

README file for OVARIAN SURVIVAL PROJECT

Author: Rino Bellocco

Date: November 14, 2006

Table of Content:

1. Introduction
2. Project Members
3. Study Design
4. Description of Folders
5. Potential Problems
6. Status of Project

README file for OVARIAN SURVIVAL PROJECT

Author: Rino Bellocco

Date: November 14, 2006

Section 1: Introduction

I have worked almost full time last week of January and first week of February 2005 to reorganize the Ovarian Cancer Survival Project (OCS). My first comment is that there was no documentation or whatsoever describing the status of the project, or where to start from, in order to understand the flow of the project itself. I had to go through several folders, and several datasets. I have tried to do my best to reorganize this. I thank Mats, Lisa, Tomas, Ingemar, Chantal, and Ann-Sophie for their help.

A major worry I had was that important matching procedures were done by using the **lopnr** coming from different datasets. I was afraid this was not correct, and I will outline this in the section on potential problems.

README file for OVARIAN SURVIVAL PROJECT

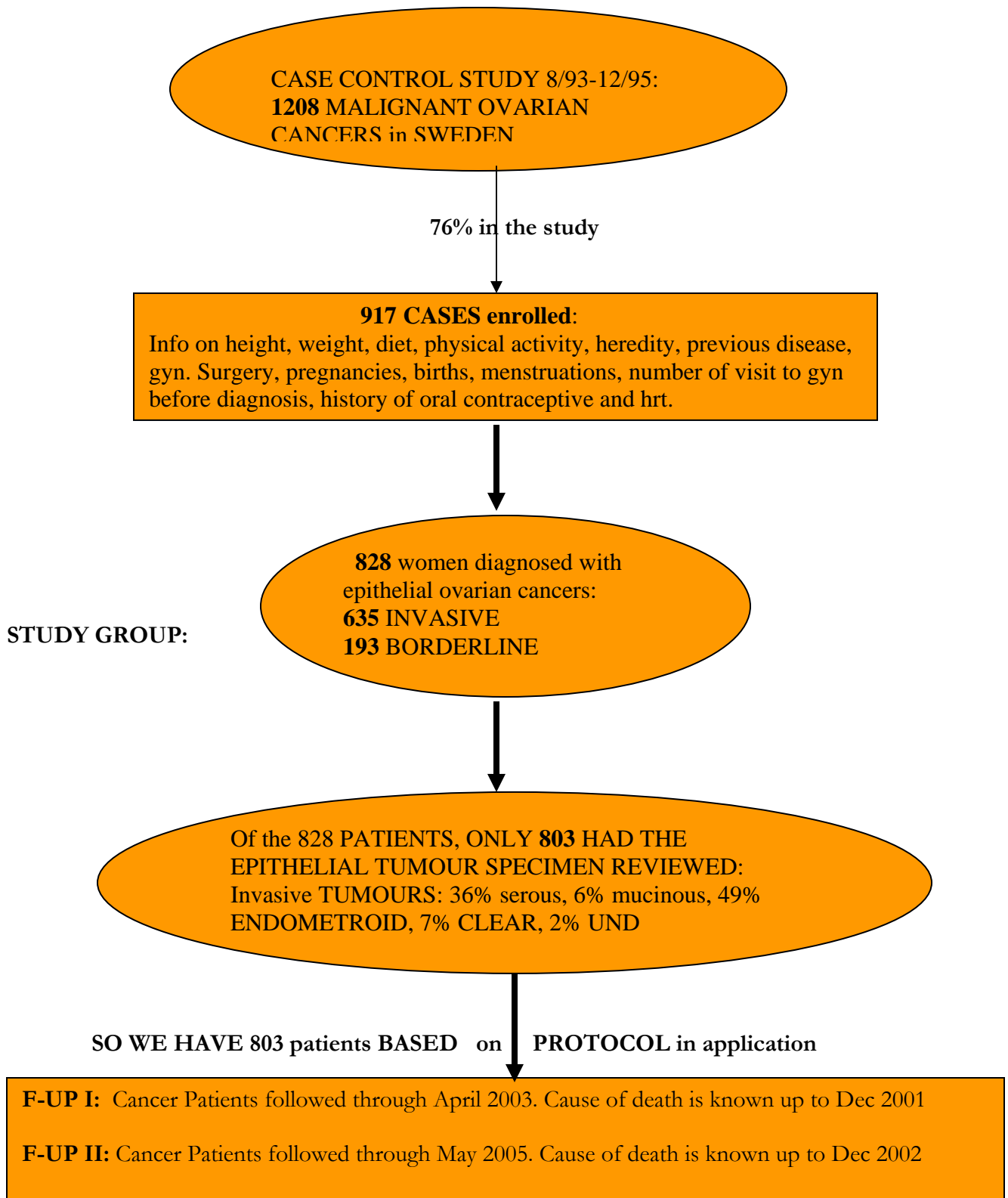
Author: Rino Bellocco

Date: November 14, 2006

Section 2: Project Members

- 1) Rino Bellocco (rino.bellocco@ki.se)
- 2) Kjell Bergfeldt (kjell.bergfeldt@onkpat.ki.se)
- 3) Åsa Klint (asa.klint@ki.se)
- 4) Mats Lambe (mats.lambe@ki.se)
- 5) Ingemar Persson (ingemar.persson@mpa.se)
- 6) Tomas Riman (tomas.riman@ltdalarna.se)
- 7) Lisa Weiderpass (elisabete.weiderpass@ki.se)
- 8) Ling Yang (ling-yang@ki.se)

FROM APPLICATION



README file for OVARIAN SURVIVAL PROJECT

Author: Rino Bellocco

Date: November 14, 2006

FROM DATA SETS

EPCT: tomas.dta

1) FILE HAS 4727 CASES / CONTROLS

2) WE KEEP CASES: 828

3) WE DROP LOPNR=12026

(From minutes August 2003: CM : My Task:

Check the duplicate pnr 193004061309 and delete lopnr 12026.

4) We checked and corrected 5 wrong PNR.

5) LEFT with 827 cases (EPITHELIAL)

EMBLA (MEDICAL RECORDS) data files: embla.dta

1) FILES (corresponding to different sections of QUEST) ARE MERGED INTO A BIG FILE KEEPING ONLY FEW VARIABLES

2) FILE HAS TOTALLY 915 CASES: 800 are eligible for analys, 115 are excluded.

DEATH FILES

We have 4 data sets death files (8796.dta, 9700.dta, 01.dta, 0203.dta), I Have appended them, created a death indicator, as well cause of death. This is considering the status of the patient at the end of follow-up.

EMBLA + OVARIAN: OVARIAN

MERGED FILE: IT STILL has 915 observations

OVARIAN 1: 799 cases: WORKING DATABASE

Selection criteria:

drop if exclude==1 (I keep epithelial, and among epithelial keep women I can review histopat tum specimen)

drop if lopnr==11999 (history of COLON CANCER)

README file for OVARIAN SURVIVAL PROJECT

Author: Rino Bellocco

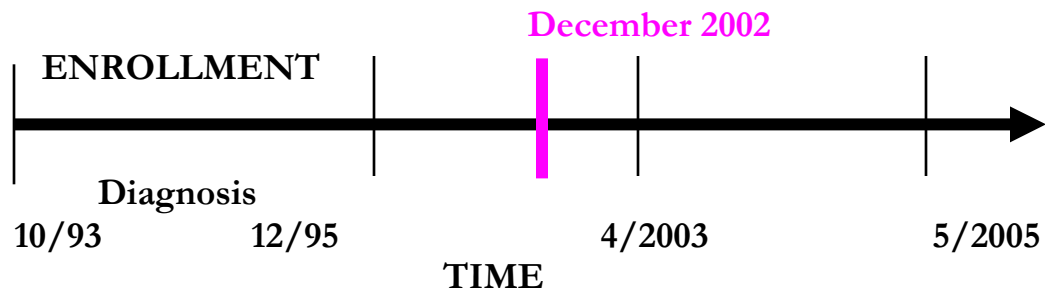
Date: November 14, 2006

There is a discrepancy of 3 cases between the PROTOCOL and the data we analyse. WHY?

README file for OVARIAN SURVIVAL PROJECT

Author: Rino Bellocco

Date: November 14, 2006



README file for OVARIAN SURVIVAL PROJECT

Author: Rino Bellocco

Date: November 14, 2006

Section 4: Description of Folders

The name of the project area on P is OCS. This area has several subdirectories. These are under constant cleaning and reorganization and as of today (14 Nov 2006)

The area has the following structure:

Data:

In here I have the data of the projects: there are 4 subfolders, embla, tomas, merged, death. The working data are in merged.

Data_old:

In here I have put much of the data left by CM as well some of the data I have used. Only for documentation.

Programs:

In here I have all the programs that have generated paper I. This directory will NOT be containing the programs for the other papers.

Programs_Chantal:

In here I have put all the programs left by CM. Only for documentation.

Programs_Rino:

In here I have left all the programs I have generated for the statistical analysis. Again only for documentation.

Results_old:

In here I have put tow folders: the first containing CM results, the second Rino's results. Only for documentation.

Documents:

Important material (data dictionary, manual in eneglish) related to the project are stored in this folder.

README file for OVARIAN SURVIVAL PROJECT

Author: Rino Bellocco

Date: November 14, 2006

Meetings:

Agenda and Minute of Meeting are stored in this folder. Please note that for some reason the dates of the file do not help to sort the files, so use the filename directly.

Literature:

In this folder we will put all papers related to our project

Paper:

In this folder we will put manuscript I

PaperII: All material for Paper II

PaperIII: All material for Paper III

README file for OVARIAN SURVIVAL PROJECT

Author: Rino Bellocco

Date: November 14, 2006

Section 5: Potential Problems

README file for OVARIAN SURVIVAL PROJECT

Author: Rino Bellocco

Date: November 14, 2006

Section 6: Status of Project

- 1. Need to complete manual.v1. bolding variables names being used**
- 2. Need to understand few variables in the QUESTIONNAIRE**
- 3. Paper I accepted**
- 4. Paper II**
- 5. Paper III**