events . . . . .	21
3.2	Adverse Events by Follow-Up Year . . . . .	22
	⋮	
<b>III</b>	<b>Supporting Material</b>	<b>35</b>
<b>1</b>	<b>Accrual</b>	<b>36</b>
1.3	Availability of Follow-up Data . . . . .	36
<b>2</b>	<b>Baseline Characteristics</b>	<b>37</b>
2.1	Demographics . . . . .	37
2.2	Medical History . . . . .	38
<b>3</b>	<b>Adverse Events</b>	<b>40</b>
3.1	Summary of Adverse Events . . . . .	40
3.2	Adverse Events by Follow-Up Year . . . . .	41

Figure BASE-1

### Baseline Demographics



Information determined by baseline interview, physical examination, and ECG administered prior to randomization. See the *Introduction* for a discussion of "Non-BHAT MI" classification. Figure ANCI-1 on page 56 displays an Open Session Report version of this page. (See Table 1: *JAMA* 247:1708, 1982.)



See Table Set BASE-1 on page 37.

## Chapter 2

# Baseline Characteristics

## 2.1 Demographics

### Table Set BASE-1

#### Baseline Demographics: Age

See Figure BASE-1 on page 18.

Trt	Total Pats	Std		Median	Q1	Q3	P5	P95	Contrast	P-Value
		Mean	Dev							
A	1916	54.69	8.48	56.00	49.00	61.00	39.00	67.00	A.B	0.45
B	1921	54.93	8.41	56.00	49.00	61.00	40.00	68.00		

#### Baseline Demographics: Gender

See Figure BASE-1 on page 18.

	Treatment Group				Contrast	P-Value
	A		B			
	N	%	N	%		
Total Patients	1916		1921		A.B	0.23
Male	1602	83.61	1633	85.01		
Female	314	16.39	288	14.99		

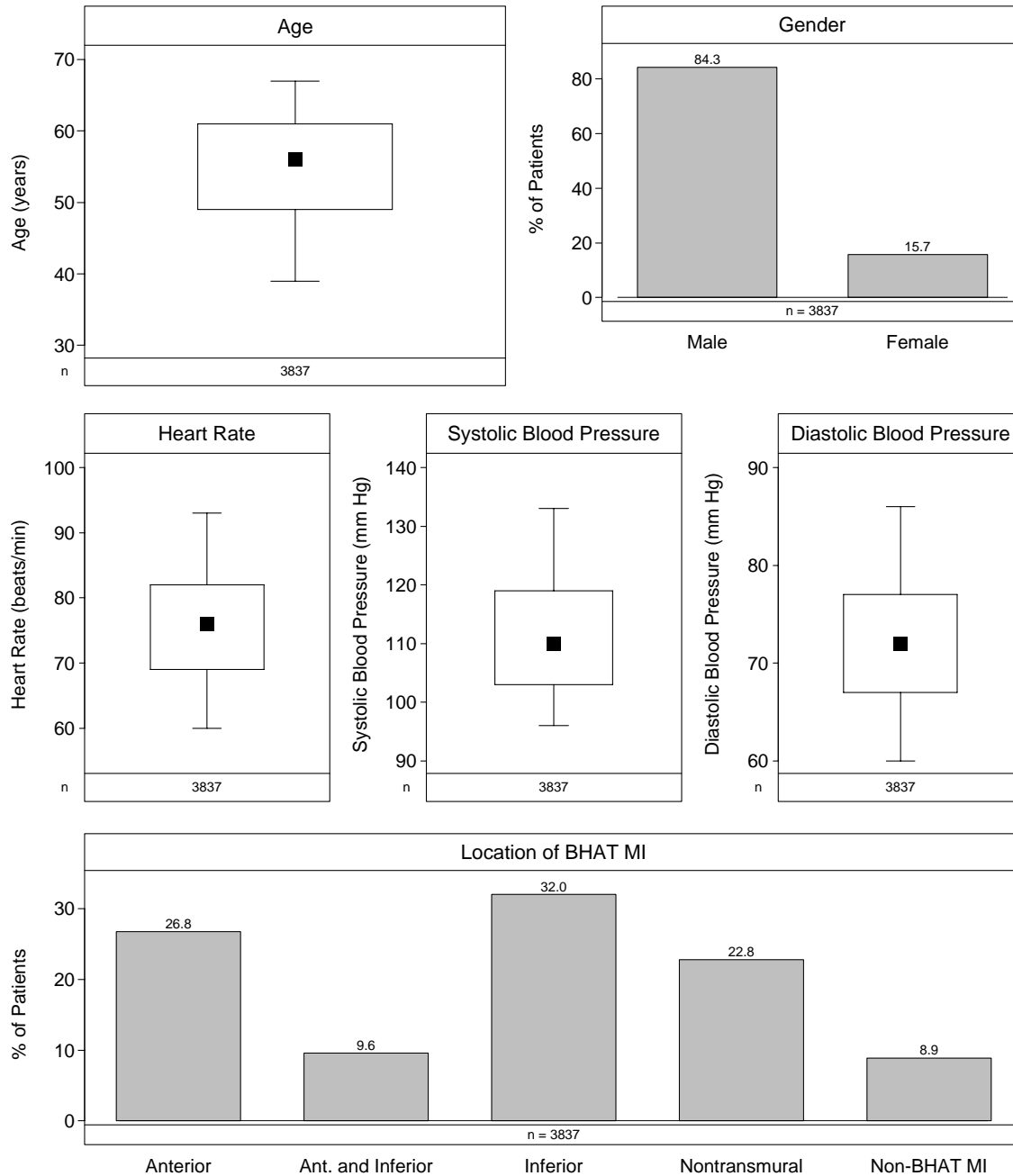
#### Baseline Demographics: Heart Rate

See Figure BASE-1 on page 18.

Trt	Total Pats	Std		Median	Q1	Q3	P5	P95	Contrast	P-Value
		Mean	Dev							
A	1916	76.17	9.80	76.00	70.00	82.00	61.00	93.00	A.B	0.16
B	1921	75.73	9.83	75.00	69.00	82.00	60.00	93.00		

Figure ANCI-1

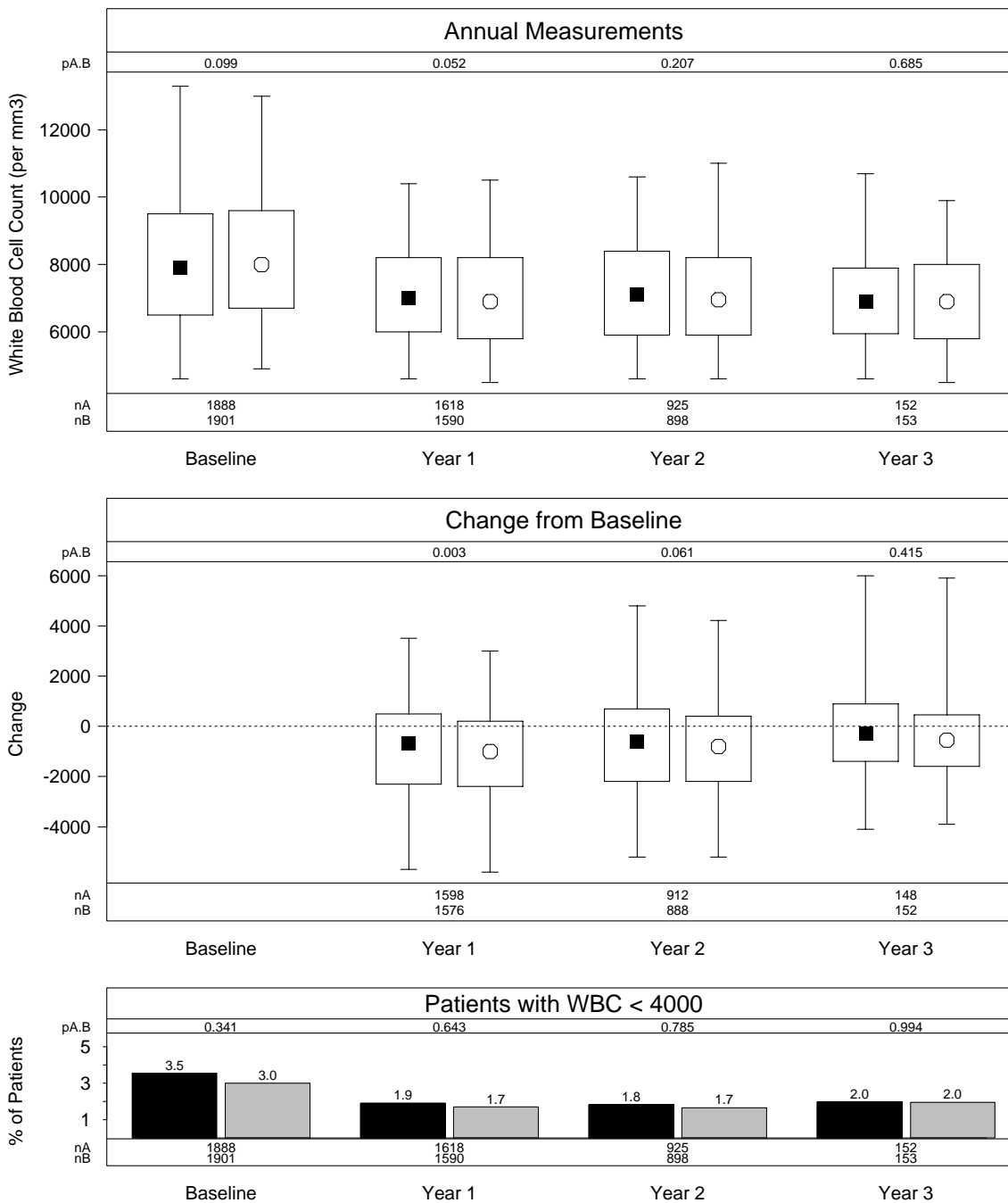
## Baseline Demographics



Information determined by baseline interview, physical examination, and ECG administered prior to randomization. ECGs for the BHAT qualifying MI were initially read at the clinical center, and later confirmed by the Resting ECG Reading Center. See the *Introduction*, page 9, for discussion of "Non-BHAT MI" classification. See also Figure BASE-1 on page 18 for the Closed Session Report version of this page.

Figure FU-2

# White Blood Cell Count



Information from the baseline examination and from visits scheduled after one, two, and three years of follow-up. Panels show measurements at each visit, absolute change from baseline, and percent of patients with white blood cell (WBC) counts < 4000. See also Figure ANCI-2 on page 57 for the Open Session Report version of this page.



See Table Set FU-2 on page 49.

## Example SAS code

Input variable specifications:

```
%inspec(basein, cards);  
cards;  
BASE    IDNO          ID      .  .  .      Patient Identification  
BASE    TRT           STRAT   .  .  .      Treatment  
BASE    MILOCATI     CAT     c  .  LOC    Location of BHAT MI  
BASE    SEX           BIN     c  .  SEX    Gender  
BASE    BPHY         BIN     c  .  YESNO  Prior Hypertension  
BASE    AGE          CTS     .  .  .      Age                      years  
BASE    WEIGHTKG     CTS     .  .  .      Baseline Weight          kilograms  
;  
%einspec;
```

## Analysis convention specifications:

```
/* create output specification */
%outspec(out, cards);
cards;
type cts    cts    univ  wilc    abschg    univ  wilc
type cat    cat    freq  chi
type ord    ord    freq  wilc
type ord+   ord    freq  wilc    sgnchg+   freq  wilc
type ord-   ord    freq  wilc    sgnchg-   freq  wilc
type bin    bin    freq  chi
type bin+   bin    freq  chi    sgnchg+   freq  wilc
type bin-   bin    freq  chi    sgnchg-   freq  wilc
;
%eoutspec;
```

## SAS driver macro call:

```
%driver(  
    in = indata.base,  
        intag = base,  
    inspec = basein,  
        catlab = basecat, chglab = spec.chg,  
    defcon = spec.con, defcgp = spec.cgp,  
    defpop = spec.pop, defblc = spec.blc,  
    outspec = spec.out,  
    out = basedata.base,  
    trtvar = trt  
    );
```

This macro call generates a series of summary output datasets.

## List of SAS output datasets:

*ct	Categorical - Summary Stats	*od	Ordinal - Summary Stats
*ctp	Categorical - P-Values	*odp	Ordinal - P-Values
		*cod	Ord Chg - Summary Stats
*nm	Numeric - Summary Stats	*codp	Ord Chg - P-Values
*nmp	Numeric - P-Values		
*cnm	Num Chg - Summary Stats	*fmt	Dictionary
*cnmp	Num Chg - P-Values	*lvs	Levels for Data
		*lvsc	Levels for Changes
*bn	Binary - Summary Stats		
*bnp	Binary - P-Values		
*cbn	Bin Chg - Summary Stats		
*cbnp	Bin Chg - P-Values		

\* = up to 4 character prefix.

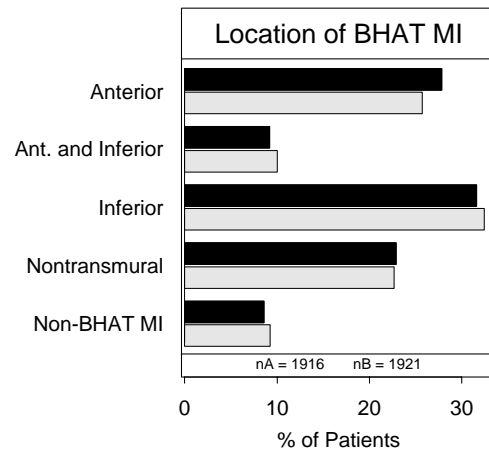
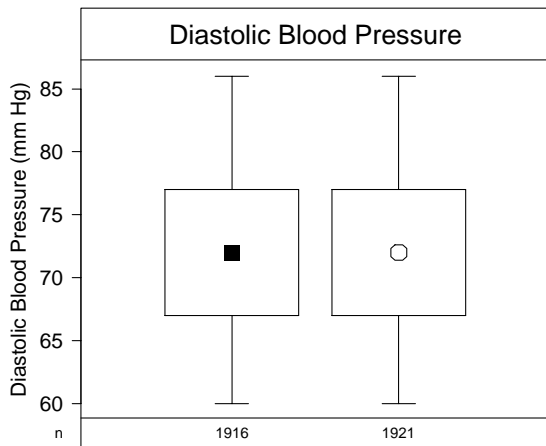
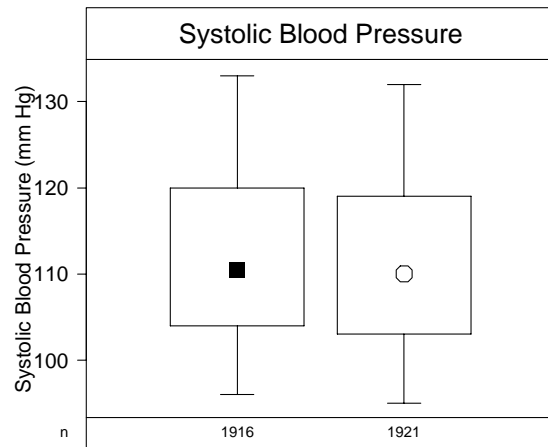
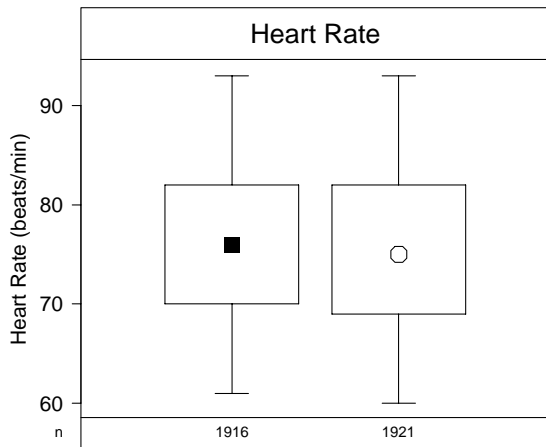
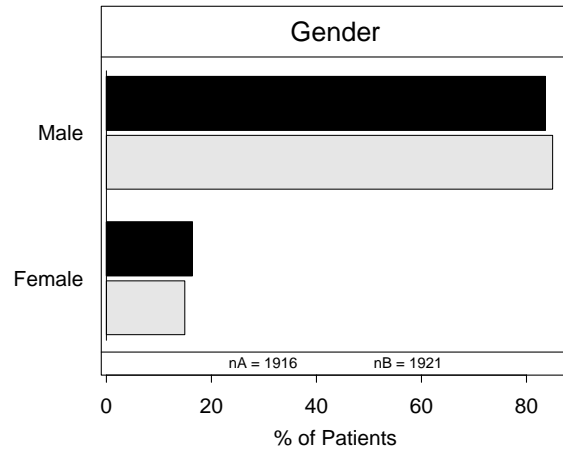
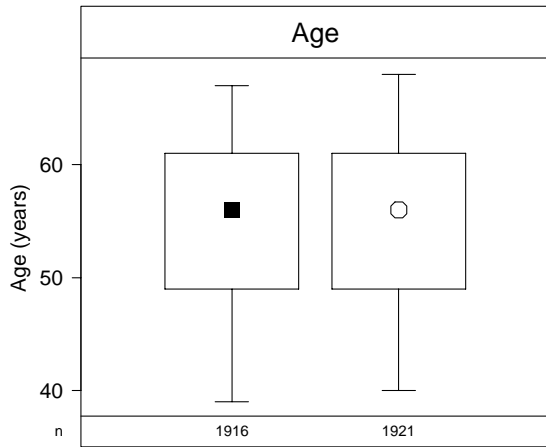
## S-PLUS program for demographics page:

```
select.sas("base", "/u/sdac/bhat/jodi/sasdata")
postscript.start("base")
page.start("base", 1, "Baseline Demographics", "")

plot.page(list("AGE", "SEX", "HRTRT", "BPSYSTO", "BPDIAST", "MILOCATI"),
  "base",
  ttl=c("Baseline Demographics"),
  contrast=T,
  forceList=T,
  horz=F,
  fig=list(NULL, NULL, c(0, .33, .33, .67), c(.33, .67, .33, .67), c(.67, 1, .33, .67),
    c(0, 1, 0, .33)),
  ylim=list(c(30, 70), NULL, c(55, 100), c(90, 140), c(60, 90), NULL),
  mgp=c(NULL, NULL, NULL, c(1.7, .3, 0), NULL, NULL),
  plt.control=list(revNam=F, cext=.75, put.N=T),
  tab.control=list(flip=T))

page.end()
```

# Baseline page using default settings:



Portion of  $\LaTeX$  file containing demographics page:

```
.  
.   
\mpages{accr_accr3}  
  
\DDchapter[base]{Baseline Characteristics}  
  
\DDsection{Demographics}  
\mpages{base_base1}  
  
\DDsection{Medical History}  
\mpages{base_base2}  
  
\DDchapter[ae]{Adverse Events}  
  
\DDsection{Summary of Adverse Events}  
\mpages{ae_ae1}  
.   
.
```

Identical structure used for both main/supporting and open/closed material.