

# EAVE II Data Dictionary



EAVE II Analysts

12 January, 2021

- 1 GP Data Demographics
- 2 GP Diagnostics
- 3 ECOSS
- 4 Health Board Lookup
- 5 Severe cases
- 6 Household identifier
- 7 QCOVID

## 1 GP Data Demographics

### 1.1 Data information

<b>Description</b>	Baseline GP extract on all patients and their key demographics
<b>Data file name</b>	EAVE_demographics.rds
<b>Type</b>	Cleaned
<b>Scripts</b>	NA
<b>Update frequency</b>	Bi-monthly to 3 months
<b>Update timesamps</b>	NA
<b>Analyses used</b>	NA

### 1.2 Variable information

Name	Label	Description	Type	Data_source	Values	Derived variable	Derivation hierarchy	Comments
EAVE_LINKNO	EAVE Link Number	Randomly assigned study index number	character	GP data	EAVE(7)digits (e.g. EAVE1234567)	NA	NA	NA
ageYear	Age	Age in years	integer	GP data	Min= 0, Max= 120	NA	NA	NA
Sex	Sex	Sex at birth	character	GP data	F, M	NA	NA	NA
simd2020_sc_quintile	SIMD Quintile	Scottish Index of Multiple Deprivation (SIMD) Quintile (2020)	integer	GP data	1,2,3,4,5	NA	NA	SIMD in 2020
ur6_2016	Urban/Rural classification	Urban/Rural 6-fold classification	integer	GP data	1,2,3,4,5,6	NA	NA	Classification in 2016
ur6_2016_name	Urban/Rural classification name	Urban/Rural 6-fold classification	character	GP data	1 Large Urban Areas, 2 Other urban Areas, 3 Accessible Small Towns, 4 Remote Small Towns, 5 Accessible Rural, 6 Remote Rural	NA	NA	Classification in 2016
hb2019	NHS Health Board Code	Codes of NHS Health Boards	character	GP data	S08000015, S08000016, S08000017, S08000019, S08000020, S08000022, S08000024, S08000025, S08000026, S08000028, S08000029, S08000030, S08000031, S08000032	NA	NA	Names of Health Boards in 2019

Name	Label	Description	Type	Data_source	Values	Derived variable	Derivation hierarchy	Comments
hb2019name	NHS Health Board Name	Names of NHS Health Boards	character	GP data	NHS Ayrshire and Arran, NHS Borders, NHS Dumfries and Galloway, NHS Fife, NHS Forth Valley, NHS Grampian, NHS Greater Glasgow and Clyde, NHS Highland, NHS Lanarkshire, NHS Lothian, NHS Orkney, NHS Shetland, NHS Tayside, NHS Western Isles	NA	NA	Names of Health Boards in 2019

## 2 GP Diagnostics

### 2.1 Data information

<b>Description</b>	Baseline GP extract of all EAVE II risk groups
<b>Data file name</b>	NAME?
<b>Type</b>	Cleaned
<b>Scripts</b>	NA
<b>Update frequency</b>	Bi-monthly to 3 months
<b>Update timesamps</b>	NA
<b>Analyses used</b>	NA

### 2.2 Variable information

Name	Label	Description	Type	Data_source	Values	Derived variable	Derivation hierarchy
EAVE_LINKNO	EAVE Link Number	Randomly assigned study index number	character	GP data	EAVE(7)digits (e.g. EAVE1234567)	NA	NA
diag	Diagnosis (GP cluster)	Medical diagnosis, clinical indicators and other health-related characteristics	character	GP data	EAVEII_dd_gp_clusters.xlsx	NA	NA

## 3 ECOSS

### 3.1 Data information

<b>Description</b>	Laboratory data on COVID-19 testing
<b>Data file name</b>	ECOSS_cohort.rds
<b>Type</b>	Cleaned
<b>Scripts</b>	NA
<b>Update frequency</b>	NA
<b>Update timesamps</b>	NA
<b>Analyses used</b>	NA

### 3.2 Variable information

Name	Label	Description	Type	Data_source	Values	Derived variable	Derivation hierarchy	Comments
EAVE_LINKNO	EAVE Link Number	Randomly assigned study index number	character	ECOSS	EAVE(7)digits (e.g. EAVE1234567)	NA	NA	NA

Name	Label	Description	Type	Data_source	Values	Derived variable	Derivation hierarchy	Comments
ECOSSID	ECOSS Link Number	Identifier	integer	ECOSS	8 digits (e.g. 12345678)	NA	NA	NA
Sex	Sex	Sex at birth	character	ECOSS	F, M, U (Unknown)	NA	NA	NA
AgeYear	Age	Age in years	integer	ECOSS	Min= 0, Max= 130	NA	NA	NA
HealthBoard	NHS Health Board	Name abbreviation of NHS Health Boards	character	ECOSS	AA, BR, DG, FF, FV, GC, GR, HG, LN, LO, OR, SH, TY, UK (Unknown), WI	NA	NA	NA
PostCode	Postcode	Postcode	character	ECOSS	7 digits/letters	NA	NA	NA
LabSpecimenNo	Number of laboratory specimen	Number of laboratory specimen	character	ECOSS	11 digits/letters	NA	NA	NA
Specimendate	date of specimen	date of specimen collection?	date	ECOSS	yyyy-mm-dd	NA	NA	NA
OriginalOrganism	Original organism	Original organism - Coronavirus 2019 positive or negative	character	ECOSS	Coronavirus 2019-nCoV, Coronavirus 2019-nCoV negative	NA	NA	NA
NCOV_RESULT	COVID-19 result	Novel Coronavirus result	character	ECOSS	Negative, Positive	NA	NA	NA
result	Result	Result	integer	ECOSS	0, 1	NA	NA	NA
SpecimenOrigin	Origin of specimen	Origin of specimen	character	ECOSS	Community, Drug Service, GP, GUM Clinic, H.M Prison, Hospital, Local Authority EH Dept, NHS-Other, Non-NHS-Other, Occupational Health, Private, Residential Care Homes	NA	NA	NA
EcosdateReceived	ECOSS date	date of laboratory result received by ECOSS?	date	ECOSS	yyyy-mm-dd	NA	NA	NA
SubLab	Sub lab	Location and name of laboratory?	character	ECOSS	ABD:GHB, AYR:ARL, BOR:ARL, DEE:UNI, DMF:ROY, EDI:RVL, FIF:ARL, GIL:BAN, GLA:GJH, GLA:REV, INV:RAI, LAN:HAR, LAN:LAW, LAN:MON, NHS:COV, ORK:BAL, STI:ROY, WES:LES	NA	NA	NA
SourceLab	Source lab	Location and name of laboratory?	character	ECOSS	ABD:ALB, ABD:GHB, ABD:ROY, AYR:ARL, BOR:ARL, DEE:UNI, DMF:ROY, DUN:ARL, EDI:ROY, EDI:RVL, EDI:SCH, EDI:UNI, EDI:WGH, FIF:ARL, GIL:BAN, GLA:GAR, GLA:GJH, GLA:NUF, GLA:REV, GLA:ROY, GLA:SCH, GLA:SCS, GLA:SGH, GLA:STR, GLA:UNI, GLA:VIC, GLA:WES, GLA:CLY, INV:RAI, LAN:ALL, LAN:HAR, LAN:LAW, LAN:MON, NHS:UKL, OBA:LAB, ORK:BAL, PAI:RAH, STI:ROY, WES:LES, WLO:BAN	NA	NA	NA
Category	Category	category?	character	ECOSS	NA	NA	NA	NA
KEYEMPLOYER	Key employer	Healthcare provider?	character	ECOSS	e.g. Angus Council	NA	NA	NA
KEYWORKERTYPE	Key employer type	?	character	ECOSS	citizen, DARA, SELF, NA	NA	NA	NA

Name	Label	Description	Type	Data_source	Values	Derived variable	Derivation hierarchy	Comments
CAREHOMEID	Care home ID	Care home identifier	character	ECOSS	NA	NA	NA	NA
NRS.Reg.date	NRS registration date	National Records of Scotland registration date	date	ECOSS	yyyy-mm-dd	NA	NA	NA
NRS.date.Death	NRS death date	National Records of Scotland death date	date	ECOSS	yyyy-mm-dd	NA	NA	NA
death28	Death within 28 days	Death within 28 days of ?	integer	ECOSS	0, 1	NA	NA	NA

## 4 Health Board Lookup

### 4.1 Data information

<b>Description</b>	Lookup of NHS Scotland Healthboards and the total population
<b>Data file name</b>	HB2019_pop_lookup.rds
<b>Type</b>	Lookup
<b>Scripts</b>	NA
<b>Update frequency</b>	NA
<b>Update timesamps</b>	NA
<b>Analyses used</b>	NA

### 4.2 Variable information

Name	Label	Description	Type	Data_source	Values	Derived variable	Derivation hierarchy	Comments
year	Year	Year of the lookup file used to determine the Health Boards	integer	Health board lookup	2019	NA	NA	NA
hb2019	Health Board code 2019	Codes of NHS Health Boards in 2019	character	Health board lookup	S08000015, S08000016, S08000017, S08000019, S08000020, S08000022, S08000024, S08000025, S08000026, S08000028, S08000029, S08000030, S08000031, S08000032	NA	NA	NA
hb2019name	Health Board name 2019	Names of NHS Health Boards in 2019	character	Health board lookup	NHS Ayrshire and Arran, NHS Borders, NHS Dumfries and Galloway, NHS Fife, NHS Forth Valley, NHS Grampian, NHS Greater Glasgow and Clyde, NHS Highland, NHS Lanarkshire, NHS Lothian, NHS Orkney, NHS Shetland, NHS Tayside, NHS Western Isles	NA	NA	NA
hb2018	Health Board code 2018	Codes of NHS Health Boards in 2018	character	Health board lookup	S08000015, S08000016, S08000017, S08000019, S08000020, S08000022, S08000024, S08000025, S08000026, S08000028, S08000029, S08000030, S08000031, S08000032	NA	NA	NA
hb2014	Health Board code 2014	Codes of NHS Health Boards in 2014	character	Health board lookup	S08000015, S08000016, S08000017, S08000019, S08000020, S08000022, S08000024, S08000025, S08000026, S08000028, S08000029, S08000030, S08000031, S08000033	NA	NA	NA
age	Age	Age in years	integer	Health board lookup	Min= 0, Max= 90	NA	NA	NA

Name	Label	Description	Type	Data_source	Values	Derived variable	Derivation hierarchy	Comments
sex	Sex (number)	Sex at birth	integer	Health board lookup	1, 2	NA	NA	NA
sex_name	Sex	Sex at birth	character	Health board lookup	M, F	NA	NA	NA
pop	Population	Number of people by year, Health Board, age and sex	integer	Health board lookup	e.g. 1508	NA	NA	NA

## 5 Severe cases

### 5.1 Data information

<b>Description</b>	EAVE II severe cases and outcomes
<b>Data file name</b>	cases_severe_dates.rds
<b>Type</b>	Derived
<b>Scripts</b>	NA
<b>Update frequency</b>	NA
<b>Update timestamps</b>	NA
<b>Analyses used</b>	NA

### 5.2 Variable information

Name	Label	Description	Type	Data_source	Values	Derived variable	Derivation hierarchy	Comments
EAVE_LINKNO	EAVE Link Number	Randomly assigned study index number	character	GP data	EAVE(7)digits (e.g. EAVE1234567)	NA	NA	NA
source	Source ID	source of ID (chi) most are from EAVE_demographics some only found in ECOSS	character	NA	EAVE_DEMOG or ECOSS	NA	NA	NA
SPECIMENdate	Specimen date	NA	date	NA	date in format yyyy-mm-dd	NA	NA	NA
CURRENT_POSTCODDE	Postcode	NA	character	case-control files	Postcode in pc7 format	NA	NA	NA
is.case	?	NA	logic	case-control files	TRUE/FALSE (should all be true)	NA	NA	NA
GP_PRAC_NO	GP Practice Number	NA	NA	NA	NA	NA	NA	NA
date_OF_DEATH	date of death	NA	date	case-control files (from ? ECOSS, NRS?)	date in format yyyy-mm-dd	NA	NA	NA
covid_ucod	COVID	NA	numeric	NA	binary 0/1	NA	NA	NA
covid_cod	COVID	NA	numeric	NA	binary 0/1	NA	NA	NA
dead28	Death within 28 days	NA	numeric	NA	binary 0/1	NA	NA	NA
RAPID	??	in RAPID dataset?	numeric	NA	binary 0/1	NA	NA	NA
inhosp	In hospital	in hospital at time of test	numeric	NA	binary 0/1	NA	NA	NA
icu	ICU admission	NA	numeric	NA	binary 0/1	NA	NA	NA

Name	Label	Description	Type	Data_source	Values	Derived variable	Derivation hierarchy	Comments
hdu	HDU admission	NA	numeric	NA	binary 0/1	NA	NA	NA
ICU_date	ICU admission date	date admitted to ICU	date	NA	date in format yyyy-mm-dd	NA	NA	NA
days_to_icu	Days to ICU	days from test to admission to icu	numeric	NA	any +ve value	NA	NA	NA
admission.date	Hospital admission date	date admitted to hospital	date	NA	date in format yyyy-mm-dd	NA	NA	NA
days_2_hosp	Days to hospital	days from test to admission to hospital	numeric	NA	any +ve value (if tested before admission, value changed to 0)	NA	NA	NA

## 6 Household identifier

### 6.1 Data information

<b>Description</b>	EAVE II Household from CHI registry
<b>Data file name</b>	EAVE_demographics.rds
<b>Type</b>	Derived
<b>Scripts</b>	NA
<b>Update frequency</b>	NA
<b>Update timesamps</b>	NA
<b>Analyses used</b>	NA

### 6.2 Variable information

Name	Label	Description	Type	Values	Data_Source	Dataset_name	Derived variable	Derivation hierarchy	Comments
EAVE_LINKNO	EAVE Link Number	Randomly assigned study index number	character	EAVE(7)digits (e.g. EAVE1234567)	GP data	EAVE_demographics.rds	NA	NA	NA
date_address_changed.x	Address change date (March)	date address last changed (march)	numeric	date as number yyyymmdd	CHILI	NA	NA	NA	date address last changed (march)
date_address_changed.y	Address change data (September)	date address last changed (sept)	numeric	date as number yyyymmdd	CHILI	NA	NA	NA	date address last changed (sept)
March_hid	Household ID (March)	Household ID number (valid for March CHILI extract)	NA	NA	CHILI	NA	NA	NA	household ID number (valid for March CHILI extract)
March_hid_count	Number of ID's in household (March)	N persons in household (March)	NA	NA	CHILI	NA	NA	NA	number of persons in household (March)

Name	Label	Description	Type	Values	Data_Source	Dataset_name	Derived variable	Derivation hierarchy	Comments
March_transfer_out_date	Move or transfer date (March)	date moved out of scotland (from March CHILI extract)	date	date	CHILI	NA	NA	NA	late: de-regi: as n in cl date mar
Sept_hid	Household ID (September)	Household ID number (valid for september extract)	numeric	NA	CHILI	NA	NA	NA	hid : mee: hou: info
Sept_hid_count	Number of ID's in household (September)	N persons in household (September)	numeric	NA	CHILI	NA	NA	NA	hid v larg: mer: may repr insti e.g. horr pris
Sept_transfer_out_date	Move or transfer date (September)	date moved out of scotland (from September CHILI extract)	date	date	CHILI	NA	NA	NA	late: de-regi: as n in cl date Sep

## 7 QCOVID

### 7.1 Data information

<b>Description</b>	EAVE II data for QCOVID validation analysis
<b>Data file name</b>	NAME?
<b>Type</b>	Derived
<b>Scripts</b>	NA
<b>Update frequency</b>	NA
<b>Update timesamps</b>	NA
<b>Analyses used</b>	NA

### 7.2 Variable information

Name	Label	Description	Type	Values	Data_Source	Derived variable	Derivation hierarchy	Comments
rowid	NA	UID for batch processor file	integer	n/a	n/a	NA	NA	NA
sex	NA	Sex	integer	0:female; 1:male	GP data or Patient demographic service	NA	NA	This is usually biological sex at birth. There is insufficient data to model those who are intersex or trans, or indeed what the long term effect of hormone treatment etc on COVID risk for those who are biologically reassigned.
age	NA	Age	integer	19-100	GP data or Patient demographic service	NA	NA	The age in years at the search date

Name	Label	Description	Type	Values	Data_Source	Derived variable	Derivation hierarchy	Comments
b2_82	NA	Have you been prescribed immunosuppressants prescribed by your GP	integer	0:false; 1:true	GP data	NA	NA	Prescribed four or more times in the previous 6 months
b2_leukolaba	NA	Are you taking anti-leukotriene or long acting beta2-agonists (LABA)?	integer	0:false; 1:true	GP data	NA	NA	Prescribed four or more times in the previous 6 months
b2_prednisolone	NA	Have you been prescribed oral steroids by your GP in the last 6 months?	integer	0:false; 1:true	GP data	NA	NA	oral prednisolone containing preparations prescribed four or more times in the previous 6 months
b_AF	NA	Do you have atrial fibrillation?	integer	0:false; 1:true	GP data	NA	NA	NA
b_CCF	NA	Do you have heart failure?	integer	0:false; 1:true	GP data	NA	NA	NA
b_asthma	NA	Do you have asthma?	integer	0:false; 1:true	GP data	NA	NA	NA
b_bloodcancer	NA	Have you a cancer of the blood or bone marrow such as leukaemia, myelodysplastic syndromes, lymphoma or myeloma and are at any stage of treatment?	integer	0:false; 1:true	GP data	NA	NA	NA
b_cerebralpalsy	NA	Do you have cerebral palsy?	integer	0:false; 1:true	GP data	NA	NA	NA
b_chd	NA	Do you have coronary heart disease?	integer	0:false; 1:true	GP data	NA	NA	NA
b_cirrhosis	NA	Do you have cirrhosis of the liver?	integer	0:false; 1:true	GP data	NA	NA	NA
b_congenheart	NA	Do you have congenital heart disease or have you had surgery for it in the past?	integer	0:false; 1:true	GP data or HES (OPCS)	NA	NA	The categorisation should be based either a Read code for congenital heart disease OR an HES OPCS code for surgery for congenital heart disease ever
b_copd	NA	Do you have chronic obstructive pulmonary disease (COPD)?	integer	0:false; 1:true	GP data	NA	NA	NA
b_dementia	NA	Do you have dementia?	integer	0:false; 1:true	GP data	NA	NA	NA
b_epilepsy	NA	Do you have epilepsy?	integer	0:false; 1:true	GP data	NA	NA	NA
b_fracture4	NA	Have you had a prior fracture of hip, wrist, spine or humerus?	integer	0:false; 1:true	GP data	NA	NA	NA
b_neurorare	NA	Do you have motor neurone disease, multiple sclerosis, myaesthesia, or Huntingtons's Chorea?	integer	0:false; 1:true	GP data	NA	NA	NA
b_parkinsons	NA	Do you have Parkinson's disease?	integer	0:false; 1:true	GP data	NA	NA	NA
b_pulmhyper	NA	Do you have pulmonary hypertension or pulmonary fibrosis?	integer	0:false; 1:true	GP data	NA	NA	NA



Name	Label	Description	Type	Values	Data_Source	Derived variable	Derivation hierarchy	Comments
b_pulmrare	NA	Do you have cystic fibrosis or bronchiectasis or alveolitis?	integer	0:false; 1:true	GP data	NA	NA	NA
b_pvd	NA	Do you have peripheral vascular disease?	integer	0:false; 1:true	GP data	NA	NA	NA
b_ra_sle	NA	Do you have rheumatoid arthritis or SLE?	integer	0:false; 1:true	GP data	NA	NA	NA
b_respcancer	NA	Do you have lung or oral cancer?	integer	0:false; 1:true	GP data	NA	NA	NA
b_semi	NA	Do you have severe mental illness?	integer	0:false; 1:true	GP data	NA	NA	NA
b_sicklecelldisease	NA	Do you have sickle cell disease or severe combined immune deficiency syndromes?	integer	0:false; 1:true	GP data	NA	NA	NA
b_stroke	NA	Have you had a stroke or TIA?	integer	0:false; 1:true	GP data	NA	NA	NA
diabetes_cat	NA	Do ypou have diabetes?	integer	0:none; 1:Type 1; 2:Type 2	GP data	NA	NA	Some patients with diabetes will have codes for type 1 and type2 both recorded in their medical record. It is likely that the most recently recorded code is the most reliable. Therefore, to categorise patients into type 1 or type 2, then identify the latest code recorded in the medical record. If it is Type 1, then code the patient as type 1. If it is type 2, then code the patient as Type 2. If a patient has two codes (one for type 1 and one for type2) recorded on the same day and that is the latest recorded entry, then assume the patient is Type2. The final categorisation should ensure that patients are EITHER Type 1 or Type 2 or neither.
b_vte	NA	Have you had a thrombosis or pulmonary embolus?	integer	0:false; 1:true	GP data	NA	NA	NA
bmi	NA	Body Mass Index	double	double	GP data	NA	NA	The most recently recorded patient BMI within the last 5 years.
chemocat	NA	Have you had chemotherapy in the last 12 months?	integer	0:none; 1: Group A; 2: Group B; 3: Group C	SACT	NA	NA	Chemotherapy prescribed in preceding 12 months as recorded on the Systemic Anti Cancer Treatment (SACT) data. Chemotherapy classified into 3 cateogries (sheet 3)
ethnicity	NA	What is your ethnic group?	integer	0 -> 17	GP data	NA	NA	See separate table in this spreadsheet for the values needed for the batch processor. For the SDK, these values are represented as text within the Ethnicity type (sheet 2)
homecat	NA	what is your housing category - care home or homeless or neither?	integer	0: neither; 1: care home; 2: homeless	GP data	NA	NA	The most recently recorded accommodation status recorded on GP record

Name	Label	Description	Type	Values	Data_Source	Derived variable	Derivation hierarchy	Comments
learnecat	NA	Do you have a learning disability or Down's Syndrome?	integer	0: neither; 1: learning disability; 2: Down's	GP data	NA	NA	The most recorded recorded value. If some has a code both for learning disability and Downs, they should be coded as Downs
p_marrow6	NA	Have you had a bone marrow or stem cell transplant in the last 6 months?	integer	0:false; 1:true	HES (OPCS)	NA	NA	NA
p_radio6	NA	Have you had radiotherapy in the last 6 months?	integer	0:false; 1:true	Radiotherapy Data Set (RTDS) OR HES (OPCS)	NA	NA	coded as having radiotherapy in the preceding 6 months on either HES or RTDS
p_solidtransplant	NA	Have you had a solid organ transplant (lung, liver, somach, pancreas, spleen, heart or thymus)?	integer	0:false; 1:true	HES	NA	NA	coded as having solid organ transplant ever on HES
renalcat	NA	Do you have kidney disease?	integer	0->5	GP data and HES (OPCS)	NA	NA	NA
town	NA	What is your postcode?	double	double (nullable)	QCovid Postcode Table	NA	NA	If given for the batch processor, this takes priority over any postcode entered
Postcode	NA	NA	string	string (nullable)	NA	NA	NA	User interface and batch processor only. When using the SDK, get the corresponding Townsend score from the Townsend deprivation database. See example code for details.