

## Solution to Exercise 1: A simple Questionnaire

### **Key Point(s):**

- Numbers can be entered into a text field but you will not be able to make any calculations with them.
- It is good practice to write a data documentation sheet before you make your actual EpiData Entry QES file.
- You should always define a value if no answer was provided to a question.
- “Date” is a reserved name in EpiData and cannot be used as a field name.

### **Task:**

- o Complete the data documentation sheet for all fields in the questionnaire. Note that you should always define a value if no answer was provided to a question.*

### **Solution:**

There are many different solutions, but for the sake of uniformity, we will be using the following (but later revise some components of it) as shown on the next page.

Field name	Field label	Field type	Field length	Field values	Value labels	Comment
serno	Laboratory serial number	I	4	1,...,9000 9001, 9002, ...		Serial number starting with 1 each year  Reserve and assign these numbers sequentially if serial number is not unique, and write a data entry note (use F5 to open a note file)
regdate	Registration date	D	10	01/01/2000, ..., 31/12/2005 01/01/1800		Range of legal registration dates Date not recorded
sex	Examinee's sex	T	1	F M 9	Female sex Male sex Sex not recorded	
age	Examinee's age in years	I	3	0, ...,125 999		Range of legal years Age not recorded
reason	Examination reason	T	1	D F 9	Diagnosis Follow-up Reason not recorded	
res1	Result of specimen 1	F	3	0.0 1.0 2.0 3.0 4.0 9.0 5.0 6.0 0.1 0.2 0.3 0.4 0.5	Negative 1+ positive 2+ positive 3+ positive 4+ positive No result recorded Positive, not quantified Scanty, not quantified Scanty, 1 AFB per 100 fields Scanty, 2 AFB per 100 fields Scanty, 3 AFB per 100 fields Scanty, 4 AFB per 100 fields Scanty, 5 AFB per 100 fields	

				0.6 0.7 0.8 0.9	Scanty, 6 AFB per 100 fields Scanty, 7 AFB per 100 fields Scanty, 8 AFB per 100 fields Scanty, 9 AFB per 100 fields	
res2	Result of specimen 2	F	3	0.0 1.0 2.0 3.0 4.0 9.0 5.0 6.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9	Negative 1+ positive 2+ positive 3+ positive 4+ positive No result recorded Positive, not quantified Scanty, not quantified Scanty, 1 AFB per 100 fields Scanty, 2 AFB per 100 fields Scanty, 3 AFB per 100 fields Scanty, 4 AFB per 100 fields Scanty, 5 AFB per 100 fields Scanty, 6 AFB per 100 fields Scanty, 7 AFB per 100 fields Scanty, 8 AFB per 100 fields Scanty, 9 AFB per 100 fields	
res3	Result of specimen 3	F	3	0.0 1.0 2.0 3.0 4.0 9.0 5.0 6.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9	Negative 1+ positive 2+ positive 3+ positive 4+ positive No result recorded Positive, not quantified Scanty, not quantified Scanty, 1 AFB per 100 fields Scanty, 2 AFB per 100 fields Scanty, 3 AFB per 100 fields Scanty, 4 AFB per 100 fields Scanty, 5 AFB per 100 fields Scanty, 6 AFB per 100 fields Scanty, 7 AFB per 100 fields Scanty, 8 AFB per 100 fields Scanty, 9 AFB per 100 fields	

Note the following here. For an unknown laboratory date (REGDATE), we must enter a legally existing (valid) date. EpiData will not accept a date 99/99/9999 nor for that matter 29/02/2001. We chose the value “9” for unknown sex, even if we have defined SEX as a character variable and could thus have used “U” (for “unknown sex”). Just note that “9” is treated as a character variable. It is a personal preference of us to usually use 9 or 99.9 or the like to define unknown values, be this for text or numeric variables. We also introduced a “legal range” for some variables like REGDATE and AGE. We did this a bit arbitrarily, but still tried to keep it within what might be expected.