

Caregiver burden data preparation

Hans, Jenny, and Derek

May 14, 2017

Overview

Our analyses focus on caregiver burden as the *outcome* from a variety of predictors, such as demographics, disorder type, cognitive function, and neuropsychological profiles.

Load data and construct data set

The observations we work with require (1) at least summary scores from Zarit's (caregiver) burden interview (ZBI; CITE) and (2) must not classify as normal or missing in final diagnosis (FINDX). Conditional to the ZBI and FINDX, we include participants who qualify as:

- Dementia (community and institution; cgcc==1 | cgcc==3),
- Alzheimer's (adcc==1),
- Parkinson's disease (parkin==1), and
- Stroke (prstroke==1).

Not included for now (because there are so few amongst this set):

- Amyotrophic lateral sclerosis (als==1),
- Epilepsy (epilepsy==1), and
- Multiple sclerosis (ms==1).

Other conditions to consider later (as pseudo-controls):

- Depression (depressn==1),
- Learning Disability (learning==1),
- Psychiatric Illness (psychiat==1), and/or
- Migraines (migraine==1).

The following lines of code will read in the data and include only the subjects outlined in the above conditions.

```
## data exist one directory above and in a separate
## folder (in part to avoid committing it to Github for
## now!)
CSHA.1991.full <- read.csv("../IPN2017_Case_Comp/Dataset.csv")

## Conditionals to extract specific participants start
## with primary outcome of interest: caregiver burden
CSHA.1991.zarit <- CSHA.1991.full[CSHA.1991.full$zarscore <
  99, ]
dim(CSHA.1991.zarit)

## [1] 1086 1724

### Now get all data for above conditions
conditions.of.interest <- c((CSHA.1991.zarit$cgcc == 1 |
  CSHA.1991.zarit$cgcc == 3 | CSHA.1991.zarit$adcc ==
  1 | CSHA.1991.zarit$parkin == 1 | CSHA.1991.zarit$prstroke ==
  1) & (CSHA.1991.zarit$finaldx != 0 & CSHA.1991.zarit$finaldx !=
```

```

9))
CSHA.1991.zarit_disorders <- CSHA.1991.zarit[conditions.of.interest,
]
dim(CSHA.1991.zarit_disorders)

```

```
## [1] 851 1724
```

We now have a subset of observations we should make a pseudo-design matrix of the conditionals above. We want to denote which of the aforementioned categories each individual belongs to:

```

cols.for.design <- c("CASEID", "cgcc", "adcc", "parkin",
"prstroke", "finaldx")
CSHA.1991.zarit_disorders_design <- CSHA.1991.zarit_disorders[,
cols.for.design]

dim(CSHA.1991.zarit_disorders_design)

```

```
## [1] 851 6
```

```
head(CSHA.1991.zarit_disorders_design)
```

```
##      CASEID cgcc adcc parkin prstroke finaldx
## 1  200004    3    1     2         8         2
## 2  200005    3    0     2         1         3
## 4  200008    1    0     2         2         3
## 11 200032    3    0     2         2         3
## 26 200086    3    0     8         8         3
## 35 200113    3    0     1         9         5
```

We can recode some of the values in these columns to something more sensible with a design matrix. The code below to perform the recoding is hidden from the output.

```
head(CSHA.1991.zarit_disorders_design)
```

```
##      CASEID  DEMENTIA  ALZ  PD  STROKE  FIN.DX
## 200004 200004  INS.CASE CASE  NO      DNK  PROB.ALZ
## 200005 200005  INS.CASE N/A  NO      YES  POSS.ALZ
## 200008 200008  COMM.CASE N/A  NO      NO   POSS.ALZ
## 200032 200032  INS.CASE N/A  NO      NO   POSS.ALZ
## 200086 200086  INS.CASE N/A  DNK      DNK  POSS.ALZ
## 200113 200113  INS.CASE N/A  YES  MISSING OTHER.DEM
```

```
head(CSHA.design)
```

```
##      DEMENTIA.INS.CASE DEMENTIA.COMM.CASE ALZ.CASE ALZ.N/A ALZ.MISSING
## 200004                1                  0          1         0          0
## 200005                1                  0          0         1          0
## 200008                0                  1          0         1          0
## 200032                1                  0          0         1          0
## 200086                1                  0          0         1          0
## 200113                1                  0          0         1          0
##      PD.NO PD.DNK PD.YES PD.MISSING STROKE.DNK STROKE.YES STROKE.NO
## 200004    1     0     0          0          1          0          0
## 200005    1     0     0          0          0          1          0
## 200008    1     0     0          0          0          0          1
## 200032    1     0     0          0          0          0          1
## 200086    0     1     0          0          1          0          0
## 200113    0     0     1          0          0          0          0
```

```
##          STROKE.MISSING FIN.DX.PROB.ALZ FIN.DX.POSS.ALZ FIN.DX.OTHER.DEM
## 200004          0          1          0          0
## 200005          0          0          1          0
## 200008          0          0          1          0
## 200032          0          0          1          0
## 200086          0          0          1          0
## 200113          1          0          0          1
##          FIN.DX.VDEM FIN.DX.UNCLASS.DEM
## 200004          0          0
## 200005          0          0
## 200008          0          0
## 200032          0          0
## 200086          0          0
## 200113          0          0
```

```
colSums(CSHA.design)
```

```
## DEMENTIA.INS.CASE DEMENTIA.COMM.CASE          ALZ.CASE
##          539          312          314
##          ALZ.N/A          ALZ.MISSING          PD.NO
##          473          64          612
##          PD.DNK          PD.YES          PD.MISSING
##          27          47          165
##          STROKE.DNK          STROKE.YES          STROKE.NO
##          45          233          432
##          STROKE.MISSING FIN.DX.PROB.ALZ FIN.DX.POSS.ALZ
##          141          346          215
##          FIN.DX.OTHER.DEM          FIN.DX.VDEM FIN.DX.UNCLASS.DEM
##          53          168          69
```

We now have a pseudo-design matrix to identify observation disorder classifications. Now we want to extract just particular sets of columns as predictors and outcomes from our subset.

```
## predictors demographics
demographics <- c("studysex", "agestrat", "studyage", "clinage",
  "studied", "eduyear", "edulevel", "region", "race",
  "ethnic1", "ethnic2", "wbocc", "working", "institut")
### caregiver demographics
cg.demographics <- c("relat", "cgsex", "cgage", "cgeth1",
  "cgeth2", "cgedyrs", "cgedlev")
### cognition
cognition <- c("langabil", "mmms", "score3ms")
### behavioral disturbance
beh.disturb <- c("apathy", "wander", "violenc", "disinhi",
  "otherbe")

### full neuropsych neuropsych batteries
neuropsych <- c("neurdone", "languagn", "doi", "occup",
  "materlan", "difficul", "othrcomn", "psyid", "clinic",
  "training", "experien", "coopern", "facility", "fatigabi",
  "inattn", "affectn", "articul", "tension", "appearn",
  "reaction", "effort", "express", "memoryn", "restless",
  "insight", "gaitn", "agitat", "persever", "impulsiv",
  "socialn", "tangent", "comprehe", "confusi", "latency",
  "rulev", "tolerate", "visionn", "hearingn", "physical",
```

```

"gaिटdiso", "tremdiso", "dyskdiso", "psycdiso", "slowdiso",
"hearaid", "glasses", "wheel", "wechsler", "buschfr1",
"buschcr1", "buschtr1", "buschfr2", "buschcr2", "buschtr2",
"buschfr3", "buschcr3", "buschtr3", "buschfr", "buschcr",
"buschtr", "delaytim", "reya1", "reya2", "reya3", "reya4",
"reya5", "reya6", "reyb1", "trueposi", "truenega", "correct",
"benmirro", "digitspa", "lag1", "lag2", "waisimil",
"waisjudg", "tokenes", "verbal", "animal", "buschke",
"visualn", "tokencol", "waisbloc", "digit", "popsiz",
"nshorter", "nlongter", "nverbal", "njudgeme", "naphasia",
"napraxia", "nagnosia", "nconstru", "ndisturb", "delirium",
"majordep", "diagdeme", "profile", "corticr", "corticl",
"subcort", "cnocoglo", "ccogloss", "cad", "cvasdem",
"cother", "cunclass", "daily", "ndiag", "severity",
"shrtlosn", "longlosn", "clanguag", "cdofint", "cintid",
"c235", "c236", "c236a", "c237", "c238", "c239", "c240",
"c241", "c242", "c243", "c244", "c245", "c246", "c247",
"c248", "c249", "c250", "c251", "c252", "c253", "c253am",
"c253ay", "c253bm", "c253by", "c254", "c255", "c256",
"c257", "c258", "c259", "c260", "c261", "c262", "c263",
"c264", "c265", "c266", "c267", "c268", "c269", "c270",
"c271", "c272", "c273", "c274", "c275", "c276", "c277",
"c278", "c279", "c280", "c281", "c282", "c283", "c284",
"c285", "c286", "c286a", "c287", "c287a", "c288", "c289",
"c290", "c291", "c292", "c293", "c294", "c295", "c296",
"c297", "c298", "c299", "c300", "c301", "c302", "c303",
"c304", "c305", "c306", "c307", "c308", "c309", "c310",
"c311", "c312", "c313", "c314", "c315", "c316", "c317",
"c318", "c319", "c320", "c321", "c322", "c323", "c324",
"c325", "c326", "c327", "c328", "c329", "c330", "c331",
"c332", "c333")

#### neuropsych: behavioral ratings
np.beh_rate <- c("coopern", "facility", "fatigabi", "inattn",
"affectn", "articul", "tension", "appearn", "reaction",
"effort", "express", "memoryn", "restless", "insight",
"gaيتn", "agitat", "persever", "impulsiv", "socialn",
"tangent", "comprehe", "confusi", "latency", "rulev",
"tolerate", "visionn", "hearingn", "physical", "hearaid",
"glasses", "wheel")

#### neuropsych: movement disorders
np.move_disorder <- c("gaिटdiso", "tremdiso", "dyskdiso",
"psycdiso", "slowdiso")

#### neuropsych: memory, fluency, etc...
np.battery <- c("wechsler", "buschfr1", "buschcr1", "buschtr1",
"buschfr2", "buschcr2", "buschtr2", "buschfr3", "buschcr3",
"buschtr3", "buschfr", "buschcr", "buschtr", "delaytim",
"reya1", "reya2", "reya3", "reya4", "reya5", "reya6",
"reyb1", "trueposi", "truenega", "correct", "benmirro",
"digitspa", "lag1", "lag2", "waisimil", "waisjudg",
"tokenes", "verbal", "animal", "buschke", "visualn",
"tokencol", "waisbloc", "digit")

#### neuropsych: impairments & diagnoses

```

```

np.impair_diagnoses <- c("nshorter", "nlongter", "nverbal",
  "njudgeme", "naphasia", "napraxia", "nagnosia", "nconstru",
  "ndisturb", "delirium", "majordep", "diagdeme", "profile",
  "corticr", "corticl", "subcort", "cnocoglo", "ccogloss",
  "cad", "cvasdem", "cother", "cunclass", "daily", "ndiag",
  "severity", "shrtlosn", "longlosn")
#### neuropsych: personality
np.personality <- c("c238", "c239", "c240", "c241", "c242",
  "c243", "c244", "c245", "c246")
#### neuropsych: memory
np.memory <- c("c247", "c248", "c249", "c250", "c251", "c252",
  "c253", "c253am", "c253ay")
#### neuropsych: general function
np.gen_func <- c("c253bm", "c253by", "c254", "c255", "c256",
  "c257", "c258", "c259", "c260", "c261", "c262")
#### neuropsych: every day activities
np.everyday <- c("c263", "c264", "c265", "c266", "c267",
  "c268", "c269")
#### neuropsych: delerium
np.delerium <- c("c270", "c271", "c272", "c273", "c274")
#### neuropsych: depression
np.depress <- c("c275", "c276", "c277", "c278", "c279")
#### neuropsych: sleep
np.sleep <- c("c280", "c281", "c282", "c283", "c284", "c285")
#### neuropsych: paranoia
np.paranoia <- c("c286", "c286a", "c287", "c287a")
#### neuropsych: cardiovascular
np.cv <- c("c288", "c289", "c290", "c291")
#### neuropsych: summary
np.gen_sum <- c("c292", "c293", "c294")
#### neuropsych: past medical history
np.history <- c("c295", "c296", "c297", "c298", "c299",
  "c300", "c301", "c302", "c303", "c304", "c305", "c306",
  "c307", "c308", "c309", "c310")

## variables that exist somewhere between predictors and
## additional variables of interest dementia
dementia <- c("shorterm", "longterm", "abstract", "judgemen",
  "aphasia", "apraxia", "agnosia", "construc", "change",
  "work", "social", "relation")
### alzheimer's
alz <- c("addement", "adcognit", "adworsen", "adconsci",
  "onsetage", "absence")

## possible predictors/mediators/variables of interest or
## confounds. Family history
famhistory <- c("twin", "samesex", "identic", "familyhs",
  "alz1", "alz2", "alz3", "sen1", "sen2", "sen3", "par1",
  "par2", "par3")
### languages spoken

```

```

languages <- c("english", "french", "italian", "german",
  "spanish", "iceland", "ukrain", "chinese", "japanese",
  "danish", "arabic", "urdu", "otherlan")
### health issues
health.drugs <- c("thycond", "attack", "oheart", "leukemia",
  "cancer", "proxdiab", "hbpyr", "drug", "height", "sweight")
### regular substance use -- COMBINE SMOKING & DRINKING &
### CAFFEIENE/SUGAR
substances <- c("coffee", "tea", "drinks", "smoke", "pipe",
  "cigars", "beer", "wine", "spirits")
### head injuries
head.inj <- c("consloss", "consage")
### depression
depress <- c("sad", "interest", "appetite", "lossweig",
  "asleep", "awaken", "allday", "tiredall", "move", "worthles",
  "suicide", "most", "impress")

## outcomes caregiver questions/situation
cg.unpaid <- c("askhelp", "notime", "feelstre", "feelemba",
  "feelangr", "affects", "afraid", "dependen", "strained",
  "suffered", "privacy", "soclife", "friends", "expect",
  "expenses", "unable", "lostctrl", "leave", "uncertai",
  "doing", "better", "burdened", "zarscore")
### caregiver 'felt'
cg.felt <- c("bother", "poorapp", "shake", "good", "mind",
  "depressd", "eveffort", "hopeful", "failure", "fearful",
  "restls", "happy", "talkless", "lonely", "unfriend",
  "enjoylif", "cryspell", "feltsad", "dislike", "getgoing",
  "cesscore")

targeted.predictors <- CSHA.1991.zarit_disorders[, c("CASEID",
  demographics, cg.demographics, cognition, beh.disturb,
  np.battery)]

full.predictors <- CSHA.1991.zarit_disorders[, c("CASEID",
  demographics, famhistory, cg.demographics, languages,
  health.drugs, substances, head.inj, cognition, beh.disturb,
  depress, dementia, alz, neuropsych)]

outcomes <- CSHA.1991.zarit_disorders[, c("CASEID", cg.unpaid,
  cg.felt)]

dim(targeted.predictors)

## [1] 851 68
dim(full.predictors)

## [1] 851 330
dim(outcomes)

## [1] 851 45

```

For our project we should begin with just the two sets of variables identified as `target.predictors` and `outcomes`. At this point we have some re-coding to do to convert missing and NA codes, and to convert YES/NO responses or scales to a more coherent format.