

## **PRACTICAL**

### **Phenotypic Data Fields**

#### **Core data:**

1. **StudyAcronym** = PRACTICAL Study acronym
2. **SubStudy** = Identifier within study strata (e.g. sub cohorts, multi-ethnic cohort, different recruitment groups, or other strata)
3. **SampleID** = PRACTICAL unique person identified
4. **CaCo** = Case control status (**0**=Control, **1**=Case, **9**=QC)
5. **Former\_control** = Patients first recruited as controls but who has developed cancer (**0**=No, **1**=Yes)
6. **Ethnicity** = Ethnic origin (**1**=European, **2**=East Asian, **3**=African American, **4**=Hawaiian, **5**=Hispanic American, **6**=South Asian, **7**=Black African, **8**=Black Caribbean, **0**=Other (to specify in *Ethnicity\_extra*), **U**=unknown)
7. **DateBirth** = Date of birth (yyyy/mm/dd)
8. **Interview** =
  - a. **DateInt** = Date at interview/questionnaire for cases and controls (yyyy/mm/dd)
  - b. **AgeInt** = Age at interview/questionnaire for cases and controls (years)
9. **Diagnosis** =
  - a. **DateDiag** = Date at diagnosis for cases (yyyy/mm/dd)
  - b. **AgeDiag** = Age at diagnosis for cases (years)
10. **FamHist** = Family history of prostate cancer in a first degree relative (**0**=No, **1**=Yes (at least one *FIRST* degree relative), **U**=Unknown)
  - a. **FamHistN** = Number of affected first degree relatives

#### **Survival data:**

11. **Follow Up** =
  - a. **DateLastFU** = Date last follow-up (yyyy/mm/dd)
  - b. **MonthsToLastFU** = Months between diagnosis and last follow-up (*Months*)
  - c. **AgeLastFU** = Age at last follow-up (*Years*)
  - d. **VitalStatus** = Vital status at last follow-up (**0**=alive, **1**=dead, **U**=unknown)
12. **Death** =
  - a. **DateDeath** = Date of death (yyyy/mm/dd)
  - b. **MonthsToDeath** = Months between diagnosis and death (*Months*)
  - c. **AgeDeath** = Age at death (*Years*)
  - d. **CauseDeath** = Cause of Death (**1**=prostate cancer, **2**=other, **U**=unknown)

#### **Clinical / Pathological data:**

13. **Blood Draw** =
  - a. **DateBloodDraw** = Date of blood draw (yyyy/mm/dd)
  - b. **MonthsToBloodDraw** = Months from diagnosis to blood draw
  - c. **AgeBloodDraw** = Age at blood draw (*Years*)
14. **Gleason** =
  - a. **GleasonGrade\_1** = Gleason Pattern from the most common (primary) tumor (**1**=Small, uniform glands. **2**=More stroma between glands. **3**=Distinctly infiltrative margins. **4**=Irregular masses of neoplastic glands. **5**=Only occasional gland formation, **U**=unknown)

- b. **GleasonGrade\_2** = Gleason Pattern from the next most common (secondary) tumor (**1**=Small, uniform glands. **2**=More stroma between glands. **3**=Distinctly infiltrative margins. **4**=Irregular masses of neoplastic glands. **5**=Only occasional gland formation, **U**=unknown)
- c. **GleasonScore** = Gleason Score or Gleason Sum (2 to 10, **U**=unknown)
- d. **GleasonScore\_Range** = Range based on Gleason Score (**1**=Gleason Score<5, **2**=Gleason Score 5,6,7, **3**=Gleason Score 8,9,10, **4**=undifferentiated, **U**=unknown)
- e. **GleasonScore\_Source** = Source of Gleason Score to clarify if the Gleason score recorded is from prostate biopsy or prostatectomy. If available, Gleason score from prostatectomy should be preferentially recorded (**1**=biopsy, **2**=prostatectomy, **U**=unknown)

**15. Cancer stage =**

- a. **Tstage** = T-Stage: Size or direct extent of the primary tumour ('T')
- b. **Nstage** = Degree of spread to regional lymph nodes ('N')
- c. **MStage** = Presence of metastasis ('M')
- d. **SEERStaging** = Prostate cancer SEER staging (**Local**=Confined to the prostate, **Regional**=Direct extension involving adjacent local structures and local lymph node, **Distant**=Direct extension or metastasis, **U**=unknown)

**16. PSA =**

- a. **PSADiag** = Prostate specific antigen (PSA) (ng/ml) at diagnosis
- b. **PSA\_FT\_ratio** = Free to total prostate-specific antigen (PSA) ratio

**17. IPSS =** International Prostate Symptom score (at diagnosis)

**18. MethodDetect =** Method of Detection (Screen or clinically detected)

(**1**=screen detected, **2**=clinically detected, **U**=unknown)

**19. Treatment =**

- a. **T\_Prostatectomy** = Radical prostatectomy (includes both open or robotic prostatectomy) as treatment (**0**= No (did not receive), **1**= Yes (did receive that therapy), **U**=unknown)
- b. **T\_ProstatectomyDate** = Date of prostatectomy (yyyy/mm/dd)
- c. **T\_Radiation** = Radiation therapy (includes both external beam and brachytherapy, with or without androgen deprivation therapy (**0**= No (did not receive), **1**= Yes (did receive that therapy), **U**=unknown)
- d. **T\_RadiationStartDate** = Radiation start date (yyyy/mm/dd)
- e. **T\_RadiationStopDate** = Radiation stop date (yyyy/mm/dd)
- f. **T\_ADT** = Androgen deprivation therapy as treatment (**0**= No (did not receive), **1**= Yes (did receive that therapy), **U**=unknown)
- g. **T\_ADTStartDate** = ADT start date (yyyy/mm/dd)
- h. **T\_ADTStopDate** = ADT stop date (yyyy/mm/dd)
- i. **T\_Surveillance** = Active surveillance (defined as primary treatment if no active treatment has been administered within 12 months after diagnosis date). (**0**= No (did not receive), **1**= Yes (did receive that therapy), **U**=unknown)
- j. **T\_SurveillanceStartDate** = Surveillance start date (yyyy/mm/dd)  
**T\_SurveillanceStopDate** = Surveillance stop date (yyyy/mm/dd)
- k. **T\_Chemo** = Chemotherapy (**0**= No (did not receive), **1**= Yes (did receive that therapy), **U**=unknown)
- l. **T\_ChemoStartDate** = Chemo start date (yyyy/mm/dd)
- m. **T\_ChemoStopDate** = Chemo stop date (yyyy/mm/dd)
- n. **T\_Other** = Other (cryotherapy, etc.)
- o. **T\_OtherDate** = Other treatment date (yyyy/mm/dd)

**Epi Data:**

1. **Procedence =**
  - a. **CountryBorn** = Country born  
([http://www.nationsonline.org/oneworld/country\\_code\\_list.htm](http://www.nationsonline.org/oneworld/country_code_list.htm),  
*U=unknown*)
  - b. **Al\_CountryBorn** = Always live in the country born (*1=no, 2=yes, U=unknown*)
  - c. **Years\_live** = Number of years lived in born country
2. **Educ** = Education (*1=none, 2=primary or secondary school, 3= university degree, 4=professional qualification, U=unknown*)
3. **Marital** = Marital status (*1=married, 2=window, 3=single, 4=divorced, 5=separated, 6=partner, U=unknown*)
4. **Occu** Occupational history
5. **P\_Activity** = Overall physical activity (*1=low activity or sedentary, 2=moderate activity, 3=high or energetic activity, U=unknown*)
6. **Smoking\_S** = Smoking status (*1=never, 2=ex-smoker, 3=smoke occasionally, 4=smoke daily, U=unknown*)
  - a. **Cig\_S** = Cigarette smoke per day
  - b. **S\_period** = Smoking period
7. **Alc\_intake** = Alcohol intake (*1=no, 2=yes, U=unknown*)
  - a. **Beer\_intake** = Average beer intake (*1=Never or less than once a month, 2= 1-3 times per month, 3= once a week, 4= 2-4 times per week, 5= 5-6 times per week, 6= once a day, 7= 2-3 times per day, 8= 4-5 times per day, 9= 6+ times per day, U=unknown*)
  - b. **Wine\_intake** = Average wine intake (*1=Never or less than once a month, 2= 1-3 times per month, 3= once a week, 4= 2-4 times per week, 5= 5-6 times per week, 6= once a day, 7= 2-3 times per day, 8= 4-5 times per day, 9= 6+ times per day, U=unknown*)
  - c. **Spirits\_intake** = Average spirits intake (*1=Never or less than once a month, 2= 1-3 times per month, 3= once a week, 4= 2-4 times per week, 5= 5-6 times per week, 6= once a day, 7= 2-3 times per day, 8= 4-5 times per day, 9= 6+ times per day, U=unknown*)
8. **Acne\_H** = Acne history (*1=no, 2=yes, U=unknown*)
  - a. **Age\_Acne** = Age developed acne
9. **Vasect** = Vasectomy (*1=no, 2=yes, U=unknown*)
  - a. **Age\_Vasect** = Age at vasectomy
10. **Baldness** = Male baldness (*1=no baldness, 2=only frontal baldness, 3=only vertex baldness, 4=both frontal and vertex baldness, U=unknown*)
11. **female\_breast\_cancer** = Family history of female breast cancer (*1=no, 2=yes, U=unknown*)
  - a. **BrCa\_mother** = Mother had breast cancer (*1=no, 2=yes, U=unknown*)
  - b. **BrCa\_sister** = Sister/s had breast cancer (*1=no, 2=yes, U=unknown*)
  - c. **BrCa\_daughter** = Daughter/s had breast cancer (*1=no, 2=yes, U=unknown*)
  - d. **BrCa\_other** = Other female relatives had breast cancer (*1=no, 2=yes, U=unknown*)
12. **BE** = Ever undertaken Barium enema (*1=no, 2=yes, U=unknown*)
  - a. **N\_BE** = Number of barium enema taken
13. **BM** = Barium meal (*1=no, 2=yes, U=unknown*)
  - a. **N\_BM** = number of times
14. **Lower\_Abdomen\_Xray** = Ever undertaken X-ray around lower abdominal area such as Hip or pelvic or lower back x-ray (*1=no, 2=yes, U=unknown*)
  - a. **N\_Lower\_Abdomen\_Xray** = Number of abdo x-ray taken
15. **H\_hypertension** = History of hypertension (*1=no, 2=yes, U=unknown*)

- a. **Age\_Hypertension** = Age at developing hypertension
16. **H\_Heart** = History of heart disease (*1=no, 2=yes, U=unknown*)
- a. **Age\_Heart** = Age at developing heart disease
17. **Diabetes** =
- a. **T1\_diabetes** = History of type 1 diabetes (clinically diagnosed by doctor) (*1=no, 2=yes, U=unknown*)
- b. **Age\_T1\_diabetes** = Age at developing type 1 diabetes
- c. **T2\_diabetes** = History of type 2 diabetes (clinically diagnosed by doctor) (*1=no, 2=yes, U=unknown*)
- d. **Age\_T2\_diabetes** = Age at developing type 2 diabetes
18. **H\_prostatitis** = History of prostatitis (*1=no, 2=yes, U=unknown*)
19. **H\_BPH** = History of Benign Prostatic Hyperplasia *1=no, 2=yes, U=unknown*
20. **Sexual history** =
- a. **Age\_first\_Ej** = Age at first ejaculation
- b. **F\_Ej\_20** = Frequency of ejaculation at 20s (*1=never, 2=less than once a month, 3=1-3 times a month, 4=once a week, 5=2-3 times per week, 6=4-6 times a week, 7=daily, U=unknown, NA=not applicable*)
- c. **F\_Ej\_30** = Frequency of ejaculation at 30s (*1=never, 2=less than once a month, 3=1-3 times a month, 4=once a week, 5=2-3 times per week, 6=4-6 times a week, 7=daily, U=unknown, NA=not applicable*)
- d. **F\_Ej\_40** = Frequency of ejaculation at 40s (*1=never, 2=less than once a month, 3=1-3 times a month, 4=once a week, 5=2-3 times per week, 6=4-6 times a week, 7=daily, U=unknown, NA=not applicable*)
- e. **Age\_first\_In** = Age at first sexual intercourse
- f. **N\_partners** = Number of sexual intercourse partners
- g. **N\_women** = Number of women whom men had sexual intercourse with
- h. **Ht** = Height (*cm*)
21. **BMI** =
- a. **Wt\_at20s** = Weight during their 20s (*Kg*)
- b. **Wt\_at30s** = Weight during their 30s (*Kg*)
- c. **Wt\_at40s** = Weight during their 40s (*Kg*)
- d. **Wt\_at50s** = Weight during their 50s (*Kg*)
- e. **Wt\_at60s** = Weight during their 60s (*Kg*)
- f. **C\_wt** = Current weight (*Kg*)
- g. **BMI** = Body Mass Index = weight in kilograms divided by the square of the height in metres (*kg/m<sup>2</sup>*)
22. **Rthand** = Right hand pattern (finger length pattern) (*1=index finger is longer than ring finger, 2=index finger is equal to ring finger, 3=index finger is shorter than ring finger, U=unknown, NA=not applicable*)
23. **Statins** = Ever use of Statins regularly ( $\geq 2$  times/week) (e.g., Atorvastatin, Cerivastatin, Simvastatin) (*1=no, 2=yes, U=unknown*)
- a. **Statin\_type** = Type of statin (or brand name) that has been taken
- b. **StatinsCurrent** = Current use of Statins regularly ( $\geq 2$  times/week) (*1=no, 2=yes, U=unknown*)
- c. **StatinsStart** = Age when use of Statins started
- d. **StatinsYrs** = Total duration of use of Statins (*Number or range*)
- e. **Statin\_dosage** = Weekly dose of Statins (in mg)
- f. **Statin\_freq** = Average frequency of taking the medicine (*1=less than once a month, 2=two to four times per week, 3= once a week, 4= once a day, 5=two to four times per day*)
- g. **StatinsStop** = If stopped statin, approximate age of stopping