Data for Daffodil project

SDC
November 2016
http://bendixcarstensen.com/daf
Version 2

Compiled Thursday 10th November, 2016, 12:39
from: /home/bendix/sdc/proj/daffodil/sas/programs.tex

Bendix Carstensen Steno Diabetes Center, Gentofte, Denmark
& Department of Biostatistics, University of Copenhagen
bxc@steno.dk b@bxc.dk
http://BendixCarstensen.com

Contents

Chapter 1

Data structure

We shall identify all persons with known diabetes at some point in time. That is in the period 1 January 1995–end, where "end" (currently) is the end of 2015.

Ultimately the Daffodil project will only use persons that are defined as type 2 diabetes (T2D) patients — which will be defined below.

1.0.1 Diabetes definition

Diabetes patients are defined from:

- the Danish National Diabetes Register (NDR) using only the non-blood glucose criteria (period 1995–2012).
- the National Patient Register (NPR) using the relevant ICD codes for diabetes (period 1977–2015).
- the Danish Adult Diabetes Database (DADD) using the annual reports (period 2005–2015)
- the Register of Medicinal Product Statistics (RMPS) using diabetes drug purchase (period 1995–2015).

1.1 Dates

Based on these registers we produce a data set with all persons known as diabetes patients at some point with the following variables:

`id`	person-id
`sex`	sex (M/F)
`typ`	type of diabetes (T1D/T2D)
`doB`	date of birth
`doDM`	date of diabetes diagnosis
`doE`	date of entry to study
`doX`	date of exit from study
`doDD`	date of death

The data set contains one record per person who is known only as T1D or as T2D (defined as *not* T1D — see below); for persons known as both, there will be one record for each period, thus potentially several periods of differing status.

The last date of exit will always be either date of death, date of emigration or 2015-12-31, the latter if the person is not recorded as dead.

1.1.1 Date of diabetes

We derive the date of diabetes as the first of:

- date of inclusion in the NDR, modified to be the earliest of the dates of foot-therapy, NPR-diagnosis or second purchase date of anti-diabetic medication.
- the earliest recorded date of DM diagnosis from the NPR.
- the earliest recorded date of diagnosis in DADD; this includes the earliest status date if it is before the recorded date of diagnosis.
- date of first purchase of anti-diabetic drugs (OAD or insulin) from the RMPS.

This definition means that it is essentially only the date of foot-therapy that is used from the NDR, the other dates used in construction of the NDR are either the same (NPR) or earlier (we use first purchase from RMPS, the NDR is using the 2^{nd}).

1.1.1.1 Gestational diabetes

The NPR contain records of gestational diabetes (GDM), possibly for more than one pregnancy:

- if the date of DM diagnosis is in the first year after a GDM diagnosis, we do not include the persons as DM till 1 year after GDM diagnosis.
- if the date of DM diagnosis of DM is earlier or more than a year after GDM diagnosis, the date of GDM has no effect on the date of DM.

1.1.1.2 PCOS

Patients with PCOS are treated with metformin, and would thus be included as DM patients at the data of purchase of metformin.

Therefore, persons with a NPR-record of PCOS (ICD-8: 61520, 61521; ICD-10: E282) will need special attention:

- if the first date of metformin purchase is before the date of PCOS, the person will be included as DM patient from date of metformin purchase until date of PCOS, exiting DM status at this date.
- If metformin is first purchased within 1 year of PCOS the person will *not* be included as DM on the basis of metformin purchase at all. But possibly on the basis of other criteria.
- If metformin is first purchased more than 1 year after PCOS, the person *will* be included on the basis of metformin purchase.

This means that we must maintain a separate date of first metformin purchase from the RMPS in addition to first purchase of other OADs and insulin.

Moreover this rule has the effect that persons can cease being DM; excluding persons totally based on a future diagnosis of PCOS would introduce immortal bias, since we cannot exclude persons who gets PCOS *after* the follow-up period. This is also one of the reasons that we need a special `doX` variable which might be different from date of death.

1.2 Types of DM

The *type* of diabetes is essentially defined as T1D respectively *not* T1D, where we for convenience label the latter "T2D". Thus we must define T1D, which ideally would be done on an individual basis.

However, we are also using information from the DADD to determine if persons are T1D, which leads to the (formally illogical) phenomenon of people being classified interchangeably as T1D and T2D patients.

Thus persons are classified as T1D as follows:

- If date of DM diagnosis is before the person's 30^{th} birthday, the person is classified as T1D throughout during the entire follow-up regardless of any other criteria.
- If status is recorded as T1D in the DADD the person is assigned T1D status from that date.
- If the first status in DADD is T1D, the person is assigned T1D status also for the time prior to the first status date.
- Persons are re-assigned T2D status if a later recording as T2D patient occur, but not if subsequent status recordings are "Other" or "Unknown".

Technically, this is accommodated by having a separate record for each period where a person is classified as either T1D or T2D. Thus the DADD *may* produce several records per person alternately coded as T1D and T2D. This is the other background for the defined variables `doE` (date of Entry) and `doX` (date of eXit).

1.2.1 Intermittent types of DM?

At first it may seem a bit odd that persons can be T1D and not T1D at different times of their follow-up. However, we are using DADD to identify (and subsequently exclude) T1D patients when recorded as such in the database. The database receives clinical records annually from outpatient clinics (and GPs in due course), including a recording of each patient as T1D or T2D (or "Other" or "Unknown").

But we cannot exclude a patient's follow-up prior to a T1D registration; this would lose the comparability with those T2D patients that in the future will have a recording as T1D patients (which we for good reasons do not know about yet) — we cannot condition on the future. If we retrospectively classified persons as T1D we would selectively be assigning long-term survivors T1D status (namely those who live long enough to hit a T1D diagnosis in DADD).

Neither can we exclude persons indefinitely after a classification as T1D. If the person later is classified as T2D, we must re-introduce this person as T2D because we would

otherwise lose comparability with persons that were classified as T1D before the inception of the DADD and only recorded as T2D after this. The latter point is however of lesser importance because excluding persons indefinitely after at T1D classification would not involve conditioning on the future. Thus, assigning T1D status permanently for any persons seeing a T1D status recording in DADD would be possible.

Hence, the DADD data produces periods where the persons' follow-up is classified T1D patients. This will be converted into records of periods where persons are at risk.

1.3 Drug exposure

We have all purchases of anti-diabetic drugs in Denmark available from the RMPS. For some of the filled prescriptions there is an indication of the dosage (the prescribed daily dose). For the remaining drugs we use the DDD multiplied by a factor as the dosage. Thus for each prescription we will have:

`id`: person-id
`doP`: date of purchase
`drug`: brand of drug
`drgr`: drug group
`dose`: dosage

The variable `drgr` is merely a grouping of the `drug` variable.

By definition any `doP` for a given person (`id`) will be on or after the `doDM` for the person.

The prescription data provide the possibility to measure whether a person is exposed to a given drug at a given date — namely whether the given date is within the computed coverage period derived from the purchase history of the given drug.

Some drug purchases takes place before the expiry of the exposure period for the previous one, and in these cases we must decide if the coverage period from the last purchase is attached to the end of the previous one or if the remaining coverage at the purchase time is discarded. Of course intermediate possibilities exist such as allowing a certain fraction of the remaining or a fixed period of coverage to be carried over across purchase dates.

These considerations are however essentially part of the analysis, the data construction ends at the specification of the coverage period for each prescription.

In the practical analysis we will use the `gen.exp` function from the R-package `Epi`:

`gen.exp` *Generate covariates for drug-exposure follow-up from drug purchase records.*

Description

From records of drug purchase and possibly known treatment intensity, the time since first drug use and cumulative dose at prespecified times is computed. Optionally, lagged exposures are computed too, i.e. cumulative exposure a prespecified time ago.

Usage

```
gen.exp(purchase, id = "id", dop = "dop", amt = "amt", dpt = "dpt",
              fu, doe = "doe", dox = "dox",
          breaks,
```

```
 use.dpt = ( dpt %in% names(purchase) ),
    lags = NULL,
push.max = Inf,
pred.win = Inf,
 lag.dec = 1 )
```

Arguments

`purchase`	Data frame with columns `id`-person id, `dop` - `d`ate `o`f `p`urchase, `amt` - `a`mount purchased, and optionally `dpt` - (`d`ose `p`er `t`ime) ("defined daily dose", DDD, that is), how much is assume to be ingested per unit time. The units used for `dpt` is assumed to be units of `amt` per units of `dop`.
`id`	Character. Name of the id variable in the data frame.
`dop`	Character. Name of the date of purchase variable in the data frame.
`amt`	Character. Name of the amount purchased variable in the data frame.
`dpt`	Character. Name of the dose-per-time variable in the data frame.
`fu`	Data frame with follow-up period for each person, the person id variable must have the same name as in the `purchase` data frame.
`doe`	Character. Name of the date of entry variable.
`dox`	Character. Name of the date of exit variable.
`use.dpt`	Logical: should we use information on dose per time.
`breaks`	Numerical vector of dates at which the time since first exposure, cumulative dose etc. are computed.
`lags`	Numerical vector of lag-times used in computing lagged cumulative doses.
`push.max`	Numerical. How much can purchases maximally be pushed forward in time. See details.
`pred.win`	The length of the window used for constructing the average dose per time used to compute the duration of the last purchase
`lag.dec`	How many decimals to use in the construction of names for the lagged exposure variables

Details

Each purchase record is converted into a time-interval of exposure.

If `use.dpt` is `TRUE` then the dose per time information is used to compute the exposure interval associated with each purchase. Exposure intervals are stacked, that is each interval is put after any previous. This means that the start of exposure to a given purchase can be pushed into the future. The parameter `push.max` indicates the maximally tolerated push. If this is reached by a person, the assumption is that some of the purchased drug is not counted in the exposure calculations.

The `dpt` can either be a constant, basically translating the purchased amount into exposure time the same way for all persons, or it can be a vector with different treatment intensities for each purchase. In any case the cumulative dose is computed taking this into account.

If `use.dpt` is `FALSE` then the exposure from one purchase is assumed to stretch over the time to the next purchase, so we are effectively assuming different rates of dose per time between any two adjacent purchases. Moreover, with this approach, periods of non-exposure does not exist. Formally this approach conditions on the future, because the rate of consumption (the accumulation of cumulative exposure) is computed based on knowledge of when next purchase is made.

The intention of this function is to generate covariates for a particular drug for the entire follow-up of each person. The reason that the follow-up prior to drug purchase and post-exposure is included is that the covariates must be defined for these periods too, in order to be useful for analysis of disease outcomes.

This function is described in terme of calendar time as underlying time scale, because this will normally be the time scale for drug purchases and for entry and exit for persons. In principle the variables termed as dates might equally well refer to say the age scale, but this would then have to be true *both* for the purchase data and the follow-up data.

Value

A data frame with one record per follow-up interval between `breaks`, with columns:

`id` person id.

`dof` date of follow up, i.e. start of interval. Apart from possibly the first interval for each person, this will assume values in the set of the values in `breaks`.

`Y` the length of interval.

`tfi` time from first initiation of drug.

`tfc` time from latest cessation of drug.

`cdur` cumulative time on the drug.

`cdos` cumulative dose.

`ldos` suffixed with one value per element in `lags`, the latter giving the cumulative doses `lags` before `dof`.

Author(s)

Bendix Carstensen, `<bxc@steno.dk>`. The development of this function was supported partly through a grant from the EFSD (European Foundation for the Study of Diabetes), ""

See Also

`Lexis`, `splitLexis`

Examples

```
# Construct a simple data frame of purchases for 3 persons
# The purchase units (in variable dose) correspond to
n <- c( 10, 17, 8 )
dop <- c( 1995.2+cumsum(sample(1:4/10,n[1],replace=TRUE)),
          1997.3+cumsum(sample(1:4/10,n[2],replace=TRUE)),
          1997.3+cumsum(sample(1:4/10,n[3],replace=TRUE)) )
amt <- sample(   1:3/15, sum(n), replace=TRUE )
dpt <- sample( 15:20/25, sum(n), replace=TRUE )
dfr <- data.frame( id = rep(1:3,n),
                   dop,
                   amt = amt,
                   dpt = dpt )
round( dfr, 3 )
# Construct a simple dataframe for follow-up periods for these 3 persons
fu  <- data.frame( id = 1:3,
                   doe = c(1995,1997,1996)+1:3/4,
                   dox = c(2001,2003,2002)+1:3/5 )
round( fu, 3 )
( dpos <- gen.exp( dfr,
                     fu = fu,
                 breaks = seq(1990,2015,0.5),
```

```
                  lags = 2:3/5 ) )
( xpos <- gen.exp( dfr,
                    fu = fu,
               use.dpt = FALSE,
                breaks = seq(1990,2015,0.5),
                  lags = 2:3/5 ) )

# How many relevant columns
nvar <- ncol(xpos)-3
clrs <- rainbow(nvar)

# Show how the variables relate to the follow-up time
par( mfrow=c(3,1), mar=c(3,3,1,1), mgp=c(3,1,0)/1.6, bty="n" )
for( i in unique(xpos$id) )
matplot( xpos[xpos$id==i,"dof"],
         xpos[xpos$id==i,-(1:3)],
         xlim=range(xpos$dof), ylim=range(xpos[,-(1:3)]),
         type="l", lwd=2, lty=1, col=clrs,
         ylab="", xlab="Date of follow-up" )
ytxt <- par("usr")[3:4]
ytxt <- ytxt[1] + (nvar:1)*diff(ytxt)/(nvar+2)
xtxt <- rep( sum(par("usr")[1:2]*c(0.98,0.02)), nvar )
text( xtxt, ytxt, colnames(xpos)[-(1:3)], font=2,
                  col=clrs, cex=1.5, adj=0 )
```

1.4 Rates

1.4.1 Mortality

The mortality in the entire population is available from the data bank of Statistics Denmark as number of deaths and person-years by sex and age and calendar year in 1-year classes. We can derive the number of deaths in the same classification and additionally by diabetes type from our person-dataset. By subtraction of the total across diabetes types we can get the number of deaths and person-years in the non-diabetic population. Alternatively we could construct the mortality rates in the non-T1D, respectively non-T2D population, although these might be considered slightly odd to operate with.

This will enable analysis of mortality among diabetes patients (T1D or T2D) relative to the mortality in the non-DM population. Note that we could easily model the relative mortality among diabetes patients by diabetes duration, drug exposure duration etc. using a combination of the mortality table data for the non-DM population and the individual follow-up records from the T2D population.

1.4.2 CVD

From the NPR we have obtained all diagnoses of CVD (defined as certain ICD-codes). Since we have the id and date of diagnosis of all persons with DM we will be able to classify all CVD events by DM status (none/T1D/T2D). Combining the CVD and the DM data we can compute person years subdivided by the cross-classification by CVD (yes/no)

and DM (none/T1D/T2D), except for the group "noCVD & noDM".

As before, the latter can be obtained by subtraction of the remainder of the persons years (that is, any follow-up after CVD or DM) from the total population person-years obtained from Statistics Denmark, of course classified by sex and age and calendar time in 1-year classes.

Table 1.1: *ICD-codes used for definition of CVD events*

Tabulation of deaths and CVD events by persons' DM and CVD status will enable us to analyze mortality rates by the combination of DM status and CVD status, but also to model the CVD occurrence rates among persons without CVD by DM status.

Note that we are here referring to "CVD" as if it were a single well-defined entity. The above requires that "CVD" be defined as the union of a set of ICD-codes; the date of (first) CVD will then be the date of the earliest of these diagnoses.

That this means that we can do different analyses depending on the codes we choose to consider as being CVD. Note that if we decide to use a subdivision of CVD events in, say, stroke and otherCVD, there is no simple way to combine the two sets of rates, since they necessarily must be defined as rates of stroke resp. otherCVD *without* the other event.

In principle we could take the analyses further, into the realm of multistate models by considering first transitions to CVD1 and CVD2 defined in some way and subsequent transitions to occurrence of the other. This would also require a concurrent tracking of DM occurrence, so we would have to construct a suitable model with states "Well", "CVD1", "CVD2", "T1D", "T2D" and "Dead" and define which transitions were of particular interest.

However this would require a solid clinical underpinning of the defined CVD states.

Chapter 2

SAS programs

The following is a listing of the SAS-programs (that is the `.log` and `.lst` files) used to generate the base datasets. Each is preceded by a very brief description; the main technical points are included as comments in the program code.

2.1 all

This is just a super-program that executes all relevant SAS-programs in the correct order.

```
%inc 'optslibs.sas' ;

* Process all prescriptions to a file of grouped prescriptions
%inc '01_rmps.sas' ;

* Collect all persons from LPR, NDR amd DVDD
%inc '04_lpr_ndr_dvdd.sas' ;

%inc '10_person.sas' ;
```

2.2 optslibs

This is common set of declarative commands that defines a couple of options, the location of the raw and the derived datasets and some global macro variables holding the follow-up period. It could be included as an autoexec file.

```
* options used thoughout ;
options nocenter ovp notes nomprint
        ps = 10000 /*105*/
        ls = 90    /*160*/
       obs = max
  formchar = '           '
 /* formats we use */
 fmtsearch = ( dsfmt.times_personstatistik
               dsfmt.brancher
               dsfmt.uddannelser
               dsfmt.geokoder
               ttfmt.ttformat00
               ttfmt.ttformat10
               ttfmt.ttformat20 ) ;

* data libraries ;
libname DELPOP15 'E:\rawdata\705093\Delpopulationer\' ;
```

```
libname EKST15   'E:\rawdata\705093\Eksterne data\' ;
libname GRUND15  'E:\rawdata\705093\Grunddata\' ;
libname POPUL15  'E:\rawdata\705093\Population\' ;
libname TTDATA  'E:\workdata\705093\BxC\daffodil\DATA\' ;

* format libraries ;
libname TTfmt 'E:\workdata\705093\QSN_MLiJ\' ;
libname DSfmt 'E:\Formater\SAS formater i Danmarks Statistik\FORMATKATALOG';

* useful constants ;
%let primo   = '01JAN1995'd ;
%let cutdate = '01JAN2016'd ; * we have data including 2015 ;
%let antdage = (&cutdate.-&primo.) ;
```

2.3 00-personbase

```
NOTE: Copyright (c) 2002-2012 by SAS Institute Inc., Cary, NC, USA.
NOTE: SAS (r) Proprietary Software 9.4 (TS1M3)
      Licensed to FORSKNING 1, Site 50800722.
NOTE: This session is executing on the X64_SRV12  platform.



NOTE: Updated analytical products:

      SAS/STAT 14.1

NOTE: Additional host information:

 X64_SRV12 WIN 6.2.9200  Server

NOTE: SAS initialization used:
      real time           0.81 seconds
      cpu time            0.59 seconds

1    ************************************************************************
2    * person base ;
3    * original MLiJ, modified BxC, Oct 2016 ;
4    options source2 ;
5    %inc 'optslibs.sas' ;
NOTE: %INCLUDE (level 1) file optslibs.sas is file
      E:\workdata\705093\BXC\daffodil\sas\optslibs.sas.
6   +* options used thoughout ;
7   +options nocenter ovp notes nomprint
8   +        ps = 10000 /* 105 */
9   +        ls = 90    /* 160 */
10  +       obs = max
11  +  formchar = '           '
12  +/* format libraries we use */
13  + fmtsearch = ( dsfmt.times_personstatistik
14  +               dsfmt.brancher
15  +               dsfmt.uddannelser
16  +               dsfmt.geokoder
17  +               ttfmt.ttformat00
18  +               ttfmt.ttformat10
19  +               ttfmt.ttformat20 ) ;
20  +
21  +* data libraries ;
```

```
22  +libname DELPOP15 'E:\rawdata\705093\Delpopulationer\' ;
NOTE: Libref DELPOP15 was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\rawdata\705093\Delpopulationer
23  +libname EKST15   'E:\rawdata\705093\Eksterne data\' ;
NOTE: Libref EKST15 was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\rawdata\705093\Eksterne data
24  +libname GRUND15  'E:\rawdata\705093\Grunddata\' ;
NOTE: Libref GRUND15 was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\rawdata\705093\Grunddata
25  +libname POPUL15  'E:\rawdata\705093\Population\' ;
NOTE: Libref POPUL15 was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\rawdata\705093\Population
26  +libname TTDATA  'E:\workdata\705093\BxC\daffodil\DATA\' ;
NOTE: Libref TTDATA was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\workdata\705093\BxC\daffodil\DATA
27  +
28  +* format libraries ;
29  +libname TTfmt 'E:\workdata\705093\QSN_MLiJ\' ;
NOTE: Libref TTFMT was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\workdata\705093\QSN_MLiJ
30  +libname DSfmt 'E:\Formater\SAS formater i Danmarks Statistik\FORMATKATALOG';
NOTE: Libref DSFMT was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\Formater\SAS formater i Danmarks Statistik\FORMATKATALOG
31  +
32  +* useful constants ;
33  +%let primo   = '01JAN1995'd ;
34  +%let cutdate = '01JAN2016'd ; * we have data including 2015 ;
35  +%let antdage = (&cutdate.-&primo.) ;
36  +
NOTE: %INCLUDE (level 1) ending.
37
38   * collect all persons from the base datasets ;
39   data saml_pop;
40     merge DELPOP15.pop_dvdd         ( in = in1 )
41           DELPOP15.pop_lmdb         ( in = in2 )
42           DELPOP15.pop_lmdb_2015    ( in = in3 )
43           DELPOP15.pop_lpr          ( in = in4 )
44           DELPOP15.pop_lpr_2015     ( in = in5 )
45           DELPOP15.pop_lpruamb      ( in = in6 )
46           DELPOP15.pop_lpruamb_2015 ( in = in7 )
47           DELPOP15.pop_ndr          ( in = in8 ) ;
48     by pnr ;
49     if in1 then dvdd ="dvdd+" ; else dvdd ="dvdd-" ;
50     if in2 then lmdb ="lmdb+" ; else lmdb ="lmdb-" ;
51     if in3 then lmdbx="lmdbx+"; else lmdbx="lmdbx-";
52     if in4 then lpr  ="lpr+"  ; else lpr  ="lpr-"  ;
53     if in5 then lprx ="lprx+" ; else lprx ="lprx-" ;
54     if in6 then lpa  ="lpa+"  ; else lpa  ="lpa-"  ;
55     if in7 then lpax ="lpax+" ; else lpax ="lpax-" ;
56     if in8 then ndr  ="ndr+"  ; else ndr  ="ndr-"  ;
57   run;

NOTE: There were 88780 observations read from the data set DELPOP15.POP_DVDD.
NOTE: There were 410819 observations read from the data set DELPOP15.POP_LMDB.
NOTE: There were 239315 observations read from the data set DELPOP15.POP_LMDB_2015.
```

```
NOTE: There were 109587 observations read from the data set DELPOP15.POP_LPR.
NOTE: There were 1696 observations read from the data set DELPOP15.POP_LPR_2015.
NOTE: There were 7122 observations read from the data set DELPOP15.POP_LPRUAMB.
NOTE: There were 7159 observations read from the data set DELPOP15.POP_LPRUAMB_2015.
NOTE: There were 523770 observations read from the data set DELPOP15.POP_NDR.
NOTE: The data set WORK.SAML_POP has 628999 observations and 11 variables.
NOTE: DATA statement used (Total process time):
      real time           1.60 seconds
      cpu time            0.51 seconds


58
59   * check if there are duplicates ;
60   proc sort  data = saml_pop  out = _null_  nodupkey ;
61     by pnr ;
62   run ;

NOTE: There were 628999 observations read from the data set WORK.SAML_POP.
NOTE: 0 observations with duplicate key values were deleted.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.20 seconds
      cpu time            0.37 seconds


63
64   * the base populations ;
65   %macro getbef ;
66   data bef ;
67      merge %do i = 1986 %to 2016 ;
68            GRUND15.bef&i. ( keep = pnr koen foed_dag )
69            %end ; ;
70      by pnr ;
71   run;
72   %mend ;
73   %getbef ;

NOTE: There were 4636799 observations read from the data set GRUND15.BEF1986.
NOTE: There were 4701265 observations read from the data set GRUND15.BEF1987.
NOTE: There were 4763521 observations read from the data set GRUND15.BEF1988.
NOTE: There were 4824079 observations read from the data set GRUND15.BEF1989.
NOTE: There were 4889891 observations read from the data set GRUND15.BEF1990.
NOTE: There were 4959508 observations read from the data set GRUND15.BEF1991.
NOTE: There were 5033636 observations read from the data set GRUND15.BEF1992.
NOTE: There were 5111448 observations read from the data set GRUND15.BEF1993.
NOTE: There were 5192058 observations read from the data set GRUND15.BEF1994.
NOTE: There were 5210483 observations read from the data set GRUND15.BEF1995.
NOTE: There were 5245145 observations read from the data set GRUND15.BEF1996.
NOTE: There were 5268824 observations read from the data set GRUND15.BEF1997.
NOTE: There were 5288549 observations read from the data set GRUND15.BEF1998.
NOTE: There were 5308437 observations read from the data set GRUND15.BEF1999.
NOTE: There were 5324533 observations read from the data set GRUND15.BEF2000.
NOTE: There were 5344497 observations read from the data set GRUND15.BEF2001.
NOTE: There were 5363038 observations read from the data set GRUND15.BEF2002.
NOTE: There were 5378304 observations read from the data set GRUND15.BEF2003.
NOTE: There were 5391890 observations read from the data set GRUND15.BEF2004.
NOTE: There were 5406633 observations read from the data set GRUND15.BEF2005.
NOTE: There were 5423347 observations read from the data set GRUND15.BEF2006.
NOTE: There were 5447126 observations read from the data set GRUND15.BEF2007.
NOTE: There were 5475791 observations read from the data set GRUND15.BEF2008.
NOTE: There were 5511451 observations read from the data set GRUND15.BEF2009.
NOTE: There were 5534738 observations read from the data set GRUND15.BEF2010.
NOTE: There were 5560628 observations read from the data set GRUND15.BEF2011.
```

```
NOTE: There were 5580516 observations read from the data set GRUND15.BEF2012.
NOTE: There were 5602628 observations read from the data set GRUND15.BEF2013.
NOTE: There were 5627235 observations read from the data set GRUND15.BEF2014.
NOTE: There were 5659715 observations read from the data set GRUND15.BEF2015.
NOTE: There were 5707251 observations read from the data set GRUND15.BEF2016.
NOTE: The data set WORK.BEF has 7456102 observations and 3 variables.
NOTE: DATA statement used (Total process time):
      real time           4:44.61
      cpu time            36.95 seconds


74
75   data dodsaars ;
76     merge GRUND15.dodsaars2001 (in = in1)
77           GRUND15.dodsaasg2013 (in = in13)
78           GRUND15.dodsaasg2014 (in = in14) ;
79     by pnr ;
80     from2001 = in1 ;
81     from2013 = in13 ;
82     from2014 = in14 ;
83     * death date from two different variables
84       those without or early date of death set to start of NPR ;
85     d_dodsdato = max( d_dodsdto, d_dodsdato, "01JAN1977"d /*No small numbers*/ ) ;
86   run ;

WARNING: Multiple lengths were specified for the variable C_DODSMAADE by input data
         set(s). This can cause truncation of data.
NOTE: There were 526222 observations read from the data set GRUND15.DODSAARS2001.
NOTE: There were 130866 observations read from the data set GRUND15.DODSAASG2013.
NOTE: There were 706598 observations read from the data set GRUND15.DODSAASG2014.
NOTE: The data set WORK.DODSAARS has 1232819 observations and 16 variables.
NOTE: DATA statement used (Total process time):
      real time           1.31 seconds
      cpu time            0.54 seconds


87
88   proc contents  data = dodsaars ; run ;

NOTE: PROCEDURE CONTENTS used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


89
90   title1 'Dates of death by origin file' ;
91   proc tabulate  data = dodsaars  missing  noseps ;
92     class d_dodsdato from201 from2013 from2014 ;
ERROR: Variable FROM201 not found.
93     table d_dodsdato, ( all from201*from2013*from2014 ) * f=comma9.
94           / rts = 6 ;
95     keylabel n = " " ;
96     format d_dodsdato year4. ;
97   run ;

NOTE: The SAS System stopped processing this step because of errors.
NOTE: PROCEDURE TABULATE used (Total process time):
      real time           0.00 seconds
      cpu time            0.00 seconds

98   title1 ;
99
```

```
100  * merge population with death records and remove persons not observed
101    between primo and cutdate ;


102  data TTDATA.bef_grund
103       bef ;
104    merge bef ( in = in1 )
105          GRUND15.dod2014
106          GRUND15.dod2015 ;
107    by pnr ;
108    * must be in base population ;
109    if in1 ;
110    * born after end date ;
111    lateborn = ( foed_dag >= &cutdate. ) ;
112    * dead before start date ;
113    earlydead = ( .z < doddato < &primo. ) ;
114    output bef ;
115    * collect persons contributing risk ;
116    if ^lateborn and ^earlydead then output TTDATA.bef_grund ;
117  run;

NOTE: There were 7456102 observations read from the data set WORK.BEF.
NOTE: There were 256274 observations read from the data set GRUND15.DOD2014.
NOTE: There were 1291907 observations read from the data set GRUND15.DOD2015.
NOTE: The data set TTDATA.BEF_GRUND has 7361669 observations and 6 variables.
NOTE: The data set WORK.BEF has 7456102 observations and 6 variables.
NOTE: DATA statement used (Total process time):
      real time           7.78 seconds
      cpu time            3.46 seconds


118
119  title "All basedata --- only lateborn=0 & earlydead=0 contribute" ;
120  proc tabulate  data = bef  missing  noseps ;
121    class koen lateborn earlydead ;
122    table lateborn*earlydead all,
123          ( koen="sex" all ) * f = comma10.
124          / rts=10 ;
125    format koen koen_t. ;
126  run ;

NOTE: There were 7456102 observations read from the data set WORK.BEF.
NOTE: PROCEDURE TABULATE used (Total process time):
      real time           0.67 seconds
      cpu time            0.96 seconds


127
128  title "Base data contributing risk" ;
129  proc contents  data = TTDATA.bef_grund ;
130  run ;

NOTE: PROCEDURE CONTENTS used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


131
132  title "Base data contributing risk combined by populations" ;
133  data a ;
134    merge TTDATA.bef_grund ( in = in1 )
135          saml_pop ;
```

```
136    by pnr ;
137    bef = in1 ;
138    * avoid too small numbers in tables ;
139    foed_dag = max( foed_dag, '1JAN1900'd ) + foed_dag - foed_dag ;
140  run ;

NOTE: Missing values were generated as a result of performing an operation on missing
      values.
      Each place is given by: (Number of times) at (Line):(Column).
      59046 at 139:43
NOTE: There were 7361669 observations read from the data set TTDATA.BEF_GRUND.
NOTE: There were 628999 observations read from the data set WORK.SAML_POP.
NOTE: The data set WORK.A has 7420624 observations and 17 variables.
NOTE: DATA statement used (Total process time):
      real time           4.89 seconds
      cpu time            2.42 seconds


141
142  proc tabulate  data = a  missing  noseps ;
143    class koen foed_dag doddato
144          dvdd lmdb lmdbx lpr lprx lpa lpax ndr bef ;
145    table all doddato foed_dag,
146          ( koen="sex" all ) * f = comma10.
147          / rts=20 ;
148    table dvdd lpr lprx lpa lpax lmdb lmdbx ndr all,
149          ( koen all ) * bef * f = comma10.
150          / rts=8 ;
151    format koen koen_t.
152           foed_dag doddato year4. ;
153  run ;

NOTE: There were 7420624 observations read from the data set WORK.A.
NOTE: PROCEDURE TABULATE used (Total process time):
      real time           2.57 seconds
      cpu time            10.00 seconds
```

00-personbase.lst

Reads the files with all person ids (pnr), corresponding to each calendar year of data.

```
The SAS System                                        14:24 Friday, November 4, 2016   1

The CONTENTS Procedure

Data Set Name        WORK.DODSAARS                 Observations          1232819
Member Type          DATA                          Variables             16
Engine               V9                            Indexes               0
Created              04/11/2016 14:30:25           Observation Length    96
Last Modified        04/11/2016 14:30:25           Deleted Observations  0
Protection                                         Compressed            NO
Data Set Type                                      Sorted                NO
Label
Data Representation  WINDOWS_64
Encoding             wlatin1  Western (Windows)


                    Engine/Host Dependent Information
```

```
Data Set Page Size          65536
Number of Data Set Pages    1811
First Data Page             1
Max Obs per Page            681
Obs in First Data Page      653
Number of Data Set Repairs  0
ExtendObsCounter            YES
Filename                    F:\SASWork\_TD75204_SRVFSESAS5_\dodsaars.sas7bdat
Release Created             9.0401M3
Host Created                X64_SRV12
```

Alphabetic List of Variables and Attributes

#	Variable	Type	Len	Format	Informat	Label
2	C_DOD1	Char	4			C_DOD1
3	C_DOD2	Char	4			C_DOD2
4	C_DOD3	Char	4			C_DOD3
5	C_DOD4	Char	4			C_DOD4
6	C_DODSMAADE	Char	1			Ddsart
8	C_DODTILGRUNDL_ACME	Char	4			C_DODTILGRUNDL_ACME
9	C_DOD_1A	Char	4			C_DOD_1A
10	C_DOD_1B	Char	4			C_DOD_1B
11	C_DOD_1C	Char	4			C_DOD_1C
12	C_DOD_1D	Char	4			C_DOD_1D
13	D_DODSDATO	Num	8	DDMMYY10.		D_DODSDATO
7	D_DODSDTO	Num	8	DDMMYY10.		Ddsdato
14	from2001	Num	8			
15	from2013	Num	8			
16	from2014	Num	8			
1	pnr	Char	12	$12.	$10.	Personnummer

		sex		
		Mand	Kvinde	All
		N	N	N
lateborn	earlydead			
0	0	3,670,704	3,690,965	7,361,669
	1	47,165	47,153	94,318
1	0	61	54	115
All		3,717,930	3,738,172	7,456,102

The CONTENTS Procedure

```
Data Set Name        TTDATA.BEF_GRUND            Observations          7361669
Member Type          DATA                        Variables             6
Engine               V9                          Indexes               0
Created              04/11/2016 14:30:27         Observation Length    56
Last Modified        04/11/2016 14:30:27         Deleted Observations  0
Protection                                       Compressed            NO
Data Set Type                                    Sorted                NO
```

```
Label
Data Representation  WINDOWS_64
Encoding             wlatin1  Western (Windows)


                        Engine/Host Dependent Information

Data Set Page Size          65536
Number of Data Set Pages    6309
First Data Page             1
Max Obs per Page            1167
Obs in First Data Page      1137
Number of Data Set Repairs  0
ExtendObsCounter            YES
Filename                    E:\workdata\705093\BxC\daffodil\DATA\bef_grund.sas7bdat
Release Created             9.0401M3
Host Created                X64_SRV12


                Alphabetic List of Variables and Attributes

#    Variable     Type    Len    Format       Informat    Label

4    DODDATO      Num       8    DDMMYY10.                Dato for ddsfald
2    FOED_DAG     Num       8    DDMMYY10.                Fdselsdato
3    KOEN         Num       8                             Kn
6    earlydead    Num       8
5    lateborn     Num       8
1    pnr          Char     12    $12.         $10.        Personnummer
Base data contributing risk combined by populations     14:24 Friday, November 4, 2016   4


                                  sex

                        .         Mand      Kvinde      All

                        N          N          N          N

 All                    58,955  3,670,704  3,690,965  7,420,624
 Dato for ddsfald
 .                      58,955  3,092,593  3,093,405  6,244,953
 1995                        .     31,068     31,707     62,775
 1996                        .     30,177     30,506     60,683
 1997                        .     29,343     30,209     59,552
 1998                        .     28,830     29,325     58,155
 1999                        .     28,632     30,258     58,890
 2000                        .     28,070     29,578     57,648
 2001                        .     28,228     29,791     58,019
 2002                        .     28,177     30,172     58,349
 2003                        .     28,015     29,331     57,346
 2004                        .     27,407     28,234     55,641
 2005                        .     26,655     28,052     54,707
 2006                        .     27,087     28,188     55,275
 2007                        .     26,894     28,452     55,346
 2008                        .     26,709     27,630     54,339
 2009                        .     26,831     27,843     54,674
 2010                        .     26,640     27,507     54,147
 2011                        .     25,831     26,480     52,311
 2012                        .     25,800     26,316     52,116
 2013                        .     26,013     26,262     52,275
 2014                        .     25,546     25,543     51,089
 2015                        .     26,158     26,176     52,334
```

Fdselsdato				
.	58,955	60	31	59,046
1900	.	1,516	5,350	6,866
1901	.	749	2,396	3,145
1902	.	1,061	3,149	4,210
1903	.	1,342	3,936	5,278
1904	.	1,801	5,024	6,825
1905	.	2,261	5,979	8,240
1906	.	2,904	7,197	10,101
1907	.	3,562	8,430	11,992
1908	.	4,476	10,065	14,541
1909	.	5,470	11,498	16,968
1910	.	6,390	12,634	19,024
1911	.	7,340	13,660	21,000
1912	.	8,575	15,182	23,757
1913	.	9,601	16,146	25,747
1914	.	10,461	17,132	27,593
1915	.	11,024	17,062	28,086
1916	.	12,021	18,087	30,108
1917	.	12,816	18,737	31,553
1918	.	14,338	19,978	34,316
1919	.	14,768	19,853	34,621
1920	.	17,754	23,893	41,647
1921	.	18,100	23,907	42,007
1922	.	18,088	22,694	40,782
1923	.	19,210	24,016	43,226
1924	.	20,180	24,273	44,453
1925	.	20,258	24,240	44,498
1926	.	20,872	24,318	45,190
1927	.	21,005	23,870	44,875
1928	.	22,065	24,479	46,544
1929	.	21,804	24,046	45,850
1930	.	22,638	24,800	47,438
1931	.	22,849	24,548	47,397
1932	.	23,159	25,129	48,288
1933	.	23,645	24,803	48,448
1934	.	24,840	26,029	50,869
1935	.	25,568	26,429	51,997
1936	.	26,763	27,415	54,178
1937	.	27,729	28,456	56,185
1938	.	28,966	28,995	57,961
1939	.	29,612	29,047	58,659
1940	.	30,815	30,507	61,322
1941	.	31,921	31,327	63,248
1942	.	36,035	34,944	70,979
1943	.	38,172	37,368	75,540
1944	.	41,727	40,238	81,965
1945	.	44,110	42,035	86,145
1946	.	45,475	43,301	88,776
1947	.	43,739	42,236	85,975
1948	.	41,179	39,741	80,920
1949	.	39,135	38,019	77,154
1950	.	39,251	38,307	77,558
1951	.	38,528	36,766	75,294
1952	.	39,010	37,636	76,646
1953	.	39,879	38,247	78,126
1954	.	39,207	37,857	77,064
1955	.	40,298	38,205	78,503
1956	.	40,729	38,864	79,593
1957	.	40,714	38,446	79,160
1958	.	40,934	38,781	79,715
1959	.	40,923	38,768	79,691

1960	.	42,738	40,318	83,056
1961	.	42,831	40,425	83,256
1962	.	44,347	41,638	85,985
1963	.	47,096	44,132	91,228
1964	.	48,543	44,966	93,509
1965	.	49,477	46,429	95,906
1966	.	51,357	48,060	99,417
1967	.	48,251	45,380	93,631
1968	.	45,796	42,575	88,371
1969	.	44,347	41,667	86,014
1970	.	44,526	42,250	86,776
1971	.	46,662	43,989	90,651
1972	.	47,065	45,372	92,437
1973	.	45,702	43,760	89,462
1974	.	45,758	44,098	89,856
1975	.	46,398	45,175	91,573
1976	.	43,891	41,827	85,718
1977	.	42,729	41,097	83,826
1978	.	43,043	41,587	84,630
1979	.	42,591	40,981	83,572
1980	.	42,076	40,843	82,919
1981	.	40,195	39,127	79,322
1982	.	40,861	39,253	80,114
1983	.	40,003	39,214	79,217
1984	.	41,013	39,754	80,767
1985	.	41,979	41,036	83,015
1986	.	43,175	41,649	84,824
1987	.	43,078	41,885	84,963
1988	.	43,910	43,168	87,078
1989	.	44,238	43,572	87,810
1990	.	44,466	44,011	88,477
1991	.	43,536	43,037	86,573
1992	.	44,013	42,860	86,873
1993	.	42,346	41,243	83,589
1994	.	42,869	41,585	84,454
1995	.	41,979	40,151	82,130
1996	.	39,501	37,548	77,049
1997	.	38,525	36,520	75,045
1998	.	37,377	35,585	72,962
1999	.	36,892	35,425	72,317
2000	.	37,052	35,419	72,471
2001	.	35,954	34,464	70,418
2002	.	35,485	33,523	69,008
2003	.	35,642	34,012	69,654
2004	.	35,651	34,074	69,725
2005	.	35,475	34,062	69,537
2006	.	36,182	34,169	70,351
2007	.	35,581	33,818	69,399
2008	.	36,224	34,000	70,224
2009	.	34,708	32,896	67,604
2010	.	34,763	33,069	67,832
2011	.	31,959	30,771	62,730
2012	.	31,521	29,646	61,167
2013	.	29,895	28,460	58,355
2014	.	30,024	28,366	58,390
2015	.	29,986	28,518	58,504

Kn

	.	Mand	Kvinde	All	
	bef	bef	bef	bef	
	0	1	1	0	1
	N	N	N	N	N
dvdd					
	.	3,381,274	3,410,351	.	6,791,625
dvdd+	66	52,604	36,110	66	88,714
dvdd-	58,889	236,826	244,504	58,889	481,330
lpr					
	.	3,381,274	3,410,351	.	6,791,625
lpr+	43,203	36,056	30,328	43,203	66,384
lpr-	15,752	253,374	250,286	15,752	503,660
lprx					
	.	3,381,274	3,410,351	.	6,791,625
lprx+	6	1,040	650	6	1,690
lprx-	58,949	288,390	279,964	58,949	568,354
lpa					
	.	3,381,274	3,410,351	.	6,791,625
lpa+	.	4,236	2,886	.	7,122
lpa-	58,955	285,194	277,728	58,955	562,922
lpax					
	.	3,381,274	3,410,351	.	6,791,625
lpax+	.	4,256	2,903	.	7,159
lpax-	58,955	285,174	277,711	58,955	562,885
lmdb					
	.	3,381,274	3,410,351	.	6,791,625
lmdb+	848	218,950	191,021	848	409,971
lmdb-	58,107	70,480	89,593	58,107	160,073
lmdbx					
	.	3,381,274	3,410,351	.	6,791,625
lmdbx+	82	134,455	104,778	82	239,233
lmdbx-	58,873	154,975	175,836	58,873	330,811
ndr					
	.	3,381,274	3,410,351	.	6,791,625
ndr+	27,581	257,773	238,416	27,581	496,189
ndr-	31,374	31,657	42,198	31,374	73,855
All	58,955	3,670,704	3,690,965	58,955	7,361,669

2.4 01-ndr

```
1                                                    "Program 01-ndr.sas"

NOTE: Copyright (c) 2002-2012 by SAS Institute Inc., Cary, NC, USA.
NOTE: SAS (r) Proprietary Software 9.4 (TS1M3)
      Licensed to FORSKNING 1, Site 50800722.
NOTE: This session is executing on the X64_SRV12  platform.


NOTE: Updated analytical products:
```

```
      SAS/STAT 14.1

NOTE: Additional host information:

 X64_SRV12 WIN 6.2.9200  Server

NOTE: SAS initialization used:
      real time           0.07 seconds
      cpu time            0.07 seconds

1          ***********************************************************************
2          * common options and libnames
3          options nosource2 ;
4          %inc 'optslibs.sas' ;
NOTE: Libref DELPOP15 was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\rawdata\705093\Delpopulationer
NOTE: Libref EKST15 was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\rawdata\705093\Eksterne data
NOTE: Libref GRUND15 was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\rawdata\705093\Grunddata
NOTE: Libref POPUL15 was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\rawdata\705093\Population
NOTE: Libref TTDATA was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\workdata\705093\BxC\daffodil\DATA
NOTE: Libref TTFMT was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\workdata\705093\QSN_MLiJ
NOTE: Libref DSFMT was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\Formater\SAS formater i Danmarks Statistik\FORMATKATALOG
36
37         ***********************************************************************
38         * NDR ;
39         proc sort  data = EKST15.ndr  out = ndr ;
40            by pnr ;
41         run ;

NOTE: Input data set is already sorted; it has been copied to the output data set.
NOTE: There were 523770 observations read from the data set EKST15.NDR.
NOTE: The data set WORK.NDR has 523770 observations and 9 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.35 seconds
      cpu time            0.10 seconds


42
43         * Only persons present in the base (TTDATA.bef_grund) ;
44         data TTDATA.NDR ( keep = pnr DoNDR datNDR ) ;
45            merge ndr              ( in = in1
46                                   rename = ( d_inkldto = DoNDR ) )
47                  TTDATA.bef_grund ( in = in2 ) ;
48            by pnr ;
49            if in1 = 1 and in2 = 1 ;
50            * revised date of inclusion (ignoring blood gucose criteria) ;
51            DoNDR = min( d_lpr, d_fodt, d_ins, d_oad ) ;
52            * To avoid small cell entries in overview table ;
53            datNDR = max( doNDR, '01JAN1990'd ) + doNDR - doNDR ;
```

```
54            label datNDR="date of DM left censored 1.1.1977" ;
55         run ;

NOTE: Missing values were generated as a result of performing an operation on missing
      values.
      Each place is given by: (Number of times) at (Line):(Column).
      88845 at 51:12   88845 at 53:40
NOTE: There were 523770 observations read from the data set WORK.NDR.
NOTE: There were 7361669 observations read from the data set TTDATA.BEF_GRUND.
NOTE: The data set TTDATA.NDR has 496189 observations and 3 variables.
NOTE: DATA statement used (Total process time):
      real time           4.29 seconds
      cpu time            1.35 seconds


56
57         title1 'Persons from NDR with valid corrected DM dates' ;
58         title2 'datNDR is the DoNDR left censored at 1.1.1990' ;
59         proc tabulate  data = TTdata.NDR  missing  noseps  formchar='           ' ;
60           class datNDR ;
61           table all datNDR,
62                 n * f=comma10.
63                 / rts = 8 ;
64           format datNDR year4. ;
65           keylabel n = " " ;
66         run ;

NOTE: There were 496189 observations read from the data set TTDATA.NDR.
NOTE: The PROCEDURE TABULATE printed page 1.
NOTE: PROCEDURE TABULATE used (Total process time):
      real time           0.14 seconds
      cpu time            0.18 seconds


67         title2 ;
68
69         proc contents  data = TTdata.ndr ;
70         run ;

NOTE: PROCEDURE CONTENTS used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds

NOTE: The PROCEDURE CONTENTS printed page 2.

NOTE: SAS Institute Inc., SAS Campus Drive, Cary, NC USA 27513-2414
NOTE: The SAS System used:
      real time           5.04 seconds
      cpu time            1.76 seconds
```

01-ndr.lst

Some of the entries in these tables are formatted by `year4.` to conceal small entries (any number between 1 and 365 will be printed as "1960").

```
Persons from NDR with valid corrected DM dates          15:18 Friday, November 4, 2016   1
datNDR is the DoNDR left censored at 1.1.1990


 All       496,189
```

```
date
of DM
left
censo-
red
1.1.1-
977
.          88,845
1990       22,488
1991        9,972
1992        9,422
1993       12,046
1994       30,897
1995       12,992
1996       13,105
1997       12,773
1998       13,991
1999       14,620
2000       14,695
2001       15,125
2002       17,805
2003       19,355
2004       19,595
2005       17,490
2006       17,198
2007       18,458
2008       19,686
2009       20,647
2010       22,157
2011       28,700
2012       24,127
```

```
The CONTENTS Procedure

Data Set Name        TTDATA.NDR                    Observations          496189
Member Type          DATA                          Variables             3
Engine               V9                            Indexes               0
Created              04/11/2016 15:18:20           Observation Length    32
Last Modified        04/11/2016 15:18:20           Deleted Observations  0
Protection                                         Compressed            NO
Data Set Type                                      Sorted                NO
Label
Data Representation  WINDOWS_64
Encoding             wlatin1  Western (Windows)


                     Engine/Host Dependent Information

Data Set Page Size          65536
Number of Data Set Pages    244
First Data Page             1
Max Obs per Page            2039
Obs in First Data Page      1996
Number of Data Set Repairs  0
ExtendObsCounter            YES
Filename                    E:\workdata\705093\BxC\daffodil\DATA\ndr.sas7bdat
Release Created             9.0401M3
Host Created                X64_SRV12
```

```
                    Alphabetic List of Variables and Attributes

#    Variable    Type    Len    Format    Informat    Label

2    DoNDR       Num       8    DATE9.    DATE9.      D_INKLDTO
3    datNDR      Num       8                          date of DM left censored 1.1.1977
1    pnr         Char     12    $12.      $10.        V_CPR
```

2.5 02-dvdd

```
1                                                       "Program 02-dvdd.sas"

NOTE: Copyright (c) 2002-2012 by SAS Institute Inc., Cary, NC, USA.
NOTE: SAS (r) Proprietary Software 9.4 (TS1M3)
      Licensed to FORSKNING 1, Site 50800722.
NOTE: This session is executing on the X64_SRV12  platform.



NOTE: Updated analytical products:

      SAS/STAT 14.1

NOTE: Additional host information:

 X64_SRV12 WIN 6.2.9200  Server

NOTE: SAS initialization used:
      real time           0.12 seconds
      cpu time            0.10 seconds

1          ************************************************************************
2          * common options and libnames
3          options nosource2 ;
4          %inc 'optslibs.sas' ;
NOTE: Libref DELPOP15 was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\rawdata\705093\Delpopulationer
NOTE: Libref EKST15 was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\rawdata\705093\Eksterne data
NOTE: Libref GRUND15 was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\rawdata\705093\Grunddata
NOTE: Libref POPUL15 was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\rawdata\705093\Population
NOTE: Libref TTDATA was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\workdata\705093\BxC\daffodil\DATA
NOTE: Libref TTFMT was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\workdata\705093\QSN_MLiJ
NOTE: Libref DSFMT was successfully assigned as follows:
      Engine:        V9
      Physical Name: E:\Formater\SAS formater i Danmarks Statistik\FORMATKATALOG
36
37         ************************************************************************
38         * DVDD ;
```

```
39         proc sort data = EKST15.DVDD  out = dvdd ;
40           by pnr status_dato diag_dato diag_type ;
41         run ;

NOTE: There were 349662 observations read from the data set EKST15.DVDD.
NOTE: The data set WORK.DVDD has 349662 observations and 53 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           9.09 seconds
      cpu time            0.71 seconds


42
43         * check number of *persons* in the data set ;
44         proc sort data = dvdd  out = pers  nodupkey ;
45           by pnr ;
46         run ;

NOTE: There were 349662 observations read from the data set WORK.DVDD.
NOTE: 260882 observations with duplicate key values were deleted.
NOTE: The data set WORK.PERS has 88780 observations and 53 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.57 seconds
      cpu time            0.51 seconds


47
48         * only persons in base and included before 1.1.2015 ;
49         data dvdd     ;
50           merge dvdd             (in = in1)
51                 TTDATA.bef_grund (in = in2) ;
52           by pnr ;
53           if in1 and in2 ;
54           * remove status records with dates after the cut date ;
55           if status_dato > &cutdate. then delete ;
56         run ;

NOTE: There were 349662 observations read from the data set WORK.DVDD.
NOTE: There were 7361669 observations read from the data set TTDATA.BEF_GRUND.
NOTE: The data set WORK.DVDD has 349557 observations and 58 variables.
NOTE: DATA statement used (Total process time):
      real time           16.85 seconds
      cpu time            1.75 seconds


57
58         * clean out multiple status dates and return a date of diagnosis ;
59         data dvdd                                /* All records */
60              dvdd_fix ( keep = pnr doDVDD ) ; /* one per pnr with revised date of DM
60       ! diagnosis */
61           set dvdd     ( keep = pnr status_dato diag_dato diag_type foed_dag doddato )
61       ! ;
62           by pnr status_dato diag_dato diag_type ;
63           retain first_statusdato doDVDD ;
64           * use only the first among identical status dates within each person ;
65           if first.status_dato ;
66           * use the earliest diagnosis of DM recorded ;
67           if first.pnr then do ;
68              * set the revised DM date to the earlier of diag_dato and status dates ;
69              doDVDD = min( diag_dato, status_dato ) ;
70              * it cannot be before birth, though ;
71              if doDVDD < foed_dag then doDVDD = status_dato ;
72              * earliest status_dato kept for each person ;
```

```
73             first_statusdato = status_dato ;
74             end ;
75           else do ;
76             doDVDD = min( doDVDD, diag_dato, status_dato ) ;
77           end ;
78           * For each DVDD record we compute the difference between the diag_date and
78       ! the
79             computed date doDVDD (months) ;
80           diff = ( diag_dato - doDVDD ) / (365.25) ;
81           output dvdd     ;
82           if last.pnr then output dvdd_fix ;
83         run ;

NOTE: Missing values were generated as a result of performing an operation on missing
      values.
      Each place is given by: (Number of times) at (Line):(Column).
      3910 at 80:22
NOTE: There were 349557 observations read from the data set WORK.DVDD.
NOTE: The data set WORK.DVDD has 347408 observations and 9 variables.
NOTE: The data set WORK.DVDD_FIX has 88292 observations and 2 variables.
NOTE: DATA statement used (Total process time):
      real time           0.54 seconds
      cpu time            0.25 seconds


84
85         * add the computed doDVDD to the staus records ;
86         data dvdd     ;
87           merge dvdd
88                 dvdd_fix ;
89           by pnr ;
90         run ;

NOTE: There were 347408 observations read from the data set WORK.DVDD.
NOTE: There were 88292 observations read from the data set WORK.DVDD_FIX.
NOTE: The data set WORK.DVDD has 347408 observations and 9 variables.
NOTE: DATA statement used (Total process time):
      real time           0.14 seconds
      cpu time            0.07 seconds


91
92         title1 'DVDD records - dates of DM' ;
93         proc tabulate  data = dvdd missing  noseps ;
94           class diag_type diag_dato doDVDD diff ;
95           table all diag_dato doDVDD diff,
96                 diag_type * f=year4. /* comma10. */
97                 / rts = 16 ;
98           format diag_dato doDVDD year4.
99                  diff 5. ;
100        run ;

NOTE: There were 347408 observations read from the data set WORK.DVDD.
NOTE: The PROCEDURE TABULATE printed page 1.
NOTE: PROCEDURE TABULATE used (Total process time):
      real time           0.12 seconds
      cpu time            0.15 seconds


101        title1 ;
102
103        title1 'DVDD persons - dates of DM' ;
```

```
104        proc tabulate  data = dvdd_fix  missing  noseps ;
105          class doDVDD ;
106          table all doDVDD,
107                n * f=year4. /*comma10.*/
108                / rts = 16 ;
109          format doDVDD year4. ;
110        run ;

NOTE: There were 88292 observations read from the data set WORK.DVDD_FIX.
NOTE: The PROCEDURE TABULATE printed page 2.
NOTE: PROCEDURE TABULATE used (Total process time):
      real time           0.03 seconds
      cpu time            0.04 seconds


111        title1 ;
112
113        * DVDD will provide classification of follow-up as T2 / not T2.
114          T2 follow-up is from date of a recording in DVDD with diag_type="T2"
115          until the next recording with diag_type="T1". A re-entry to T2 is then
116          from the next recording of "T2". If unspecified type is in the first
117          status records persons are classified according to the first T1/T2
118          specification, as is the time prior to the first status record ;
119
120        * tabulation of the sequences of changes occurring;
121        proc sort  data = dvdd  out = dvdd_type ;
122          by pnr status_dato ;
123        run ;

NOTE: There were 347408 observations read from the data set WORK.DVDD.
NOTE: The data set WORK.DVDD_TYPE has 347408 observations and 9 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.12 seconds
      cpu time            0.12 seconds


124
125        data dvdd_type ( keep = pnr doDVDD status_dato typ )
126             dvdd_hist ( keep = hist ) ;
127          set dvdd_type ;
128          by pnr ;
129          length typ $ 4  hist $ 80 ;
130          retain hist ;
131          typ = substr( diag_type, 1, 2 ) ;
132          if typ eq "Ty" then typ = "T" || substr( diag_type, 6, 1 ) ;
133          if  first.pnr then hist = typ ;
134          if ^first.pnr and ( diag_type ne lag(diag_type) )
135                        then hist = trim(hist) || " " || typ ;
136          if last.pnr then output dvdd_hist ;
137          output dvdd_type ;
138        run ;

NOTE: There were 347408 observations read from the data set WORK.DVDD_TYPE.
NOTE: The data set WORK.DVDD_TYPE has 347408 observations and 4 variables.
NOTE: The data set WORK.DVDD_HIST has 88714 observations and 1 variables.
NOTE: DATA statement used (Total process time):
      real time           0.15 seconds
      cpu time            0.10 seconds


139
140        title1 'DVDD records - all types of diagnosis histories' ;
```

```
141         proc tabulate data = dvdd_hist  noseps  missing  order = freq ;
142           class hist ;
143           table hist, n * f=year4. /* avoid small numbers instead of comma10. */
144                 / rts=50 ;
145         run ;

NOTE: There were 88714 observations read from the data set WORK.DVDD_HIST.
NOTE: The PROCEDURE TABULATE printed page 3.
NOTE: PROCEDURE TABULATE used (Total process time):
      real time           0.03 seconds
      cpu time            0.04 seconds


146         title1 ;
147
148         * Split follow up by type in intervals ;
149         data TTDATA.dvdd ( keep = pnr doDVDD entry exit typ ) ;
150           set dvdd_type ;
151           by pnr ;
152           retain entry exit ;
153           if first.pnr then do ;
154              * If coded as unknown or other put in T2 class ;
155              if typ ne "T1" then typ = "T2" ;
156              * assume this has been the case from the start ;
157              entry = doDVDD ;
158              * and all the way till the end ;
159               exit = &cutdate. ;
160              * if it is also the last record output ;
161              if last.pnr then output ;
162              end ;
163           if ^first.pnr and
164              typ ne lag(typ) and
165              typ in ("T1","T2") then do ;
166               exit = status_dato ;
167               output ;
168              entry = exit ;
169               exit = &cutdate. ;
170              end ;
171         run ;

NOTE: There were 347408 observations read from the data set WORK.DVDD_TYPE.
NOTE: The data set TTDATA.DVDD has 28409 observations and 5 variables.
NOTE: DATA statement used (Total process time):
      real time           0.18 seconds
      cpu time            0.03 seconds


172

NOTE: SAS Institute Inc., SAS Campus Drive, Cary, NC USA 27513-2414
NOTE: The SAS System used:
      real time           28.12 seconds
      cpu time            4.00 seconds
```

02-dvdd.lst

Some of the entries in these tables are formatted by `year4.` to conceal small entries (numbers between 1 and 365 are printed as "1960").

```
DVDD records - dates of DM                          15:37 Friday, November 4, 2016   1
```

	Diabetestype			
	-In-gen-	And-en type	Type 1	Type 2
	N	N	N	N
All	1960	1982	2334	2514
Diagnosetidspunkt				
.	1960	1961	1961	1967
1890	.	.	.	1960
1899	.	.	.	1960
1900	.	1960	1960	1960
1901	.	.	.	1960
1909	.	1960	.	.
1919	.	.	.	1960
1920	.	.	1960	1960
1923	.	1960	.	.
1927	.	.	1960	.
1928	.	.	.	1960
1931	.	.	.	1960
1933	.	.	1960	.
1934	.	.	1960	1960
1936	.	.	1960	1960
1937	.	.	1960	.
1938	.	.	1960	1960
1939	.	.	1960	1960
1940	.	.	1960	1960
1941	.	.	1960	.
1942	.	.	1960	1960
1943	.	.	1960	.
1944	.	.	1960	.
1945	.	.	1960	.
1946	.	1960	1960	1960
1947	.	.	1960	.
1948	.	.	1960	1960
1949	.	.	1960	.
1950	.	.	1960	1960
1951	.	.	1960	.
1952	.	.	1960	1960
1953	.	1960	1960	1960
1954	.	1960	1960	1960
1955	.	1960	1961	1960
1956	.	.	1961	1960
1957	.	.	1961	1960
1958	.	1960	1961	1960
1959	.	1960	1961	1960
1960	.	1960	1962	1960
1961	.	1960	1962	1960
1962	.	1960	1962	1960
1963	1960	1960	1962	1960
1964	.	1960	1962	1960
1965	.	.	1962	1960
1966	.	1960	1963	1960
1967	.	1960	1963	1960
1968	1960	1960	1963	1960
1969	.	1960	1963	1960
1970	.	1960	1964	1960

```
1971              . 1960 1964 1960
1972              . 1960 1965 1960
1973           1960 1960 1964 1960
1974              . 1960 1965 1960
1975           1960 1960 1965 1961
1976              . 1960 1965 1961
1977           1960 1960 1965 1961
1978              . 1960 1966 1961
1979              . 1960 1966 1961
1980           1960 1960 1967 1964
1981           1960 1960 1966 1962
1982              . 1960 1966 1963
1983           1960 1960 1966 1963
1984           1960 1960 1966 1964
1985           1960 1960 1966 1967
1986           1960 1960 1967 1966
1987           1960 1960 1968 1966
1988           1960 1960 1967 1967
1989           1960 1960 1968 1967
1990           1960 1960 1969 1974
1991           1960 1960 1969 1970
1992           1960 1960 1968 1974
1993           1960 1960 1968 1973
1994           1960 1960 1969 1975
1995           1960 1960 1969 1981
1996           1960 1960 1969 1979
1997           1960 1960 1970 1980
1998           1960 1960 1970 1985
1999           1960 1960 1968 1983
2000           1960 1960 1969 1989
2001           1960 1960 1970 1983
2002           1960 1960 1968 1984
2003           1960 1960 1968 1984
2004           1960 1960 1969 1986
2005           1960 1960 1968 1988
2006           1960 1961 1969 1986
2007           1960 1961 1970 1987
2008           1960 1961 1969 1984
2009              . 1961 1968 1980
2010           1960 1961 1966 1976
2011              . 1961 1963 1972
2012              . 1961 1962 1968
2013           1960 1961 1961 1965
2014              . 1960 1960 1962
2015              . 1960 1960 1960
9999           1960    . 1960 1960
doDVDD
1890              .    .    . 1960
1900           1960 1960 1960 1960
1901              .    .    . 1960
1919              .    .    . 1960
1920              .    . 1960 1960
1923              . 1960    .    .
1931              .    .    . 1960
1933              .    . 1960    .
1934              .    . 1960    .
1936              .    . 1960    .
1937              .    . 1960    .
1938              .    . 1960    .
1939              .    . 1960 1960
1940              .    . 1960 1960
1941              .    . 1960    .
```

```
1942              .    . 1960 1960
1943              .    . 1960    .
1944              .    . 1960    .
1945              .    . 1960 1960
1946              . 1960 1960 1960
1947              .    . 1960 1960
1948              . 1960 1960 1960
1949              . 1960 1960    .
1950              .    . 1960 1960
1951              .    . 1960    .
1952              . 1960 1960 1960
1953              . 1960 1960 1960
1954              . 1960 1960 1960
1955              . 1960 1961 1960
1956              .    . 1961 1960
1957              .    . 1961 1960
1958              . 1960 1961 1960
1959              . 1960 1961 1960
1960              . 1960 1962 1960
1961              . 1960 1962 1960
1962              . 1960 1962 1960
1963           1960 1960 1962 1960
1964              . 1960 1962 1960
1965              . 1960 1962 1960
1966              . 1960 1963 1960
1967              .    . 1963 1960
1968           1960 1960 1963 1960
1969           1960 1960 1963 1960
1970              . 1960 1964 1960
1971              . 1960 1964 1960
1972              . 1960 1965 1960
1973           1960 1960 1964 1960
1974              . 1960 1965 1960
1975           1960 1960 1965 1961
1976           1960 1960 1965 1961
1977           1960 1960 1965 1961
1978              . 1960 1967 1961
1979           1960 1960 1966 1961
1980           1960 1960 1967 1964
1981           1960 1960 1967 1962
1982           1960 1960 1967 1963
1983           1960 1960 1967 1963
1984           1960 1960 1967 1964
1985           1960 1960 1967 1967
1986           1960 1960 1968 1966
1987           1960 1960 1968 1967
1988           1960 1960 1968 1968
1989           1960 1960 1968 1968
1990           1960 1960 1970 1975
1991           1960 1960 1969 1970
1992           1960 1960 1968 1975
1993           1960 1960 1968 1973
1994           1960 1960 1970 1975
1995           1960 1960 1969 1983
1996           1960 1960 1969 1980
1997           1960 1960 1970 1981
1998           1960 1960 1970 1986
1999           1960 1960 1968 1984
2000           1960 1960 1969 1991
2001           1960 1960 1970 1984
2002           1960 1960 1969 1985
2003           1960 1960 1968 1984
```

```
2004            1960 1960 1969 1986
2005            1960 1960 1968 1987
2006            1960 1961 1968 1986
2007            1960 1961 1968 1984
2008            1960 1961 1967 1981
2009               . 1961 1965 1977
2010            1960 1961 1965 1974
2011               . 1961 1963 1971
2012               . 1961 1962 1968
2013            1960 1961 1961 1965
2014               . 1960 1960 1963
2015               . 1960 1960 1960
diff
.               1960 1961 1961 1967
-111               .    .    . 1960
-110               .    . 1960 1960
-109               .    . 1960 1960
-108               .    . 1960 1960
-107               . 1960 1960 1960
-99                . 1960    .    .
-89                .    .    . 1960
-88                .    . 1960 1960
-87                .    . 1960 1960
-86                .    .    . 1960
-85                .    .    . 1960
-84                . 1960    .    .
-81                .    . 1960    .
-60                .    . 1960    .
-56                .    . 1960    .
-54                . 1960    .    .
-53                .    . 1960 1960
-47                .    . 1960    .
-45                .    . 1960    .
-44                .    . 1960    .
-41                .    . 1960    .
-40                .    . 1960    .
-39                .    . 1960    .
-37                .    . 1960    .
-36                .    . 1960    .
-35                .    . 1960    .
-33                .    . 1960    .
-29                .    . 1960    .
-24                .    . 1960    .
-21                .    . 1960    .
-20                .    . 1960    .
-19                .    . 1960    .
0               1960 1980 2311 2473
1                  . 1960 1966 1967
2                  . 1960 1961 1963
3                  . 1960 1960 1962
4                  . 1960 1960 1961
5                  . 1960 1960 1961
6                  . 1960 1960 1961
7                  . 1960 1960 1961
8                  . 1960 1960 1961
9                  . 1960 1960 1961
10                 . 1960 1960 1961
11                 . 1960 1960 1961
12                 . 1960 1960 1961
13                 . 1960 1960 1960
14                 . 1960 1960 1960
15                 . 1960 1960 1960
```

```
16              . 1960 1960 1960
17              . 1960 1960 1960
18              . 1960 1960 1960
19              . 1960 1960 1960
20              . 1960 1960 1960
21              . 1960 1960 1960
22              . 1960 1960 1960
23              . 1960 1960 1960
24              . 1960 1960 1960
25              . 1960 1960 1960
26              . 1960 1960 1960
27              . 1960 1960 1960
28              .    . 1960 1960
29              . 1960 1960 1960
30              . 1960 1960 1960
31              .    . 1960 1960
32              . 1960 1960 1960
33              . 1960 1960 1960
34              .    . 1960 1960
35              .    . 1960 1960
36              .    . 1960 1960
37              .    . 1960 1960
38              . 1960 1960 1960
39              .    . 1960 1960
40              .    . 1960 1960
41              . 1960 1960 1960
42              .    . 1960 1960
43              . 1960 1960 1960
44              .    . 1960 1960
45              .    . 1960 1960
46              .    . 1960    .
47              . 1960 1960 1960
48              .    . 1960 1960
49              .    . 1960    .
50              .    . 1960    .
51              .    . 1960 1960
52              .    . 1960 1960
53              . 1960 1960 1960
54              . 1960 1960    .
55              .    . 1960 1960
56              .    . 1960 1960
57              . 1960 1960 1960
58              .    . 1960    .
59              .    . 1960    .
60              .    . 1960    .
61              .    . 1960 1960
62              .    . 1960 1960
64              .    . 1960    .
67              .    . 1960    .
70              .    .    . 1960
71              .    . 1960 1960
73              .    . 1960    .
75              .    . 1960    .
78              .    . 1960    .
80              .    . 1960 1960
81              .    .    . 1960
88              .    .    . 1960
93              .    .    . 1960
94              .    . 1960    .
96              .    .    . 1960
100             .    . 1960 1960
104             .    .    . 1960
```

```
105               .     . 1960 1960
106               .     .     . 1960
107               .     .     . 1960
108               .     .     . 1960
109               .     . 1960 1960
111               .     .     . 1960
7992              .     . 1960 1960
7993           1960     .     . 1960
7994              .     .     . 1960
8099           1960     .     . 1960
```

```
                N

All            2201
doDVDD
1890           1960
1900           1960
1901           1960
1919           1960
1920           1960
1923           1960
1931           1960
1933           1960
1934           1960
1936           1960
1937           1960
1938           1960
1939           1960
1940           1960
1941           1960
1942           1960
1943           1960
1944           1960
1945           1960
1946           1960
1947           1960
1948           1960
1949           1960
1950           1960
1951           1960
1952           1960
1953           1960
1954           1960
1955           1960
1956           1960
1957           1960
1958           1960
1959           1960
1960           1960
1961           1960
1962           1960
1963           1960
1964           1960
1965           1960
1966           1960
1967           1960
1968           1960
1969           1960
1970           1960
```

```
1971            1960
1972            1961
1973            1960
1974            1961
1975            1961
1976            1961
1977            1961
1978            1961
1979            1961
1980            1962
1981            1961
1982            1961
1983            1961
1984            1962
1985            1962
1986            1962
1987            1962
1988            1963
1989            1963
1990            1965
1991            1963
1992            1965
1993            1964
1994            1965
1995            1967
1996            1966
1997            1966
1998            1968
1999            1967
2000            1970
2001            1968
2002            1968
2003            1969
2004            1969
2005            1970
2006            1971
2007            1970
2008            1971
2009            1969
2010            1969
2011            1968
2012            1966
2013            1966
2014            1964
2015            1960
```

```
                                    N

hist
T2                                 2121
T1                                 2021
An                                 1966
T2 T1                              1963
T1 T2                              1962
An T2                              1961
T2 T1 T2                           1961
T2 An                              1960
T1 T2 T1                           1960
An T1                              1960
```

```
T1 An                                   1960
T2 An T2                                1960
T1 An T1                                1960
T2 T1 T2 T1                             1960
An T2 An                                1960
-I T2                                   1960
-I                                      1960
T1 T2 T1 T2                             1960
-I An                                   1960
-I T1                                   1960
An T2 T1                                1960
T1 T2 An                                1960
An T1 T2                                1960
T2 An T1                                1960
An T1 An                                1960
T2 T1 An                                1960
T2 T1 T2 T1 T2                          1960
T1 An T2                                1960
T2 -I T2                                1960
T1 -I T1                                1960
An T2 An T2                             1960
T2 T1 An T1                             1960
T2 T1 T2 An                             1960
-I An T2                                1960
T1 T2 An T1                             1960
T1 T2 T1 T2 T1                          1960
T1 An T1 An                             1960
T2 T1 T2 T1 T2 T1                       1960
An T1 An T1                             1960
T2 -I                                   1960
An T2 T1 T2                             1960
T1 An T2 T1                             1960
An -I An                                1960
-I T2 T1                                1960
T2 An T2 An                             1960
An T2 An T1                             1960
An T1 T2 T1                             1960
-I T1 T2 T1                             1960
T1 T2 T1 An                             1960
-I An T1                                1960
T1 An T1 An T1                          1960
T2 T1 An T2                             1960
An T2 T1 An                             1960
-I T1 T2                                1960
T2 An T1 T2                             1960
An T1 An T2                             1960
An T2 T1 T2 T1                          1960
T1 An T2 An                             1960
-I T2 T1 T2                             1960
-I An -I An T2                          1960
T1 -I An                                1960
T1 T2 An T2                             1960
An T2 An T2 An                          1960
An T1 T2 An                             1960
T1 -I An T1 T2                          1960
-I An T2 An                             1960
T1 T2 T1 T2 T1 T2 T1 T2                 1960
T1 An T1 T2 An                          1960
T2 An T1 T2 T1 An T1                    1960
T2 -I T1 T2 T1                          1960
An -I An T2 An                          1960
T2 -I An T2 An T2 T1                    1960
```

```
 T1 -I An T2                                    1960
 T1 T2 An T2 An T1                              1960
 T1 T2 T1 T2 T1 T2 T1                           1960
 T2 An T2 T1                                    1960
 -I T2 T1 T2 T1                                 1960
 T2 An T1 T2 T1                                 1960
 -I An T1 An                                    1960
 T2 An T1 An                                    1960
```

2.6 03-npr

Processes the records from the NPR, and computes periods around recorded diagnoses of GDM that should not be included.

2.7 04-rmps

Processes the records from the NPR, and computes periods around recorded diagnoses of GDM that should not be included.

```
* Extracts all prescription records with A10 code
* Constructed MLIJ 12/15-01/16
* Revised BxC Oct 2016

* Udtraf AL A10 medicin. Bruger originale rta fra LMBD data
  Genbrug fra No0010_LMS1_al_medicin.SAS p02448
  Danner grunddata til alle beregninger pntidiabetisk medicin, samt undersgelse af
  hvilken DM-start-dato, som frste indlsning i RMPS bidrager med.

* ansvarlig   : MLIJ
* oprettet    : 01.12.2015
* tilpasset   : 05.01.2016

INDDATA: RMPS data fra ORIGINALDATA
 UDDATA:
E:\workdata\705093\medicin1 medicindata
                    \medicin3 DEN, DER BRUGES I DET VIDERE FORL
E:\workdata\705093\RMPS_start med variable: DoRMPS + Present_RMPS + Drug_user_status
                        bruges til PERSONBASE
 \sum_total   vedr. htering af beregning af DSM
 \sum_total_2 vedr. htering af beregning af DSM
 \subSU       Korrektion pr. 5. januar 2016, hvorefter vi skal kunne htere
              at A10BX02 repeglinide og A10BX03 nateglinid behandles som for selv
              i en gruppe 19 Meglitinides .
              Det fixes ved at lave et sub-dataset, som indeholder
              12 SU excl. meglitinider + 19 meglitinides
              San man tage det overordnede datasTTDATA.medicin3 og fjerne
              gruppe 12, som er den totale SU-gruppe og srstatte den med det nye sub-datas;

source2 ;
%inc 'E:\workdata\705093\BxC\sas\optslibs.sas ;

proc contents data = TTDATA.doso_tekster;   run;
proc contents data = TTDATA.medicin3;       run;
proc contents data = GRUND15.lmdb_pop2009;  run;
proc contents data = GRUND15.lmdb2009;      run;
proc print    data = GRUND15.lmdb2013 (obs=20); run;
```

```
options obs = max;
*********************************************************;
* alle lms data smed al medicin;
data medicin;
     set GRUND15.lmdb1995
     GRUND15.lmdb1996
 GRUND15.lmdb1997
 GRUND15.lmdb1998
 GRUND15.lmdb1999
 GRUND15.lmdb2000
 GRUND15.lmdb2001
 GRUND15.lmdb2002
 GRUND15.lmdb2003
 GRUND15.lmdb2004
 GRUND15.lmdb2005
 GRUND15.lmdb2006
 GRUND15.lmdb2007
 GRUND15.lmdb2008
 GRUND15.lmdb2009
 GRUND15.lmdb2010
 GRUND15.lmdb2011
 GRUND15.lmdb2012
 GRUND15.lmdb2013
 GRUND15.lmdb2014
 ;
where substr(atc,1,3) in('A10');
run;

*************************************************************************************;
**** inddeling i relevante drug groups                                               ;
*************************************************************************************;
data TTDATA.medicin1;
   set medicin;

drugkode = 9999;
      if 1 <= put(atc,$AZdrug5atc.) <= 9990 then do;
         drugk = put(atc,$AZdrug5atc.);
 end;
      drugkode=drugk;

if drugkode = 9999 or
   drugkode = .  then do;
      if 1 <= put(substr(atc,1,5),$AZdrug4atc.) <= 9990 then do;
         drugk = put(substr(atc,1,5),$AZdrug4atc.);
         end;
      drugkode=drugk;
end;

if drugkode = 9999 or
   drugkode = .  then do;
      if 1 <= put(substr(atc,1,4),$AZdrug3atc.) <= 9990 then do;
         drugk = put(substr(atc,1,4),$AZdrug3atc.);
         end;
      drugkode=drugk;
end;

* acarbose slettes ikke alligevel;
* if substr(atc,1,5) = 'A10BF' then delete;
run;
```

```
**  The data set TTDATA.MEDICIN1 has 28395681 observations;
proc contents data = TTDATA.medicin1;
run;

proc format;
value $eksttype
'DD'  = 'Dosisdispens.'
other = 'EJ dosisdispemns.'
;
run;

options obs = max;
data medicin2;
   set TTDATA.medicin1;

eksp_aar = year(eksd);

* antal DDD er antal pakker gange antal DDD'er i en pakke;
quantityDDD = apk * volume;

* antal stk er antal pakker gange den numeriske pakningsstrrelse;
quantitySTK = apk * packsize;

* antal mg er antal pakker gange den numeriske pakningsstrrelse gange styrken af hver stk;
quantityMG = 0;
  if strunit = 'MG'   or
     strunit = 'MGD'  or
 strunit = 'MGG'  or
 strunit = 'MGM'
     then do;
     quantityMG = apk * packsize * strnum;
  end;
* antal mg er antal pakker gange den numeriske pakningsstrrelse gange styrken af hver stk -
  if strunit = 'RG'   or
     strunit = 'RGD'  or
 strunit = 'RGH'  or
 strunit = 'RGT'
     then do;
     quantityMG = apk * packsize * (strnum / 1000);
  end;

** htering af om det er dosisdispenseret medicin;
eksp_type = put(ekst,$eksttype.);

** htering af missing quantityDDD p10BH01 Sitagliptin ca. 60 obs.;
if atc = 'A10BH01' and (quantityDDD = 0 or
                        quantityDDD = .    )then do;
   quantityDDD = quantityMG / 100;
   ** 'A10BH01' =  100;
   ** quantityDDD = quantityMG / (input( put( atcgrup , $DDDvalueAZ.) , 8.));
   end;

run;

proc sort data = medicin2;
   by pnr;
run;


** Der srges for at der kun er records med personer, der findes i BEF;
data medicin2_1;
   merge medicin2            (in=in1)
```

```
         TTDATA.bef_grund    (in=in2 keep = pnr)
         ;
         by pnr;

if in1=1 and in2=1;
run;

** af hensyn til pftning af overselse af doseringskoder;
proc sort data = medicin2_1;
   by doso;
run;

*********************************************************;
** sortering af dataset med overselse af doseringskoder;
proc sort data=TTDATA.doso_tekster;
     by doso;
run;

/*
ods rtf file='E:\workdata\705093\QSN_MLIJ\OUTPUT\DOSO_koder_tekster.rtf';
proc print data = TTDATA.doso_tekster;
run;
ods rtf close;
run;
*/
/*
** print af indikationskoder;
ods rtf file='E:\workdata\705093\QSN_MLIJ\OUTPUT\INDO_koder_tekster.rtf';
proc print data = INDODOSO.indo_tekster_fra_SSI;
run;
ods rtf close;
run;
*/

** pftning af overselse af doseringskoder;
data TTDATA.medicin3;
   merge medicin2_1          (in=in1)
         TTDATA.doso_tekster (in=in2)
         ;
         by doso;

if in1=1;

** denne division med 1000 skal checkes en ekstra gang; ** Den SKAL IKKE bruges;
** dagdosis_ny = dagdosis/1000;

** kun DOSO-koder, med DOSOGRUPPE = '01' skal bruges - '00' er malplacerede - '99' er blank
dosogruppe = put(doso,$dosogrp.);
if (doso = '0000000' or
    doso = ' '          ) and dosogruppe = '00' then dosogruppe = '99';

dagdosisgrp = 0;
   if dagdosis = . then dagdosisgrp = 1;
   if dagdosis = 0 then dagdosisgrp = 2;
   if dagdosis > 0 then dagdosisgrp = 3;

** VARIGHED af den udleverede mde;
** kbt mde (prim i antal styk) divideret med dgnmde (i samme enhed) beregnet ud fra lns dos
   giver antal dage medicinen br re til;
** virker kun fra og med 2004, da DOSO blev introduceret 1. april 2004;
** DOSO-koden skal v tablet-orienteret og den skal give en positiv vi;
** Doseringskoder, der ikke giver mening i forhold til tabletbehandling bliver ignoreret;
```

```
if dagdosis > 0 and dosogruppe = '01' then do;
   varighed = (packsize * apk) / dagdosis;
   end;

** sig beregning for transaktioner med DD (dosisdisensering), men ej for insulin og ej for
*************************************************************************************
if ekst = 'DD' and drugkode ne 15 and drugkode ne 17 then do;

   ** Kategorisering af hvilken Bekendtgrelse transaktionen hrer til under;
   bek = 0;
   if                  eksd <= '30sep2001'd then bek = 1;
   if '01oct2001'd <= eksd <= '16feb2003'd then bek = 2;
   if '17feb2003'd <= eksd                 then bek = 3;

   ** den oprindeligt beregnede varighed gemmes i OPR_VARIGHED;
   opr_varighed = varighed;

   ** kun de obs., hvor der er en gyldig beregnet varighed over 0 dage;
   if opr_varighed > 0 then do;

      ** der rundes op til et helt antal dage i varighed;
      varighed = ceil(opr_varighed);

      ** i BEK 837 fra 2001 mer kun dispenseres til 14 dage ad gangen;
      if bek = 2 and opr_varighed > 14 then varighed = 14;
      ** i BEK 80 fra 2003 mer max dispenseres til 28 dage ad gangen;
      if bek = 3 and opr_varighed > 28 then varighed = 28;
   end;

   ** kun de obs., hvor mangler en gyldig beregnet varighed over 0 dage;
   if opr_varighed = 0 or
      opr_varighed = . then do;

      varighed = 14;
   end;

end; ** slut posis dispensering;
run;


*proc print data = TTDATA.medicin3 (obs=10);
*run;

*************************************************************************************
*************************************************************************************
****  Korrektion pr. 5. januar 2016, hvorefter vi skal kunne htere at A10BX02 repeglinide o
                                                                        A10BX03 nateglinid
      i en gruppe 19 Meglitinides .
      Det fixes ved at lave et sub-dataset, som indeholder 12 SU excl. meglitinider
                                                           +
                                                           19 meglitinides
      San man tage det overordnede datasTTDATA.medicin3 og fjerne gruppe 12, som er den tot
      den med det nye sub-datas
*************************************************************************************
*************************************************************************************
data TTDATA.subSU;
   set TTDATA.medicin3;

where drugkode = 12;

if atc in('A10BX02' , 'A10BX03') then do;
   drugkode = 19;
```

```
   drugk    ='19';
   end;

run;
*******************************************************************************************
*****  SLUT PSU korrektion
*******************************************************************************************




*******************************************************************************************;
** overblik over Acarbose
** 12.8.2015: MLiJ bryder sig ikke om at Acarbose negligeres, for det giver et
   forkert billede af de patienter, som udover acarbose ogsandre drugs.
   Det er i 7.134 personer i lbet af 1955-2009 (ca. 3%) - dog beskedent - men RODET!!!
   Efterflgende blev det besluttet at medtage Acarbose som en 8. gruppe!!!!!!!!!!!!!

   Det er i 7.460 personer i lbet af 1955-2014 med i alt 112.711 recepter, dvs. 15,1 recept
*******************************************************************************************;

data acarbose;
   set TTDATA.medicin3 (keep = pnr eksd atc drugkode drugk eksp_aar);
   where substr(atc,1,5) in('A10BF');
run;

proc sort data = acarbose;
  by pnr eksd;
run;

data acarbose;
   set acarbose;
   by pnr eksd;

if first.pnr;
run;

options obs = max;
ods rtf file='E:\workdata\705093\QSN_MLIJ\OUTPUT\acarbose.rtf';
proc tabulate data = acarbose missing;
class drugkode atc eksp_aar
  ;
*var volume
    ;
table (all='Total' drugkode   ='Drug-gruppe'      *
       (all='Total' atc        ='ATCkode'          ))
   ,
   (all='Total' eksp_aar='Year') * N=' ' * f=comma18.0
       /
   Box = _page_
       rts = 80;
   ;
format drugkode antidiagrouptxt.
       ;
title1    'acarbose.rtf - Overview: Antal patienter og deres initiering af Acarbose i 1995-
run;
title1    ' ';
run;
ods rtf close;
run;

*******************************************************************************************;
```

```
** SLUT p
** overblik over Acarbose
*************************************************************************************;

options obs = max;
ods rtf file='E:\workdata\705093\QSN_MLIJ\OUTPUT\total_RMPS.rtf';
proc tabulate data = TTDATA.medicin3 missing;
class drugkode atc eksp_aar
  ;
*var volume
    ;
table (all='Total' drugkode   ='Drug-gruppe'      *
       (all='Total' atc        ='ATCkode'          ))
   ,
   (all='Total' eksp_aar='Year') * N=' ' * f=comma18.0
       /
   Box = _page_
       rts = 80;
   ;
format drugkode antidiagrouptxt.
       ;
title1    'total_RMPS.rtf - Overview: al indlst medicin 1995-2014 - dataset: TTDATA.medicin
run;
title1    ' ';
run;
ods rtf close;
run;


*************************************************************************************;
*************************************************************************************;
*** OMfanget af patienter  -  og der laves dataset til PERSONBASE med DoRMPS          ;
*************************************************************************************;
*************************************************************************************;
proc sort data = TTDATA.medicin3;
   by pnr eksd;
run;


** For de frste 20 i RMPS 1995-2014 er der reistreret 410.817  P-numre.
   Nder reduceres til DK - residents, sbliver det til 409.958  P-numre.  For 1995-2009 var
data RMPS_start;
   set TTDATA.medicin3 (keep = pnr eksd atc drugkode drugk eksp_aar);
   by pnr eksd;
   if first.pnr;
run;


*****************************************;
*****************************************;
**  dette dataset skal bruges i PERSONBASE ;
*****************************************;
*****************************************;
data TTDATA.RMPS_start (keep = pnr DoRMPS Present_RMPS Drug_user_status_1y
                                                       Drug_user_status_2y);
   set RMPS_start;
   by pnr eksd;

** datoen for frste optrn i RMPS;
DoRMPS = eksd;

** Marker for om man findes i RMPS;
```

```
Present_RMPS = 1;

** Category of blood glucose-lowering drug user status - 2 S REGEL - passer til FINLAND;
                                                        *********************************
   1 = incident user
   2 = prevalent user
   3 = non user;
if '01jan1995'd <= DoRMPS <  '01jan1997'd then Drug_user_status_2y = 2; * prevalent user;
if                 DoRMPS >= '01jan1997'd then Drug_user_status_2y = 1; * incident user;

** Category of blood glucose-lowering drug user status - 1 S REGEL - passer til SVERIGE;
                                                        *********************************
   1 = incident user
   2 = prevalent user
   3 = non user;
if '01jan1995'd <= DoRMPS <  '01jan1996'd then Drug_user_status_1y = 2; * prevalent user;
if                 DoRMPS >= '01jan1996'd then Drug_user_status_1y = 1; * incident user;

run;


options obs = max;
ods rtf file='E:\workdata\705093\QSN_MLIJ\OUTPUT\entry_RMPS.rtf';
proc tabulate data = RMPS_start missing;
class drugkode atc eksp_aar
  ;
*var volume
    ;
table (all='Total' drugkode   ='Drug-gruppe'      *
       (all='Total' atc        ='ATCkode'          ))
   ,
   (all='Total' eksp_aar='Year') * N=' ' * f=comma18.0
       /
   Box = _page_
       rts = 80;
   ;
format drugkode antidiagrouptxt.
       ;
title1    'entry_RMPS.rtf - Overview: Antal patienter og deres initieringsdrug i lbet af No
run;
title1    ' ';
run;
ods rtf close;
run;
*******************************************************************************************;
*******************************************************************************************;
*** Slut p
*** OMfanget af patienter;
*******************************************************************************************;
*******************************************************************************************;


*******************************************************************************************;
*******************************************************************************************;
***  Gennemgribende gennemgang af DOSO-koderne
*******************************************************************************************;
*******************************************************************************************;
options obs = max;


** TTDATA.TEST3 has 6.502.458 records- dvs en record med dagdosis > 0;
data TTDATA.test3;
```

```
   set TTDATA.medicin3;

if doso = ' ' then delete;
if dagdosis > 0 ;
run;

options obs=max;
data TTDATA.test4;
   set TTDATA.test3;

dosogruppe = put(doso,$dosogrp.);
grup = 1;
run;

*******************************************************************************
*** her kommer en re test af DOSO - som volder voldsomme problemer
*******************************************************************************;
***  TABEL 1 ;
options obs = max;
ods rtf file='E:\workdata\705093\QSN_MLIJ\OUTPUT\DOSO_koder.rtf';
proc tabulate data = TTDATA.test3 missing;
class drugkode atc doso dosering_kort dosering_total ekst
  ;
*var volume
    ;
table (drugkode='Drug-gruppe' * atc='ATC' * dosering_kort='DOSO-tekst' * dosering_total='DO
      ,
      ekst='Eksp-type' * N='Antal ' * f=comma15.0
       /
   Box = _page_
       rts = 80;
   ;
format drugkode antidiagrouptxt.
       ;
title1    'DOSO_koder - Overview: DRUGS 1995-2014 - dataset: TTDATA.test3 fra TTDATA.medici
run;
ods rtf close;
run;



***  TABEL 2 ;
options obs = max;
ods rtf file='E:\workdata\705093\QSN_MLIJ\OUTPUT\DOSO_koder2.rtf';
proc tabulate data = TTDATA.test3 missing;
class drugkode atc doso dosering_kort dosering_total ekst aktiv_inaktiv dosogruppe
  ;
*var volume
    ;
table (drugkode='Drug-gruppe' * dosogruppe ='0/1'          * aktiv_inaktiv='A/I' *
       DOSO='DOSO-koder'      * dosering_kort='DOSO-tekst' * dosering_total='DOSO-tekst')
      ,
      ekst='Eksp-type' * N='Antal ' * f=comma15.0
       /
   Box = _page_
       rts = 80;
   ;
format drugkode antidiagrouptxt.
       ;
title1    'DOSO_koder2 - Overview: DRUGS 1995-2014 - dataset: TTDATA.test3 fra TTDATA.medic
run;
ods rtf close;
```

```
run;


options obs = max;
ods rtf file='E:\workdata\705093\QSN_MLIJ\OUTPUT\DOSO_koder3.rtf';
proc tabulate data = TTDATA.test3 missing;
class drugkode atc doso dosering_kort dosering_total ekst aktiv_inaktiv dosogruppe
  ;
*var volume
    ;
table (all='Total' dosogruppe ='0/1' * (all='sub-total' (aktiv_inaktiv='A/I' *
       DOSO='DOSO-koder' * dosering_kort='DOSO-tekst' *  dosering_total='DOSO-tekst')))
      ,
      (all='Total' ekst='Eksp-type') * N='Antal ' * f=comma15.0
       /
   Box = _page_
       rts = 80;
   ;
format drugkode antidiagrouptxt.
       ;
title1    'DOSO_koder3 - Overview: DRUGS 1995-2014 - dataset: TTDATA.test3 fra TTDATA.medic
run;
ods rtf close;
run;

proc sort data = TTDATA.medicin3;
   by pnr eksd;
run;

options obs = max;
ods rtf file='E:\workdata\705093\QSN_MLIJ\OUTPUT\druggrupper.rtf';
proc tabulate data = TTDATA.medicin3 missing;
class drugkode atc eksp_aar
  ;
*var volume
    ;
table (all='Total' drugkode   ='Drug-gruppe'      *
       atc        ='ATCkode'          )
   ,
   all='Total' eksp_aar='Year' * N=' ' * f=comma15.0
       /
   Box = _page_
       rts = 80;
   ;
format drugkode antidiagrouptxt.
       ;
title1    'druggrupper - Overview: DRUGS 1995-2014 - dataset: TTDATA.medicin3. prg: No00000
run;
ods rtf close;
run;
*******************************************************************************;
*******************************************************************************;
***  Slut p
***  Gennemgribende gennemgang af DOSO-koderne
*******************************************************************************;
*******************************************************************************;

*******************************************************************************;
*******************************************************************************;
***************   DD Dosis Dispenseret ****************************************;
*******************************************************************************;
*******************************************************************************;
```

```
** The data set TTDATA.test85 has   588.216 observations for 1995-2009           ;

**              TTDATA.MEDICIN3 har 1.539.562 observations med ekst = 'DD'
   The data set TTDATA.TEST85   has 1.539.553 observations for 1995-2014 , nman smider 9 ob
data test85;
set TTDATA.medicin3 (keep = pnr atc drugkode drugk
                            eksd eksp_aar ekst bek rinr
                            quantityDDD quantityMG quantitySTK
                            apk volume packsize doso dagdosis dosogruppe
varighed opr_varighed
                            strnum strunit );

where ekst in('DD');

if substr(atc,1,4) = 'A10A' then delete;

ny_varig = varighed;
if  0 < varighed < 14 then ny_varig = 1.5;
if 14 < varighed < 28 then ny_varig = 17.5;
if 28 < varighed      then ny_varig = 30.5;


run;

* 25.1.2016: rettet i tabellerne pga af for fbs. i cellerne;
ods rtf file='E:\workdata\705093\QSN_MLIJ\OUTPUT\DosisDiso_DOSO_1_3.rtf';
/* bliver for detaljeret
proc tabulate data = TTDATA.test85 missing;
class varighed dagdosis bek eksp_aar rinr atc drugkode;
table all='I alt' varighed='DSM'
      ,
  all='I alt ' bek='Bek-version' * eksp_aar='eksp-aar' * N='Antal' * f=comma10.0
      /
   Box = _page_
      rts = 80;
   ;
format drugkode antidiagrouptxt.;
title 'prg: No000002_RMPS_1.SAS - OUTPUT\DosisDiso_DOSO_1_1.rtf - Dosisdisp. trans: beregne
run;

title '';
proc tabulate data = test85 missing;
class varighed dagdosis bek eksp_aar rinr atc drugkode;
table all='I alt' drugkode='Gruppe' * (all='Subtotal' (atc='ATC' * (varighed='DSM' all='ATC
      ,
  all='I alt ' bek='Bek-version' * eksp_aar='eksp-aar' * N='Antal' * f=comma10.0
      /
   Box = _page_
      rts = 80;
   ;
format drugkode antidiagrouptxt.;
title 'prg: No000002_RMPS_1.SAS - OUTPUT\DosisDiso_DOSO_1_2.rtf - Dosisdisp. trans: ATC X b
run;
*/
title '';
proc tabulate data = test85 missing;
class varighed dagdosis bek eksp_aar rinr atc drugkode dosogruppe;
table all='I alt' dosogruppe='Brugbare DOSOkoder'
      ,
      all='I alt' drugkode='Gruppe'
      ,
  all='I alt ' /*bek='Bek-version' * */ eksp_aar='eksp-aar' * N='Antal' * f=comma12.0
```

```
      /
   Box = _page_
      rts = 80;
   ;
format drugkode antidiagrouptxt.;
title 'prg: No000002_RMPS_1.SAS - OUTPUT\DosisDiso_DOSO_1.rtf - Dosisdisp. trans: DOSO x AT
run;
title '';

ods rtf close;
run;


proc sort data = TTDATA.test85;
*  by atc pnr eksd;
  by descending dagdosis;
run;

proc print data = TTDATA.test85 (obs=100);
run;

*************************************************************************************;
** test86 + test88 + test89 er summen af alle med DD som ekspeditionstype = TTDATA.test85;
*************************************************************************************;
** test86   --     195.362 obs. har tablet-dosokode og en DAGDOSIS og en varighed, der er f
                   546.639 obs  -""-  for 1995-2014
;
data test86;
   set TTDATA.test85;
where opr_varighed ne 14;
if dagdosis   = 0 or
   dagdosis   = . then delete;
if dosogruppe = '00' then delete; * nonsens DOSOkoder';
run;

ods rtf file='E:\workdata\705093\QSN_MLIJ\OUTPUT\DosisDiso_test86.rtf';
proc tabulate data = test86 missing;
class varighed dagdosis bek opr_varighed;
table dagdosis='dagdosis' * opr_varighed='Varighed'
      ,
  bek='Bek-version' * N='Antal' * f=comma10.0
      /
   Box = _page_
      rts = 80;
   ;
format drugkode antidiagrouptxt.;
run;
ods rtf close;
run;

** test af de oprindeligt beregnede varigheder - inden jeg stopper 14/28 dages algoritmen p
   problemet er at ene masse varigheder er med en masse decimaler;
proc sort data = test86;
   by opr_varighed;
run;

data test87;
   set test86;
   by opr_varighed;

if first.opr_varighed;
run;
```

```
proc print data = test87;
run;


** test88   --     3.612 obs varighed = 14 dage (og dermed en dagdosis)
                180.191  -""-  for 1995-2014
;
data test88;
   set TTDATA.test85;
where opr_varighed   = 14   and
      dosogruppe     = '01'   ;
run;

** test89   --     389.242 obs har helt manglende DOSO kode blank, tom eller nonsens
            --     812.922 obs   -""-  for 1995-2014
;
data test89;
   set TTDATA.test85;

if dagdosis   = 0    or
   dagdosis   = .    or
   dosogruppe = '00'  ;

*where dagdosis in(.,0);
run;

***************   SLUT                 ****************************************;
*******************************************************************************;
***************   DD Dosis Dispenseret ****************************************;
*******************************************************************************;
*******************************************************************************;

*******************************************************************************;
*******************************************************************************;
***************   Diverse test         ****************************************;
*******************************************************************************;
*******************************************************************************;

** check af om der er udregnet antal DDD'er for alle trans.;
** OK 29.7.2015 p995-2009 dataset = 0 records;

** OK 7.12.2015 p995-2014 dataset = 0 records;
*************************************************************;
data testDDDberegn;
   set TTDATA.medicin3 (keep = atc vnr quantityDDD eksd ekst);

where quantityDDD in(0, .);
run;

** check af hvor mange trans., der er gennemfrt efter indfrsel af DOSOkoden;
** OK 29.7.2015 p995-2014 dataset =  8.523.793 records efter 1.4.2004

** OK 7.12.2015 p995-2014 dataset = 18.787.185 records efter 1.4.2004
;
*************************************************************************;
data testDOSO;
   set TTDATA.medicin3 (keep = eksd drugkode drugk atc indo);

if eksd >= '01apr2004'd;
run;
```

```
options obs = max;
ods rtf file='E:\workdata\705093\QSN_MLIJ\OUTPUT\DosisDiso_testDOSO.rtf';
proc tabulate data = testDOSO missing;
class drugkode indo;
table all='Total' indo='Indikationskode'
      ,
  (all='Total' drugkode='Drug gruppe') * N='Antal obs.' * f=comma18.0
  /
  Box = _page_
      rts = 80;
  ;
format drugkode antidiagrouptxt.;
run;
ods rtf close;
run;


*******************************************************************************;
*******************************************************************************;

** test af Metf. + Saxagliptin, som har henh. 1 og 2 tabletter som DDD-value afhig af varen
** kan frst checkes, ndata for efter 2009 bliver leveret
*************************************************************************************
data testMetSaxa;
   set TTDATA.medicin3 (keep = vnr atc volume apk strnum strunit packsize);

where atc = 'A10BD10';
run;

proc sort data = testMetSaxa;
  by vnr;
run;

data testMetSaxa;;
   set testMetSaxa;;
   by vnr packsize strnum strunit volume;
if first.volume;
run;

proc print data = testMetSaxa;;
run;
*************************************************************************************



*************************************************************************************
*************************************************************************************
** htering af DOSO koder generelt
*************************************************************************************
*************************************************************************************
proc contents data = TTDATA.medicin3;
run;

** fordeling af samtlige records efter DOSO-gruppe og DAGDOSIS-vi.
   De records som har en valid DOSO-kode (doso-gruppe = '01')og en dagdosis > 0 skal unders
ods rtf file='E:\workdata\705093\QSN_MLIJ\OUTPUT\Dagdosis_gruppering.rtf';
proc tabulate data = TTDATA.medicin3 missing;
class dosogruppe drugkode atc doso dosering_kort dosering_total ekst varighed dagdosisgrp
  ;
*var volume
    ;
table (all='Total' dosogruppe='Dos.gruppe' * dagdosisgrp='Gruppering af dagdosis')
```

```
      ,
      (all='Total' drugkode='Drug-gruppe' * atc='ATC')
      ,
      all='Total' ekst='Eksp-type' * N='Antal ' * f=comma15.0
       /
   Box = _page_
       rts = 80;
   ;
format drugkode    antidiagrouptxt.
       dosogruppe  $dosotxt.
   dagdosisgrp dosisgrptxt.
       ;
title1    'Dagdosis_gruppering - Overview: DRUGS 1995-2014 - dataset: TTDATA.test3. prg: No
run;
title1    ' ';
run;
ods rtf close;
run;


** test af de DAGDOSIS vier, som er knyttet til valid DOSO-koder - der ses bort fra DD-eksp
** det er 3,070,733 - 198,974 = 2.871.759  (1995-2009)

** WORK.TEST_DOSO1 has 5.607.009 observations (1995-2014)
*********************************************************************************************
data test_doso1;
   set TTDATA.medicin3;

where dosogruppe = '01' and dagdosis > 0;
if ekst = 'DD' then delete;
if substr(atc,1,4)='A10A' then delete;
** Acarbose slettes ikke alligevel;
* if substr(atc,1,5)='A10BF' then delete;

DSM = ceil(varighed);
run;


** Fordeling af dagdosis-vier;
ods rtf file='E:\workdata\705093\QSN_MLIJ\OUTPUT\Dagdosis_vaerdier.rtf';
proc tabulate data = test_doso1 missing;
class dosogruppe drugkode atc dagdosis doso dosering_kort dosering_total ekst varighed dagd
  ;
*var volume
    ;
table (all='Total' DSM='DSM' * dagdosis='Dagdosis')
      ,
      (all='Total' drugkode='Drug-gruppe' * atc='ATC') * N='Antal ' * f=comma15.0
       /
   Box = _page_
       rts = 80;
   ;
format drugkode    antidiagrouptxt.
       dosogruppe  $dosotxt.
   dagdosisgrp dosisgrptxt.
       ;
title1    'Dagdosis_vaerdier - Overview: DRUGS 1995-2014 - dataset: TTDATA.test3. prg: No00
run;
title1    ' ';
run;
ods rtf close;
run;
```

```
*************************************************************************
*** TEST AF DEL HELT HE DSM vier
*************************************************************************;
data hoj_DSM;
   set test_doso1;
if DSM > 365;
run;

proc sort data = hoj_DSM;
   by drugkode atc DSM strnum packsize volume dagdosis dosering_total;
run;


data hoj_DSM;
   set hoj_DSM;
   by drugkode atc DSM strnum packsize volume dagdosis dosering_total;

   if first.dosering_total then antal = 1;
   if first.dosering_total ne 1 then do; antal + 1; end;
run;


data hoj_DSM;
   set hoj_DSM;
   by drugkode atc DSM strnum packsize volume dagdosis dosering_total;

   if last.dosering_total;
run;

/*
** denne liste skal lige afstemmes MED KLINIKERNE.
proc print data = hoj_DSM ;*(obs=10);
var eksd atc dsm apk vnr quantityDDD packsize volume dagdosis dosering_total strnum antal;
run;
*/
/*
data test_0255560;
   set TTDATA.medicin3 (keep= eksd atc varighed apk vnr packsize quantityDDD volume dagdosi
where vnr in('025560');
run;
proc print data = test_0255560 (obs=100);
run;
*/
*************************************************************************
*** TEST AF DEL HELT HE DSM vier
*************************************************************************;

*************************************************************************
*** UNDERSELSE AF FREKVENS AF DE FORSKELLIGE DSM beregnede vier
*************************************************************************;
proc sort data = test_doso1;
   by DSM;
run;
proc means data = test_doso1 noprint;
   by  DSM;
   var apk;
   output out = TTDATA.sum_total
          sum = apk;
run;
```

```
proc print data = TTDATA.sum_total (obs=10);
run;

data TTDATA.sum_total_2;
   set TTDATA.sum_total;

antal = _freq_;
run;
proc print data = TTDATA.sum_total_2 (obs=10);
run;

proc gchart data = TTDATA.sum_total_2;
vbar DSM / sumvar = antal;
run;
quit;
*************************************************************************
*** SLUT
*** UNDERSELSE AF FREKVENS AF DE FORSKELLIGE DSM beregnede vier
*************************************************************************;



proc contents data = TTDATA.medicin3;
run;


data test;
   set TTDATA.medicin3 (keep = drugkode eksd varighed);

for_efter = 0;
if eksd >= '01apr2004'd then for_efter = 1;

var = 0;
if varighed > 0 and
   varighed ne . then var = 1;

run;

** Oversigt over hvor mange recepter, som der kan beregnes DEM p
ods rtf file='E:\workdata\705093\QSN_MLIJ\OUTPUT\Andel_med_DEM_26jan2016.rtf';
proc tabulate data = test missing;
class drugkode for_efter var
  ;
*var volume
    ;
table  (all='Total' drugkode='Drug-gruppe')
       ,
       (all='Total' for_efter='Fr/efter 1.4.2004' * var='Med/uden DEM') * N='Antal ' * f=co
       /
   Box = _page_
       rts = 80;
   ;
format drugkode    antidiagrouptxt.
       ;
title1    'Overview: DRUGS 1995-2014 - beregninger af DEM - Dataset: test3 Prg: No000002_RM
run;
title1    ' ';
run;
ods rtf close;
run;
```