

Access 2013: An introduction to databases for researchers

Practical workbook

Aims and Learning Objectives

This course provides an introduction to databases for researchers. By the end of this course, you will be able to:

- decide what type of database tool might best suit your research needs and understand the importance of including database design and development costings in your Data Management Plan;
- manage and filter data for analysis in a pre-existing small scale research database using Microsoft Access, the University's site-licensed, small-scale, relational database software;
- explain what you should use the database for and when to use specialist analysis software;
- explain open data format requirements in relation to relational databases and output data from the database in acceptable open formats;
- explain the need for a rigorous set of metadata when sharing/archiving your data at the end of the project as well as use the database software to create some of this metadata.

Document information

Course files

This document and any associated practice files (if needed) are available on the web. To find these, go to www.bristol.ac.uk/it-services/learning/resources and in the **Keyword** box, type the document code given in brackets at the top of this page.

Related documentation

Other related documents are available from the web at:

<http://www.bristol.ac.uk/it-services/learning/resources>



This document is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 2.0 UK: England & Wales Licence (<http://creativecommons.org/licenses/by-nc-sa/2.0/uk/>). Its “original author” is the University of Bristol which should be acknowledged as such in any derivative work.

Introduction

In a relational database data is stored in tables. Each table holds data about a specific 'thing' and is related to at least one other table. Although the data is held in tables, the best way to access and edit data is not within these tables, but by using data entry forms. Queries are used to filter the data and also to do other things, some of which we will see on this course. Reports let you group the data together in printer-friendly format.

The software used on this course is Microsoft Access 2013, the University's site-licensed relational database application, which can be installed on any University-owned machine at the discretion of zonal IT staff.

Prerequisites

You should be familiar with recent versions of MS Office.

Data security

If you store or process restricted UoB data or sensitive personal information on a database, you should be aware of UoB regulations and legal requirements for protecting that data. Further information can be found on the **University Secretary's** website at <http://www.bristol.ac.uk/secretary/> and on the **Information security** website at <http://www.bris.ac.uk/infosec/>.

Contents

Document information

Slide show: what database tool is relevant for my research needs? If using a relational database, what do I need to know about sharing my data?

Task 1	Opening the course database	1
Task 2	Managing the data.....	2
Task 3	Filtering datasets for use in analysis.....	5
	Single table query: adding simple criteria, showing, sorting.....	5
	Wildcards and operators.....	6
	And/Or criteria	7
	Running and modifying saved queries	8
Task 4	Filtering data from multiple tables	9
	Multi-table query with different table joins	9
Task 5	Bulk-changing data values (if we have time).....	10
Task 6	Exporting datasets for analysis and sharing	11
	Exporting data for analysis.....	11
	Exporting data for sharing/archiving	11
Task 7	Creating metadata from within Access.....	12

Task 1 Opening the course database

Objectives To open the course database and select the correct security setting.

Comments If you forget to choose the correct security setting, just close the database and reopen it, this time choosing the correct setting.

On opening Access, the **Getting Started with Microsoft Office Access** window is displayed. From here you can create a new database, access pre-created templates (not recommended) or, as we want to do, open a pre-existing database file.

1.1 To open the Intro database:

- Click on **Open...**, then navigate to where the training files are stored. If you are in the Training Room, you will be told where this is. If you are doing this course in your own time, see the notes on the back of the front page of this document for information on downloading the practice file.
- Double-click on **AnimalFarm.accdb** to open the database.

1.2 To select the correct response to the Security Warning:



Figure 1 – select Enable Content

- On the **Security Warning** message box, select **Enable Content**.
- (If a further **Security Warning** opens asking if you want to make the file a “Trusted Document”, select **No**.)
- Note that the database opens directly onto a data entry form – this is because the **Autoexec Macro** (circled below) tells it to do this when the database opens.

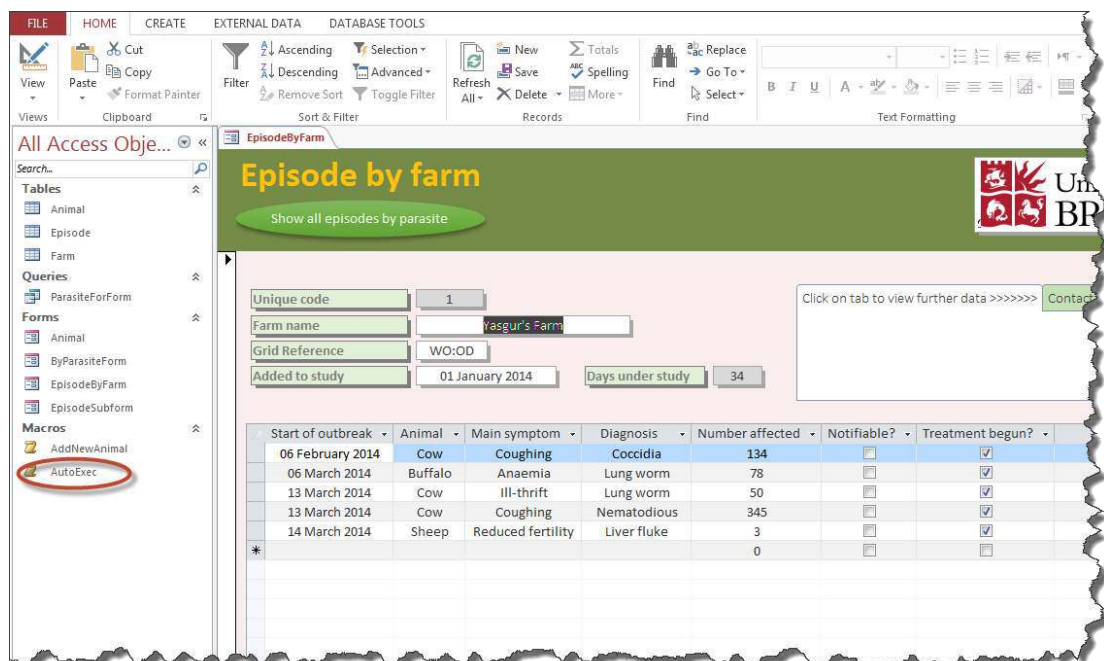


Figure 2 – the work environment showing the Ribbon (top) and Navigation Pane (left) with the Autoexec Macro circled

Task 2 Managing the data

Objectives To find, enter, edit and delete data using a data entry form. To use the form's Datasheet View to quickly sort data.

Comments You can do all of the above in tables, but it is very easy to make mistakes, such as deleting or updating the wrong record. By using data entry forms, you have much more control over what you are doing.

2.1 To move from field to field within the same record:

- Move from the **Farm name** field to the next field by using the **TAB** key on the keyboard.
- Continue tabbing until you reach the **Added to study** field, then press down the **SHIFT** key (keep it held down) and use the **TAB** key again – this time it moves you to the previous field.

2.2 To move from record to record:

- Method 1: tab again to the **Added to study** field and then tab once more – notice that it doesn't take you to the subform or to the tabbed display but takes you directly to the next record. This all depends on what tab rules you set up when creating the form.

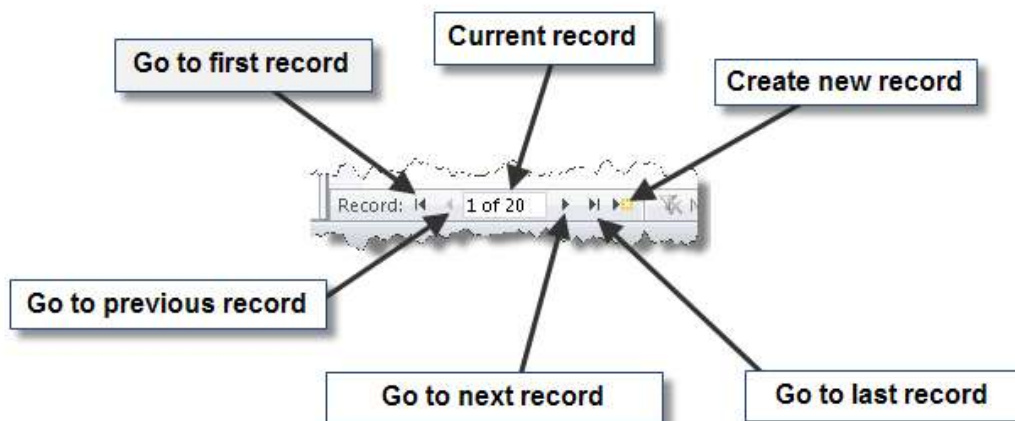


Figure 3 - navigation buttons

- Method 2: navigate between records by clicking on the various **Record navigation** buttons at the bottom of the form window. (Notice that there are buttons at the top of the form to do some of these actions – again, you could create buttons to do all. Larger buttons are generally easier to use than the **Record navigation** buttons at the bottom of the screen.)

2.3 To enter a new record:

- Click on the **New (blank) record** icon at the bottom of the screen.
- Enter some data for your new record – just make it up.
- Click on the **Contact details** tab in the tabbed display instead.
- Click on the label for **Contact** to highlight the data entry field next to it and enter an address – note that you can now tab to the field below as well.


Start of outbreak	Animal	Main symptom	Diagnosis	Number affected	Notifiable?	Treatment begun
03 February 2014	Cow	Ill-thrift	Scour worm	4	<input type="checkbox"/>	<input checked="" type="checkbox"/>
06 February 2014	Sheep	Coughing	Liver fluke	5	<input type="checkbox"/>	<input checked="" type="checkbox"/>
*				0	<input type="checkbox"/>	<input type="checkbox"/>

Figure 4 – Enter some data into the subform

- Click in the **Start of outbreak** field in the subform and add a couple of records – again, just make it up, though notice that in the first field you can choose the date from a small calendar icon that appears to the right of the field. Also notice that you can only choose items from the dropdown lists in the next three fields
- Tab to the **Notifiable?** tick box and press down the **SPACEBAR** a few times to toggle the tick on and off.




Figure 5 – the pencil icon indicates that a record is being edited and is not yet saved

- Save the record by clicking on the column down the left-hand side of the form (if you tab to the next record or close the form, Access automatically saves any unsaved record).
- 2.4** To view your new record in the underlying table:
- Close the **EpisodeByFarm** form (right-click on the tab and select **Close**).
 - In the **Navigation Pane**, double-click on the **Episode** table to open it.  Episode
 - Scroll to the bottom of the record set and view your new record.

Note Remember that everything you do within a data entry form is reflected in the underlying table – so if you delete or alter anything it is gone/changed for good (unless you have sensibly made a backup of the database before altering data).

- Close the **Episode** table and reopen the **EpisodeByFarm** form.

2.5 To find a specific record to edit:

- Click on the **Find** icon in the **Home** tab of the **Ribbon** to open the **Find and Replace** box.  Find
- Enter **mag** in the **Find What** box, change **Look In** to **Current document** and **Match** to **Any Part of Field**, then click the **Find Next** button – we are looking for the record for **Maggie's Farm**.
- Once you have found the correct record press the **Cancel** button on the **Find and Replace** box.

2.6 To edit existing data:

- In the record for **Maggie's Farm**, change **Added to study** to **12 February 2014**.
- Click in the first field in the top record in the **subform** and use the **Find** option to find all instances of **li** (that's a lower case "L" first, by the way).

Note By clicking on *Find Next*, you should have found nine instances of "li". Not just in *Liver fluke* and *Lickey end*, but also in *Reduced fertility*.

2.7 To delete a record:

- Move to the last record using the **Navigation buttons** at the bottom of the screen.
- Click somewhere on the form to select the record.
- On the drop-down arrow next to the **Delete** icon on the **Ribbon**, and choose **Delete Record** – notice that you cannot delete the record because you have added data to the subform.
- Click **OK** then click on the grey box to the left of the first record in the subform (if you have added more than one record to the subform, drag the mouse down the grey boxes to highlight all records).
- Delete the subform record(s) as described above and select **Yes** when asked for confirmation.
- Now try to delete the new farm record again.

2.8 To view the Datasheet View of the form:

Note Datasheet View looks like a table, but includes only fields included in the form. If you do not have the option of *Datasheet View* in your own version of Access, see the appendices for instructions on how to set up this view.

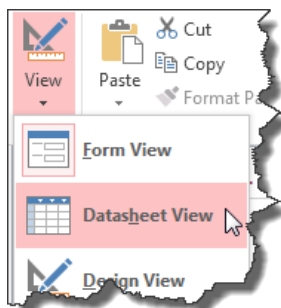


Figure 6 – you can toggle between Form View and Datasheet View of the form

- Click on the downward arrow underneath the **View** icon in Home tab of the **Ribbon**.
- Select **Datasheet View** and notice that **Days under study** is included in this view (**Days under study** does not appear in the **Farm** table because it is a calculated field).

2.9 To sort records:

Note You can sort in both form and datasheet views (and also in tables and queries), but datasheet view may be more useful because you can see many records at once.

- Click somewhere in the **Farm name** column.
- Sort in ascending order (A to Z) followed by descending order (Z to A) using the icons in the **Sort and Filter** section of the **Home** tab in the **Ribbon**.
- If you feel the need to sort the records back into the original order, click on the **Remove Sort** icon underneath the other sort options.



Task 3 Filtering datasets for use in analysis

Objectives To create simple single and multi-table queries and export query results to Excel.

Comments Data are displayed according to selection criteria that you choose. Queries can be saved and run at any time – results will change based on the current set of data.

Single table query: adding simple criteria, showing, sorting

3.1 To create a simple single table query:

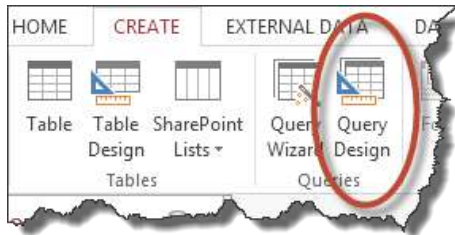


Figure 7 – Query Design in the Create tab of the Ribbon

- Close the form then click on the **Create** tab in the **Ribbon** and click **Query Design**.
- In the **Show Table** dialogue box, double-click on **Episode** then click **Close**.
- Add **FarmID**, **AnimalID**, **SymptomID** and **ParasiteID** from the **Episode** table by double-clicking on each in turn.

Field:	FarmID	AnimalID	SymptomID
Table:	Episode	Episode	Episode
Sort:			
Show:	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Criteria:			

Figure 8 – add the above fields

- Check the result using either the **Datasheet View** icon or the **Run** icon.
- 3.2 To find all records that match a specific criteria:
- Click on the **Design View** icon to return to the grid, then click on the **Criteria** row in the **AnimalID** column.



Field:	FarmID	AnimalID	SymptomID
Table:	Episode	Episode	Episode
Sort:			
Show:	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
Criteria:		"cow"	
or:			

Figure 9 - setting criteria

- Type **Cow** or **cow** – queries are not case sensitive (speech marks are added automatically when you click outside of the field).

- Check the result using either the **Datasheet View** icon or the **Run** icon then switch back to **Query Design**.

3.3 To exclude a field from appearing in the query results:

- Deselect the tick in the **Show** row for **AnimalID** in the above query.
- Check the result in **Datasheet View** then return to **Query Design**.

3.4 To sort the query by a single field:

- Add the tick back into the **Show** row of the **AnimalID** field.

Field:	FarmID	AnimalID	SymptomID	
Table:	Episode	Episode	Episode	
Sort:		Ascending		
Show:	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
Criteria:				

Figure 10 – select sort order from the Sort row dropdown

- Click in the **Sort** row for **AnimalID** and choose **Ascending** or **Descending** from the dropdown box.
- Check the result in **Datasheet View** then return to **Query Design**.

3.5 To discover annoying problems with certain dropdown fields:

- Sort **FarmID** in **Ascending** order and view the query results. It doesn't seem to have worked does it?
- Remove the sort and in the **Criteria** row, enter **animal farm** – now view the query results. Hm, it can't find a match and throws up a 'data mismatch' error message, doesn't it?
- Remove **animal farm** and enter **1** instead, then run the query.

Note The problem here is that *FarmID* is looking up the farm code in the *Farm* table. It shows you the name of the farm, but it is really storing the code in the field. This means that you have to know the codes if you want to see records for a specific farm. This is easily solved, but you need to add another table to the query – see the following task.

- Remove the **1** from **FarmID**.

3.6 To save the query:

- Click on the **Save** icon above the **File** tab and name it **AnimalsAscending**. 
- Close the query.

Wildcards and operators

Note See the appendices for explanations of these and more operators and wildcards.

3.7 To use the * wildcard:

- Set up a new query as before (**Create** tab, **Query Design**), adding **Episode** from the **Show Table** dialogue box.
- Add the same four fields as before then add **NoAffected**, **TreatmentStarted** and **EpisodeStart**.

SymptomID	ParasiteID	NoAff
Episode	Episode	Episo
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	Like "s**"	

Figure 11 – add “l” and the wildcard symbol

- In the criteria field for **ParasiteID** add **s*** and run the query. You should get just the one record that starts with “s”.
- Return to **Design View**, alter the criteria to ***s** then run the query again. You should now get lots of records that end with “s”.
- Return to **Design View**, alter the criteria to ***s*** then run the query again. You should now get every record that includes an “s” anywhere within the field.

3.8 To use the Greater Than operator:

- Remove the criteria from above and in **NoAffected** enter **>10**, then view the query results.
- Back in **Design View**, amend this to **>=10** and view the query results.

3.9 To use the Between operator:

- Remove the criteria from above and in **NoAffected** enter **between 10 and 50**, then view the query results.
- Back in **Design View**, remove the above criteria and add **between l and s** in the criteria field for **ParasiteID**, then run the query again.

And/Or criteria

You can set criteria on more than one field. If criteria are placed on the same line, all criteria must be true. This is called an *AND query*.

3.10 To find episodes where the parasite starts with L and where treatment has started:

- Remove all operators or wildcards (and anything else you might have added) from the above query and enter **l*** (*that’s a lower case L*) into the **ParasiteID** criteria field.
- Enter **Yes** into the **TreatmentStarted** criteria field and check the query results (tick boxes can be matched against “Yes” or “No”).

SymptomID	ParasiteID	NoAffected	TreatmentStarted
Episode	Episode	Episode	Episode
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Like "l**"		Yes

Figure 12 - using more than one criteria

If you want to broaden your results to show records that match one of two (or more) different sets of criteria you can perform an *OR query*.

3.11 To find episodes where either the parasite starts with L or those where treatment has started:

- Remove the **Yes** from **TreatmentStarted**.

SymptomID Episode	ParasiteID Episode	NoAffected Episode	TreatmentStarted Episode
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	Like "I*"		Yes

Figure 13 - using 'and/or' criteria

- In the second criteria row (the OR row) in **TreatmentStarted**, enter **Yes**.
- Check the query results then close the query, saving it as **TreatmentOr** when prompted.

Running and modifying saved queries

3.12 To run a saved query:

- Double-click on **AnimalsAscending** in the **Navigation pane** and note the number of records returned (at the bottom of the screen amongst the **Record Navigation** buttons).

Maggie's Farm	Sheep	Ill-thrift	Licke
Maggie's Farm	Cow	Reduced fertil	Liver
Maggie's Farm	Cow	Reduced fertil	Liver

Figure 14 – click on the grey box

- Delete one of the records (click on the grey box to the side of the record, hit the **DELETE** key and say **Yes** on the message box) then close the datasheet.
- **Run** the query from the **Navigation pane** again and note that there is now one record less – yes, you can update, add and delete data in an open query!
- Close the query again.

Task 4 Filtering data from multiple tables

Objectives To create queries using more than one table. To use different table joins and see the difference in data returned by the query based on join.

Comments BEWARE: the default query setting may provide incomplete data for analysis.

4.1 To create a multi-table query to be used later as the basis for a report:

- Create a new query and add all three tables.

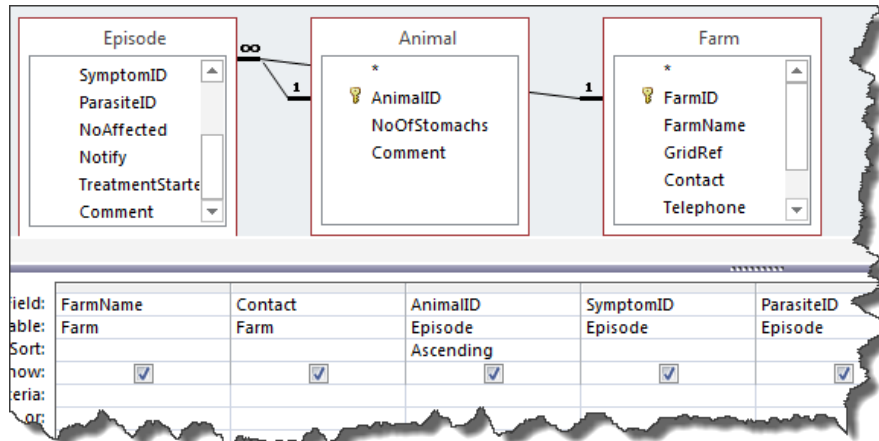


Figure 15 – create a three-table query

- From **Farm** add: **FarmName**; **Contact**: from **Episode** add: **AnimalID**; **SymptomID**; **ParasiteID**; **NoAffected**; **TreatmentStarted**: from **Animal** add: **NoOfStomachs**.
- **Run** the query, then close it, saving when prompted as **QueryForReport**.

Multi-table query with different table joins

4.2 To create a multi-table query and see the results based on different table joins:

- Create a new query and add the **Episode** and **Animal** tables.
- From **Episode** add **FarmID** and from **Animal** add **AnimalID**.
- **Run** the query, check the results and return to **Design View**.

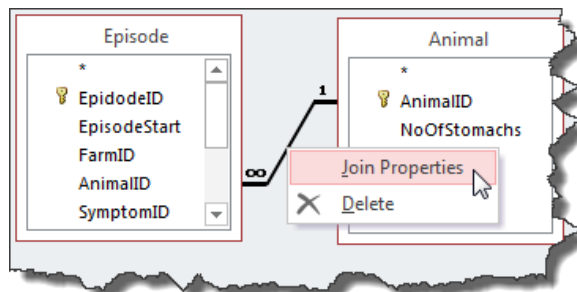


Figure 16 – right-click on the relationship line (this can be fiddly)

- Right-click on the relationship between the two tables and select **Join Properties**.
- Select the second option, click **OK** then run the query to see the different results – you can now see all animals irrespective of whether they have been notified as ill.
- Now try the third option – what is this doing? – and close the query without saving it.

Task 5 Bulk-changing data values (if we have time)

Objectives To create a query to change data in the underlying table.

Comments Make sure that the field you want to change can accept the updated data – for example, you cannot update a Number field with text. If you try to update a field that is too small to accept the update text, the update text will be truncated.

Warning! Action queries can seriously damage the data in your database. Always make a backup of your database before creating new (or running saved) action queries. Never double-click on an action query in the Navigation Pane just ‘to see what it does’ because it will do it. Always check in Design View first by right-clicking on it in the Navigation Pane.

5.1 To change the current entries in the Comments field in the Animal table:

- Create a new query and add the **Animal** table.
- Select **Update** from the **Query Type** section of the **Ribbon**.
- Add the **AnimalID** and **Comments** fields to the QBE grid.
- In **AnimalID** type *antelope* the **Criteria** box, then in the **Or** row below type *buffalo*; in the next **Or** row type *sheep*.



Field:	AnimalID	Comment	
Table:	Animal	Animal	
Update To:		"Stomach has four chambers"	
Criteria:	"antelope"		
or:	"buffalo"		
	"sheep"		

Figure 17 – the QBE grid should look like this

- Click in the **Update To** box in the **Comment** field (notice that the animal names you just typed have speech marks added around them automatically) type ***Stomach has four chambers***.
- View the values that be will changed by selecting **Datasheet View** (do not select the Run icon yet) – you should see a column of three rows with various text in each.
- Return to **Design View** and select **Run**, followed by **Yes** to confirm.
- To see what you have done double-click on **Animal** in the **Tables** section of the **Navigation Pane**.
- Close the **Animal** table and close the query, saving it as ***StomachUpdate***.
- Notice that the saved query has a different icon to the other queries.



Warning! Never double-click an action query to see what it does – because it will do it! Always look to see what it does in Design View first.

5.2 To view an update query in Design View:

- Right-click on **StomachUpdate** in the **Navigation Pane** and select **Design View**.
- Close the query.

Task 6 Exporting datasets for analysis and sharing

Objectives To export data for analysis using specialist software. To export all required data (excluding metadata) in an acceptable open format for sharing at end of project.

Comments Although Access lets you do a certain amount of analysis within the database itself, you really have to know what you are doing if you don't want to end up doing your analysis on incomplete or otherwise flawed data. See note below.

Exporting data for analysis

Note It is much easier to do analysis on data in other packages such as Excel or SPSS. Access is great at storing your data, whilst queries let you filter just the data that you need to do your analysis, but it is not a good package for beginners to try their luck at doing complicated analysis.

6.1 To export a query in Excel format for analysis:

- Right-click on **TreatmentOr**, then select **Export**, followed by **Excel**.
- If in the training room click on **Browse** and save the file to your own **MyFiles** space: if elsewhere it is easiest to save to the **Desktop** – click **Close** on the splash-screen that opens asking if you want to save the export steps.
- Navigate to where you saved the file and double-click on it to open it in **Excel**.

Exporting data for sharing/archiving

6.2 To export data in an acceptable open format for sharing/archiving in accordance with funding requirements (information only – do not do this in the training room):

- Do exactly as above, but for all tables and queries.
- Name each file with the same name as the table or query.
- Open each file in Excel and save as .csv format.

Note You could instead export each table and query as XML, which is an open format that has the added benefit of letting you maintain the structure of the data as well as exporting the data itself. However, XML comes with some problems attached, which is that XML data has to be actively managed to ensure that it remains usable. Access lets you export in XML format, but not all relational databases provide this option. In any case with both formats you will need to create specific metadata so that subsequent researchers can recreate your data structure if required.

Task 7 Creating metadata from within Access

Objectives To create as much of the required metadata as possible from within Access.

Comments Further metadata will need to be created outside of Access so that subsequent researchers can come to your data and replicate the exact datasets as used for analysis. However, Access can create much of this metadata automatically.

7.1 To open the Database Documenter:

- Open your database and close any open objects.
- Click on the **Database Tools** tab and select **Database Documenter**.

7.2 To create the base report that includes much of the required metadata:

- In the **Tables** tab select **Select All** and click on **Options...**
- In the **Include for Table** section select only **Relationships**; in **Include for Fields** select **Names, Data Types and Sizes**; in **Include for Indexes** select **Names and Fields**, then click on **OK**.
- In the **Queries** tab select **Select All** and click on **Options...**
- In **Include for Query** select only **SQL** and select **Nothing** for the following two sections, then click on **OK**.
- In the **Macros** tab select **Select All** and click on **Options...**
- Select only **Actions and Arguments** and click on **OK**.
- (If you have any **Modules**, select all and leave options as default.)
- Click on **OK** to produce the **Object Definition** report.

7.3 To export the **Output Definition** report in an open format:

- Right-click on the report and select **Export**.
- Choose **Text File** and if asked any further questions on saving, choose defaults (other than choosing where you want it saved, of course).

Note You will have to open the resulting file and do some tidying up because Access tends to format it in an 'interestingly messy' way.

Appendix A Common query operators/wildcards

Criteria operators		
Operator	Meaning	Example
<	Less than	< 2000 would find all values up to and including 1999
<=	Less than or equal to	<= 2000 would find all values up to and including 2000
>	Greater than	> 2000 would find all values from 2001 upwards
>=	Greater than or equal to	>= 2000 would find all values from 2000 upwards
<>	Not equal to	<> 0 would find all positive and negative (non-zero) numbers
Between X and Y	Any of the given values	Between 1500 and 2000 would find all values from 1500 to 2000 including both 1500 and 2000

- For a list of all available operators, see <http://office.microsoft.com/en-gb/access-help/table-of-operators-HA010235862.aspx>

Wildcards		
Character	Meaning	Example
?	Matches one character	B?t would find Bat, But, Bet, but not Boat, etc.
*	Matches any number of characters	B*t would find Bat, Boat, Bracket, etc.
#	Matches a single digit in a <i>text</i> field (this <i>does not</i> work on a numeric field)	3#5 would find 375, 385, etc. #05 would find 105, 205, etc.

- For a list of all available wildcards, see: <http://office.microsoft.com/en-gb/access-help/access-wildcard-character-reference-HP005188185.aspx>