

Importing Data in BioNumerics



Jen Castleman
April 2011

Overview

- Import Demographic Information
 - Create ODBC Link
 - Download Data
 - Remove Timestamps

Import Demographic Information

- What can be imported?
 - Data located in Microsoft Excel or Microsoft Access
 - Database fields (ex. patient age, source site)
 - Character data (antimicrobial profiles—not covered in this presentation)
 - Sequence data (not covered in this presentation)

Import Demographic Information

Microsoft Excel - Book1

File Edit View Insert Format Tools Data Window Help Adobe PDF

B18

Isolate Number	Serotype	Country	State	County	SourceSite	SourceType	PatientAge
07E01705	Cough	USA	PA	Butler	Urine	Human	84 FEMALE
07E01714	Infantis	USA	PA	Philadelphia	Stool	Human	25 MALE
07E01715	Saintpaul	USA	PA	Dauphin	Urine	Human	63 FEMALE
07E01716	Agona	USA	PA	Lycoming	Stool	Human	37 FEMALE
07E01718	Sandiego	USA	PA	Philadelphia	Stool	Human	25 FEMALE
07E01719	Montevideo	USA	PA	Cumberland	Stool	Human	40 FEMALE
07E01720	Typhimurium	USA	PA	Lancaster	Stool	Human	42 FEMALE
07E01721	Typhimurium var. O 5 - (Copenhagen)	USA	PA	Berks	Stool	Human	42 MALE
07E01722	Montevideo	USA	PA	Philadelphia	ankle bone	Human	33 MALE
07E01723	Litchfield	USA	PA	Philadelphia	Stool	Human	17 FEMALE
07E01724	Typhi	USA	PA	Lancaster	Blood	Human	19 MALE

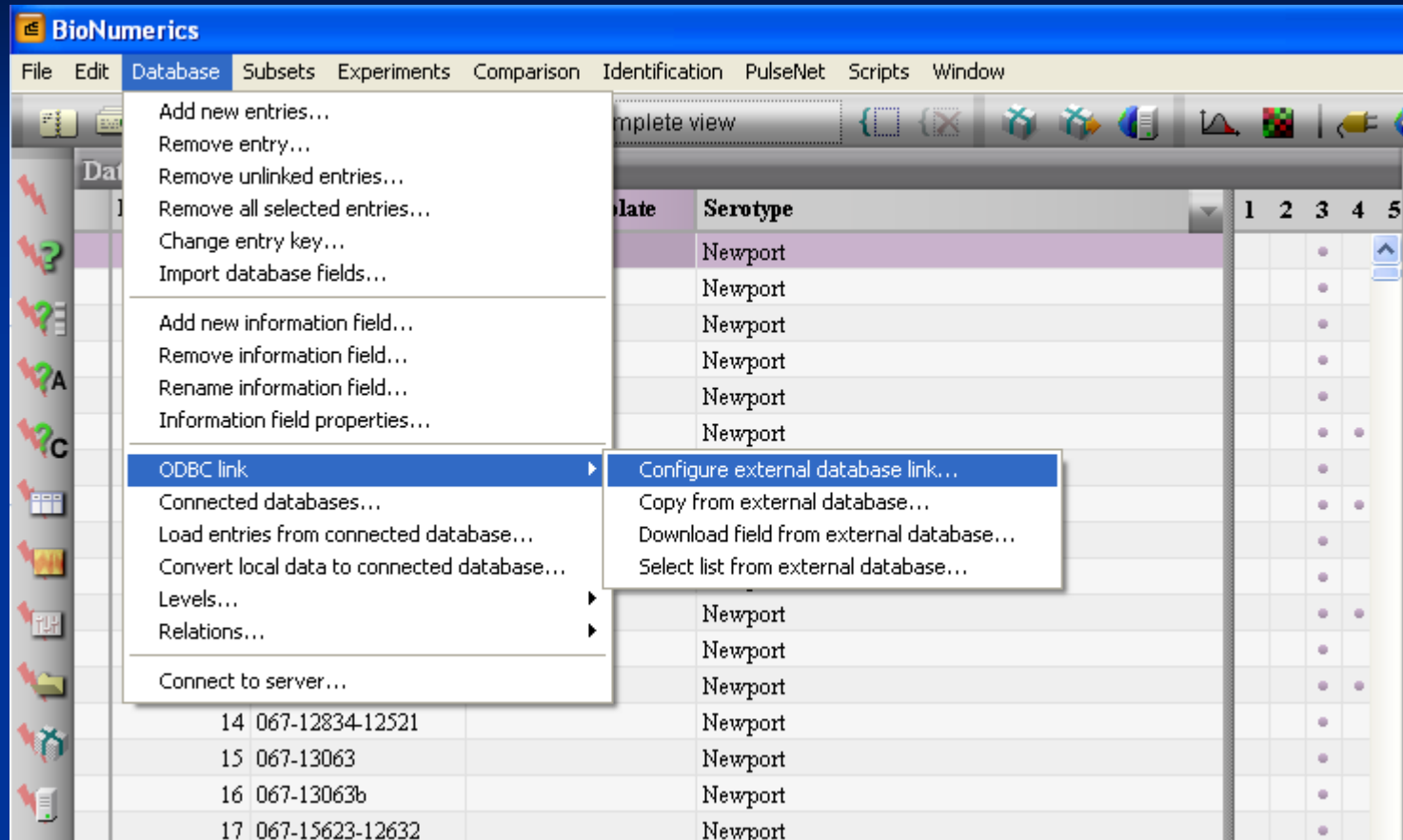
Sheet1

First row of Excel should contain database field names (they do not have to be the same as those found in BioNumerics)

There must be a column containing isolate key numbers

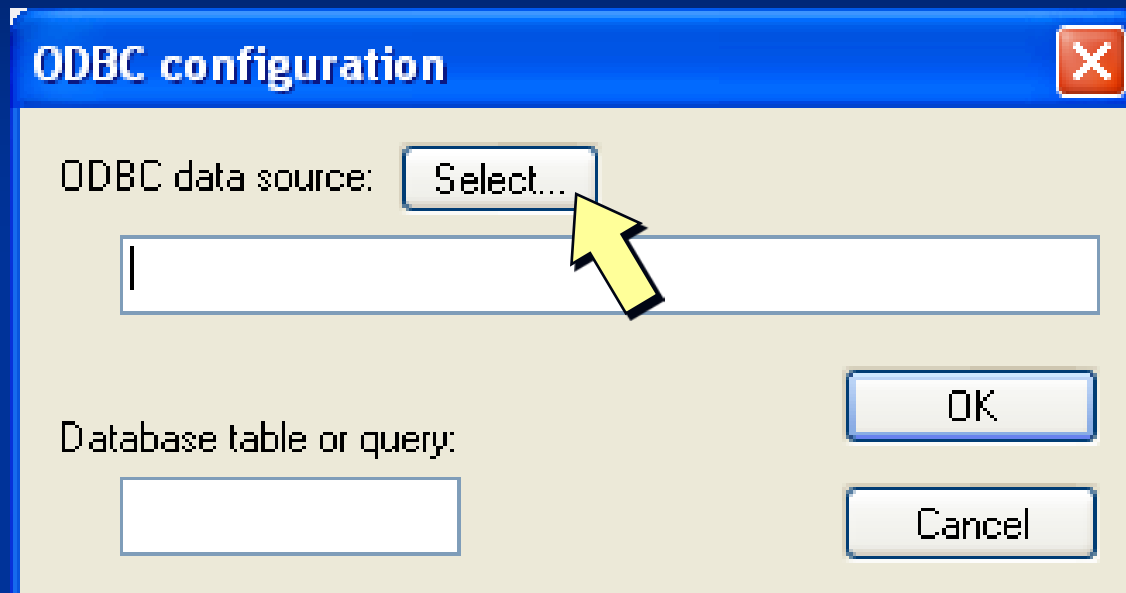
Open the Excel file and note the name of the sheet (default: Sheet1)

Import: Create ODBC Link



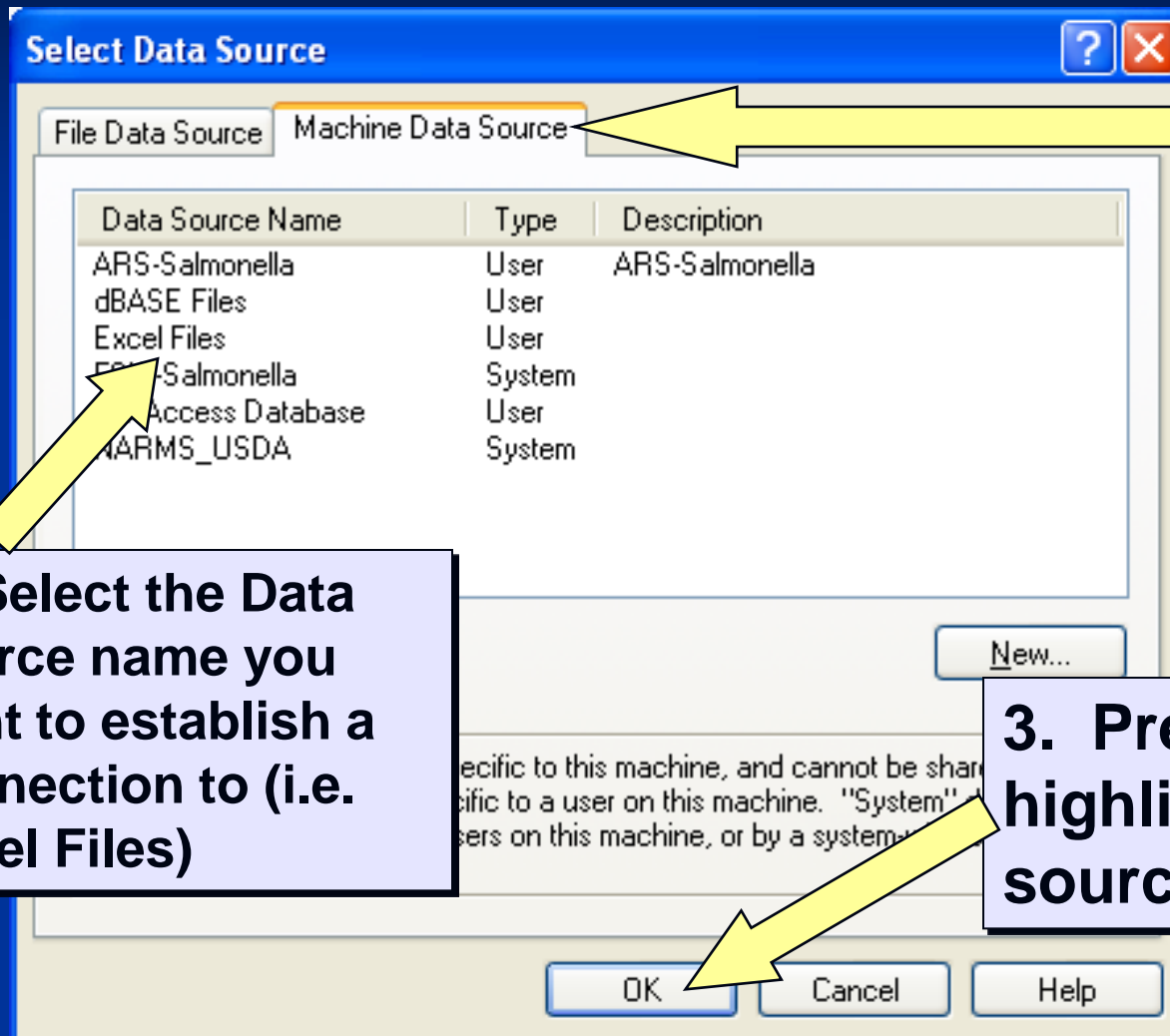
In the main screen of BioNumerics, select Database → ODBC link → Configure external database link...

Import: Create ODBC Link



Press “Select” to bring up the Windows ODBC configuration wizard

Import: Create ODBC Link



1. Select the Machine Data Source tab

2. Select the Data source name you want to establish a connection to (i.e. Excel Files)

3. Press "OK" after highlighting the data source name

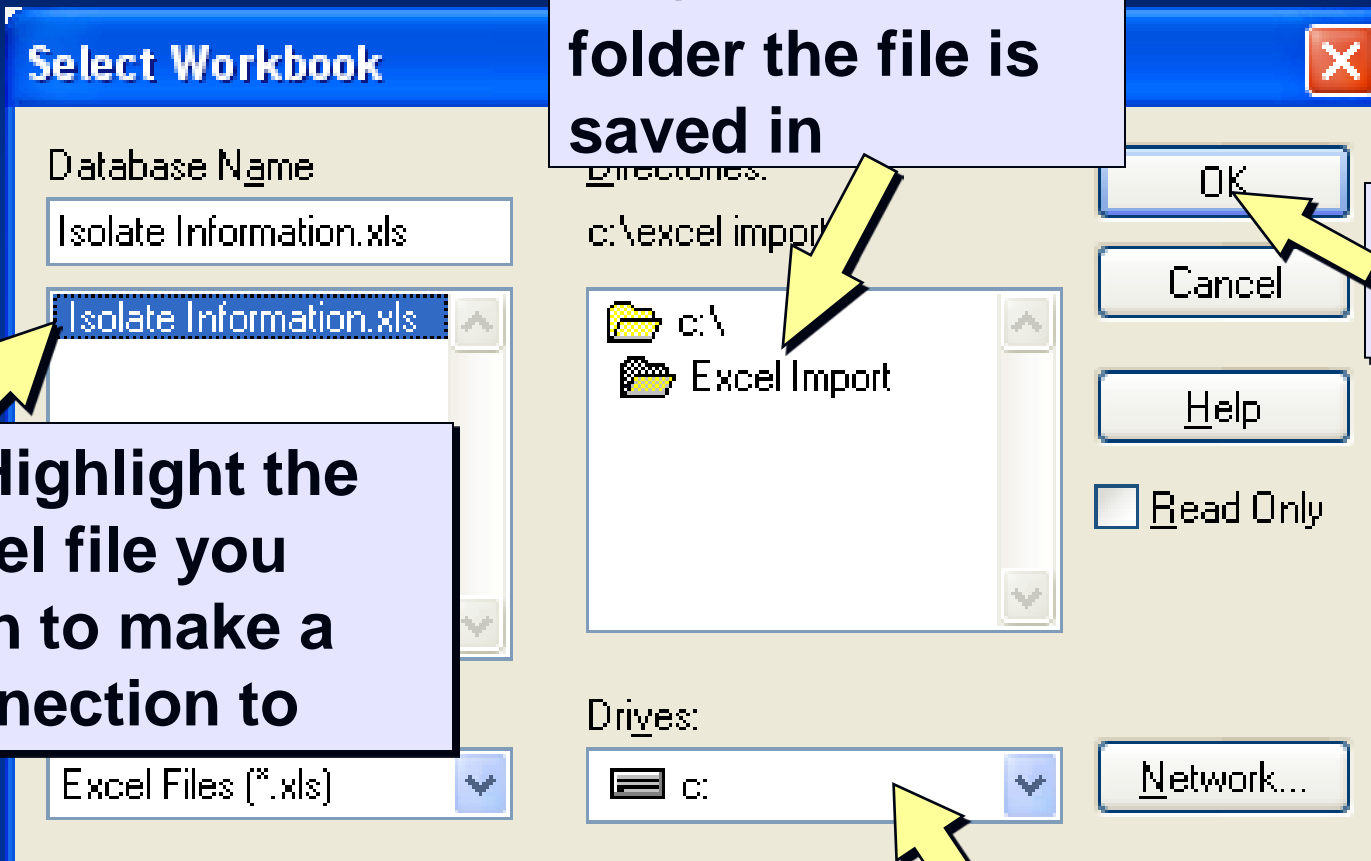
Import: Create ODBC Link

2. Select the folder the file is saved in

4. Press "OK"

3. Highlight the Excel file you wish to make a connection to

1. Navigate to the appropriate drive



Import: Create ODBC Link

ODBC configuration

ODBC data source:

DSN=Excel Files;DBQ=C:\Excel Import\Isolate Information.

Database table or query:

2. Press
"OK"

1. Enter the name of the sheet with a
'\$' sign added (ex. sheet1\$)

Import: Create ODBC Link

ODBC database import

BioNumerics	External database
Key	Isolate Number
LabID	
PatientAge	Patient Age
PatientSex	
UploadDate	
Serotype	
SourceType	
SourceSite	

Serotype

Link Unlink Ok Cancel

1. Select a BioNumerics field in the left column and the corresponding Excel field in the right column

2. Press "Link" when both fields are highlighted

Import: Create ODBC Link

BioNumerics	External database
Key	Isolate Number
LabID	
PatientAge	Patient Age
PatientSex	
UploadDate	
Serotype	Serotype
SourceType	
SourceSite	

Link Unlink Ok Cancel

Press "OK"
after all
appropriate
fields are
linked

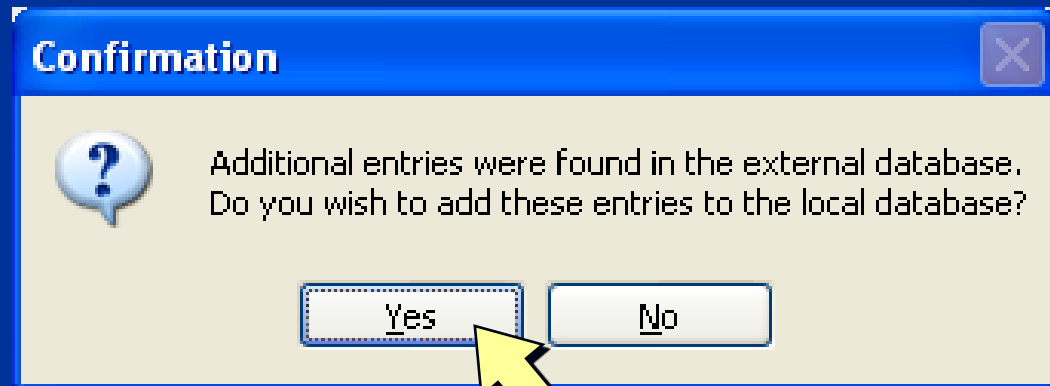
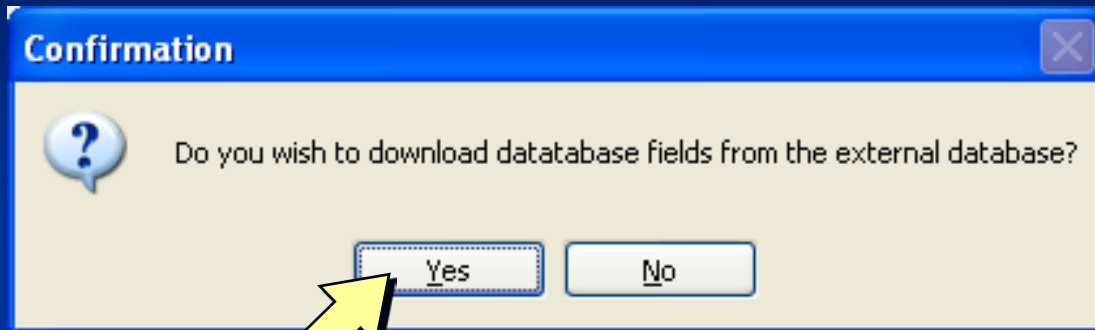
Import: Download Data

The screenshot shows the BioNumerics software interface. The 'Database' menu is open, and the 'ODBC link' option is selected, which has opened a sub-menu. In this sub-menu, the 'Copy from external database...' option is highlighted. The background shows a data table with columns for 'Date' and 'Serotype'. The 'Serotype' column contains entries like 'Newport', 'Montevideo', and 'Heidelberg'.

Date	Serotype
	Newport
	Newport
	Newport
	Newport
	Newport
	Newport
14	067-12834-12521
15	067-13063
16	067-13063b
17	067-15623-12632
18	07-39790
19	07-39791

Select Database → ODBC link → Copy from external database...

Import: Download Data



Now the data from Excel is present in your BioNumerics database

Import: Remove Timestamps

The screenshot shows the BioNumerics software interface. The 'PulseNet' menu is open, and 'Standardize Date' is selected. A yellow arrow points to this menu item. Below, a 'Standardize Dates' dialog box is open, with 'ReceivedDate' selected in the list of fields. A yellow arrow points to 'ReceivedDate', and another yellow arrow points to the 'OK' button.

Database entries

IsolatDate	ReceivedDate	UploadDate	UploadModifiedDate	AntigenForm	...
2009-05-01	2009-05-06 00:00:00	2009-05-13	2009-05-13		
2009-05-02	2009-05-05 00:00:00	2009-05-13			
2009-04-24	2009-05-05 00:00:00	2009-05-13			
2009-04-28	2009-05-05 00:00:00	2009-05-13			
2009-04-29	2009-05-06 00:00:00	2009-05-13			
2009-04-25	2009-05-06 00:00:00	2009-05-13			
2009-04-27	2009-05-06 00:00:00	2009-05-13			
2009-04-30	2009-05-06 00:00:00	2009-05-13			
2009-05-01	2009-05-06 00:00:00	2009-05-13			
2009-04-29	2009-05-06 00:00:00	2009-05-13			
2009-04-29	2009-05-06 00:00:00	2009-05-13			
2000-11-30		2006-02-27	2006-04-26		

Standardize Dates

Select field to change:

- ReceivedDate
- IsolatDate
- PFGE-Xbal-rundate
- PFGE-Blnl-rundate
- PFGE-Spel-rundate

OK Cancel

Select PulseNet → Select Standardize Date →
Select field to change → Press OK

Import: Remove Timestamps

The screenshot shows the BioNumerics interface with a table of database entries. A confirmation dialog box is overlaid on the table, asking for confirmation to change the 'ReceivedDate' field for 11 selected entries. A yellow arrow points to the 'OK' button in the dialog box.

IsolatDate	ReceivedDate	UploadDate	UploadModifiedDate	AntigenForm
2009-05-01	2009-05-06 00:00:00	2009-05-13		
2009-05-02	2009-05-05 00:00:00	2009-05-13		
2009-04-24	2009-05-05 00:00:00	2009-05-13		
2009-04-28	2009-05-05 00:00:00	2009-05-13		
2009-04-29	2009-05-06 00:00:00	2009-05-13		
2009-04-25	2009-05-06 00:00:00	2009-05-13		
2009-04-27	2009-05-06 00:00:00	2009-05-13		
2009-04-30	2009-05-06 00:00:00	2009-05-13		
2009-05-01	2009-05-06 00:00:00	2009-05-13		
2009-04-29	2009-05-06 00:00:00	2009-05-13		
2009-04-29	2009-05-06 00:00:00	2009-05-13		


Confirm number of entries to be changed for the selected date field and press OK.

Now the timestamps are gone

The screenshot shows the BioNumerics interface with the same table of database entries. The 'ReceivedDate' column now only contains dates without timestamps, indicating that the timestamps have been removed.

IsolatDate	ReceivedDate	UploadDate	UploadModifiedDate
2009-05-01	2009-05-06	2009-05-13	2009-05-13
2009-05-02	2009-05-05	2009-05-13	2009-05-13
2009-04-24	2009-05-05	2009-05-13	2009-05-13
2009-04-28	2009-05-05	2009-05-13	2009-05-13
2009-04-29	2009-05-06	2009-05-13	2009-05-13
2009-04-25	2009-05-06	2009-05-13	2009-05-13
2009-04-27	2009-05-06	2009-05-13	2009-05-13
2009-04-30	2009-05-06	2009-05-13	2009-05-13
2009-05-01	2009-05-06	2009-05-13	2009-05-13
2009-04-29	2009-05-06	2009-05-13	2009-05-13
2009-04-29	2009-05-06	2009-05-13	2009-05-13

Import Data into BioNumerics

- Link imported isolates to the appropriate lane on the analyzed TIFF
- Check for spelling errors for serotype and demographic information
- Click on  to make sure:
 - No extra spaces
 - No spelling errors
 - Information matches the pull-down menus
 - Information is in the proper format
 - i.e. YYYY-MM-DD date format and 00:00:00 does NOT appear after the date
- Download can be performed each time the excel file is modified

Questions?



Thank you for your attention

The findings and conclusions in this presentation are those of the author and do not necessarily represent the views of the Centers for Disease Control and Prevention

Exercise 7