

## Exercise 6: Upgrading an EpiData 3.1 REC/CHK file pair to an EPX file

At the end of this exercise you should be able to:

- a. Understand how you import an EpiData Entry 3.1 REC / CHK file pair into EpiData Manager
- b. How to edit the new file and add the link to the metadata from the old CHK file to the new EPX file.

### Importing an EpiData REC / CHK file containing data

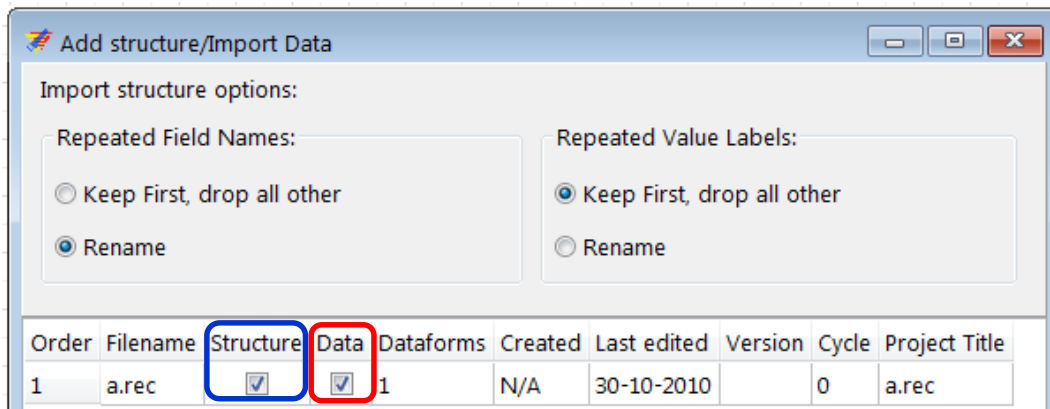
In the required zip folder you find 4 data sets, a, b, c, and d, complete with all 16 files, i.e. for the “a” files a.qes, a.rec, a.chk, and a.eix, containing 75 records, and analogously for the “b”, “c”, and “d” sets. It contains microscopy laboratory data from four laboratories in Yangon (Myanmar), courtesy Dr Ti Ti. If you look at the A.QES file in EpiData Entry 3.1, you see the visual format:

```
id          Unique laboratory identifier  _____
serno       Laboratory serial number  #####
labcode     Laboratory code             _
regdate     Registration date         <dd/mm/yyyy>
age         Examinee's age in years        ##
sex         Examinee's sex           #
reason      Examination reason         #
res1        Result of specimen 1       #.#
res2        Result of specimen 2       #.#
res3        Result of specimen 3       #.#
```

We note that the results were defined as floats with the following values (the definition of which we can best see by opening the A.CHK file (e.g. in a text editor):

```
LABEL label_result
0.0 Negative
0.1 "Scanty, 1 AFB per 100 fields"
0.2 "Scanty, 2 AFB per 100 fields"
0.3 "Scanty, 3 AFB per 100 fields"
0.4 "Scanty, 4 AFB per 100 fields"
0.5 "Scanty, 5 AFB per 100 fields"
0.6 "Scanty, 6 AFB per 100 fields"
0.7 "Scanty, 7 AFB per 100 fields"
0.8 "Scanty, 8 AFB per 100 fields"
0.9 "Scanty, 9 AFB per 100 fields"
1.0 "1+ positive"
2.0 "2+ positive"
3.0 "3+ positive"
4.0 "Positive, not quantified"
5.0 "Scanty, not quantified"
9.0 "No result recorded"
END
```

In EpiData Manager, go to File | Import File... and pick the a.rec to see it then in an editable box:



You note that we encountered this box in an earlier exercise when we imported a value-label set. Note that we can import the structure only or both the structure and the data, as we plan to do here. Once confirmed with OK we get:

and see at the bottom left that there are indeed 75 records which we can browse in **Dataform | Browse Data** (shortcut **CTRL+D**) (first lines only shown):

Record No:	id	serno	labcode	regdate	age	sex	reason	res1	res2	res3
1	A-1001	1001	A	24-10-2003	35	1	5	0	0	9
2	A-1002	1002	A	24-10-2003	20	2	0	0	0	0
3	A-1003	1003	A	24-10-2003	30	1	5	0	0	9
4	A-1004	1004	A	24-10-2003	24	2	2	0	0	9

Note that the button Show Values / Labels is non-functional because the labels are not yet available, we thus see only the values at this point in time. Save the file when prompted as a\_ex06a.epx.

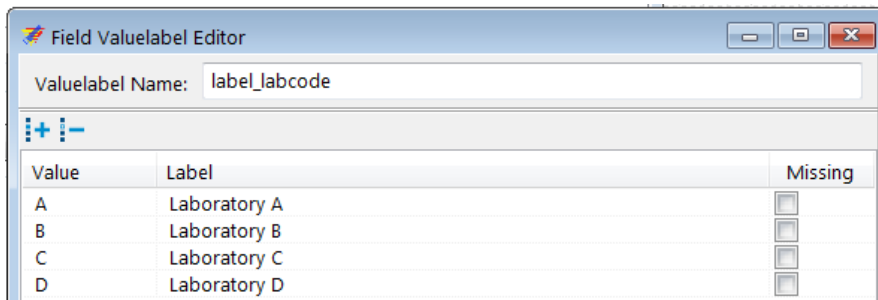
### Make the metadata from the CHK file available to the EPX file

We work our way through the entry form from top to bottom.

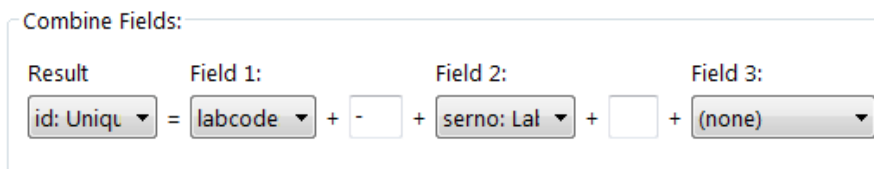
The Unique laboratory identifier (field name `id`) is a calculated value from the Laboratory code and the Laboratory serial number. It is thus a field set to No Enter and calculated after both components that make it up are available.

The Laboratory serial number (field name `serno`) is Must Enter and we Add leading zeros.

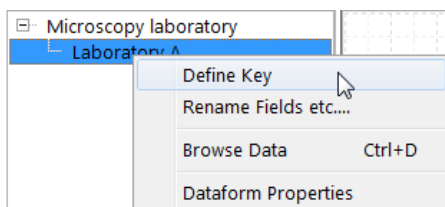
The Laboratory code (field name `labcode`) is Must Enter. We noted above in browsing that the value was “A”. The field does also not have a Label Block (the Laboratory code was added only after data entry in the analysis module). As we will have also laboratories B, C, and D, we will add here a Label block with string values to later accommodate the three additional laboratories. Thus create a New Value Label as:



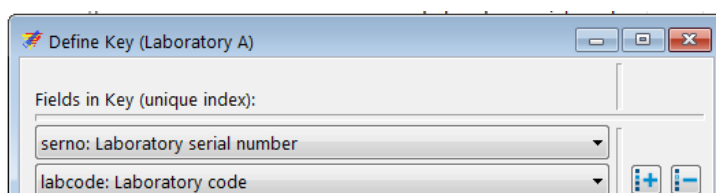
As we do with variables with label blocks, we tick Always show picklist during entry. As we have now all the necessary information we Combine Fields in the Calculate tab:



Let's also make both `labcode` and `serno` Key fields (right-click the data form [which we have here already properly relabeled]):

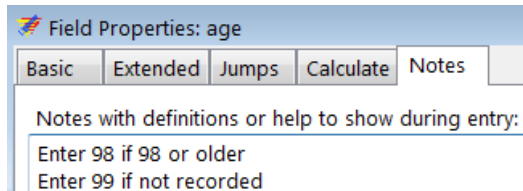


and make `serno` and `labcode` Key fields:



The Registration date (field name `regdate`) is Must Enter and as all laboratory results are from the year 2003 we might as well add a Range from 01/01/2003 to 31/12/2003.

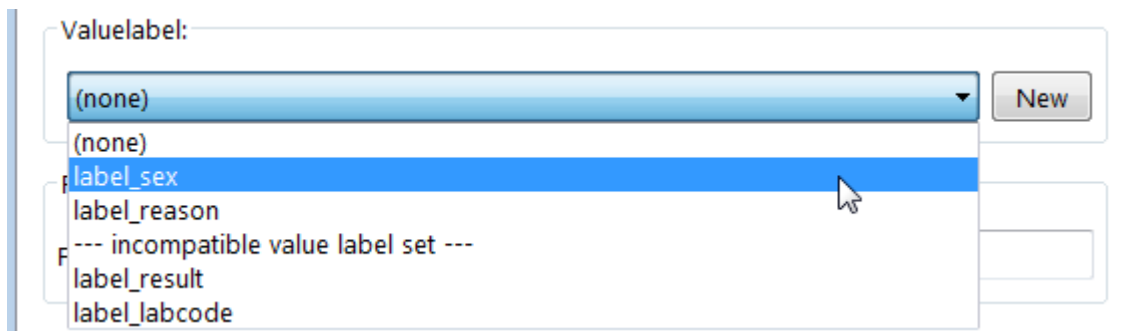
The Examinee's age in years (field name `age`) is Must Enter and as a continuous variable of length 2 we may add Notes:



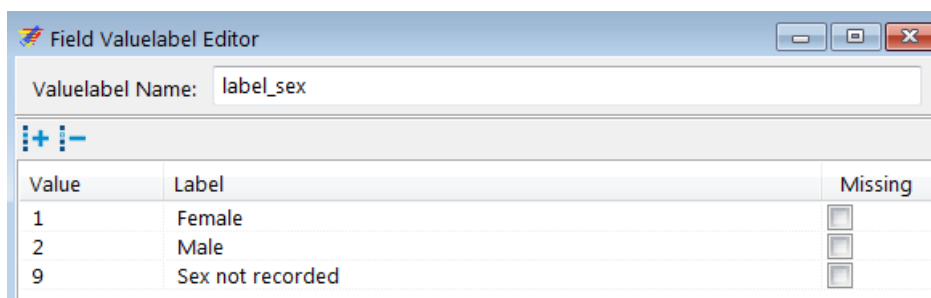
The Examinee's sex (field name `sex`) has no Valuelabel ticked:



However, if you pick the drop-down list, you see that it actually exists (it is in the CHK file) and all we need to do is to pick it:



and then we can view it with Edit that replaces New once picked:



The list of Value labels also shows that we have `label_reason` and `label_result` which we can use for the last four variables to complete updating the new EPX file with the metadata from the old EpiData Entry CHK file.

Finally, once the above is done, we may edit it nicely and check functionality in the EpiData EntryClient (without saving the record!!):

## Tuberculosis Microscopy Laboratory Laboratory A

Unique laboratory identifier

Laboratory serial number

Laboratory code  Laboratory A

Registration date

Examinee's age in years

Examinee's sex  Female

Examination reason  Diagnosis

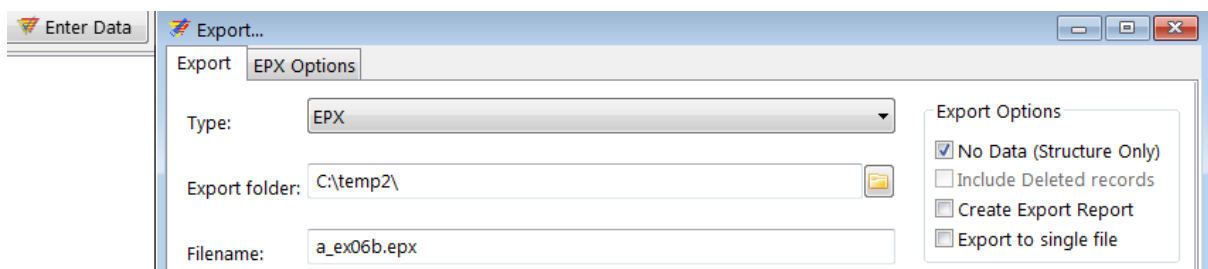
Result of specimen 1  1+ positive

Result of specimen 2  2+ positive

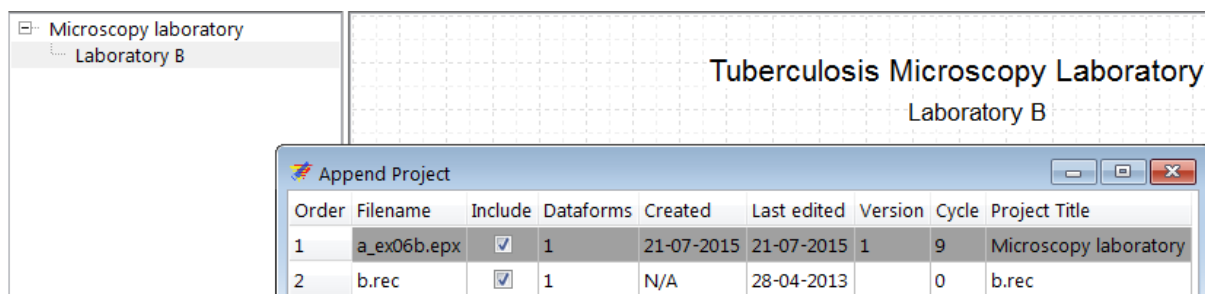
Result of specimen 3  No result recorded

### Exporting the structure of an EPX file and importing REC file data into the new structure

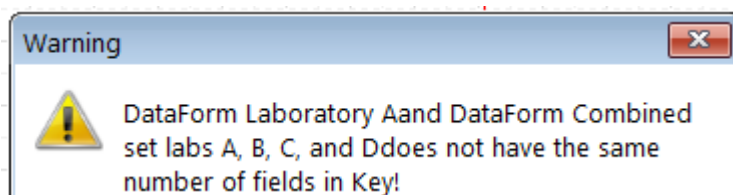
The data form structure of the now created a\_ex06a.epx file can be exported without the data using Export and ticking the appropriate box. We export it to a new file a\_ex06b.epx:



We then open this a\_ex06b.epx file, edit it (changing the name of the data form and the header to become “Laboratory B”) and use Tools | Append and pick the b.rec file:

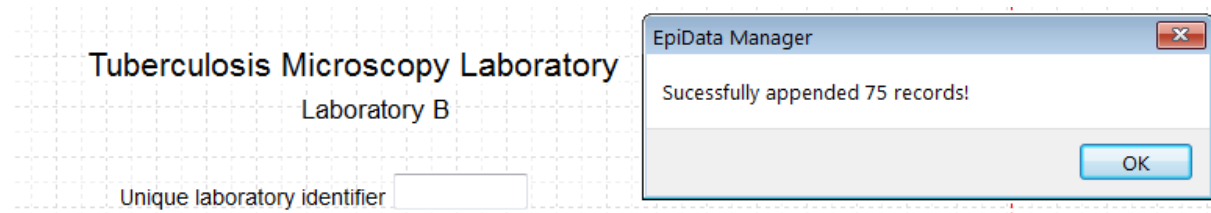


However, doing so, we get an error message:



The reason is that in the EpiData Entry REC file serno is the key, and in this EpiData Manager EPX file serno and labcode are the keys. The best approach is to remove the b.chk file

and to remove both keys in the EPX file and subsequently add them again after successful import which is prompt if these actions are done:



If we wanted to have all four laboratories in one single a\_ex06abcd .epx file, we could have done that simply by adding all four REC files (after removing keys in the EPX file and removing the CHK files for the four REC files after successfully using their metadata). However, it cannot be done in one sweep, each must be added on its own.

**Task:**

- o Create the entire set of four EPX files, so that in the end we have a\_ex06a.epx, a\_ex06b.epx, a\_ex06c.epx, and a\_ex06d.epx*