

Clustering *Campylobacter* strains based on antibiotics data using the BioNumerics® software

K. Janssens, J. Dombrecht, B. Pot, L. Vauterin
Applied Maths NV, Keistraat 120, B-9830 Sint-Martens-Latem, Belgium

Introduction: Various methods allow the analysis of large numbers of bacterial isolates for susceptibility to a range of antibiotics. Consistent data storage and detailed statistical analysis of these data, together with proper visualization of overall results and correlations, is crucially important to detect trends and predict future evolutions in resistance. Using a new plugin tool in BioNumerics, cut-off values can be set and antibiotic resistance data are automatically converted to RIS categories (Resistant, Intermediate, Susceptible), which can be used to cluster entries and perform statistical analyses. Using the data from Kinana *et al.* (*AEM*, 72:3309-13; 2006), the workflow for this plugin tool is illustrated.

Add / modify RSI breakpoints in the database: type manually or import from a text file

Agent	Abbr	S	I	R
Ciprofloxacin	CIP	>=21	16-20	<=15
Nalidixic acid	NAL	>=19	14-18	<=13
Tetracycline	TET	>=19	15-18	<=14
Gentamicin	GEN	>=15	13-14	<=12

The three categories are automatically created: Susceptible, Intermediate, Resistant

Import new antibiotic resistance data (MIC values) at any time

	A	B	C
1 Isolate	CIP	NAL	
2 103C	>=21	>=19	
3 104C	>=21	>=19	
4 106B	<=15	<=13	
5 110A	<=15	<=13	
6 110B	<=15	<=13	
7 112A	<=15	<=13	

Isolate	CIP	NAL
103C	>=21	>=19
104C	>=21	>=19
106B	<=15	<=13
110A	<=15	<=13

Automatic data conversion and data storage

Character	Value	Mapping
CIP	3	R
NAL	2	I

Adding breakpoints: In order to convert antibiotic resistance data to categories, the cut-off values for the disk diffusion (zone diameter standards) or MIC values (Minimum Inhibitory Concentrations) need to be defined once. When more information becomes available, cut-off levels can be adjusted, if necessary.

Categories: Based on these cut-off values, BioNumerics automatically creates three categories (called 'mapping') for each antibiotic: 1 = susceptible, 2 = intermediate, and 3 = resistant. Different test methods can be processed and stored for later validation and comparative statistics.

Data import and data conversion: After having defined the breakpoints, the antibiotics data can be imported in the database. The data can be imported from a text file or an ODBC-compatible source (e.g. Microsoft Excel/Access/SQLServer, Oracle, MySQL, PostgreSQL, DB2, etc).

Data conversion: During the import, the data are automatically converted to categories (S, I, or R). The categories for each entry are displayed in the experiment cards: with their values (left) or with their mapping (right).

CLUSTER ANALYSIS

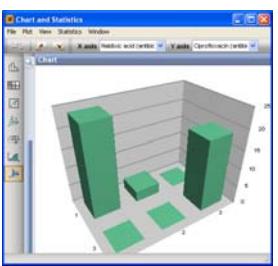
Entries are clustered based on their resistance categories using a categorical coefficient, which treats different values as different states. The colors in the dendrogram window correspond to the color of each antibiotic category (susceptible, intermediate or resistant). MIC values are shown in the central panel. Clusters can be used to create working groups for hypothesis testing. Groups can also be created based on information fields in the database (time of isolation, geography, patient type, hospital, etc...)

CHARTS and STATISTICS

A large number of relevant chart tools are available in BioNumerics, which can directly be applied to the resistance data stored in the database. For each type of chart, one or more standard statistical tests are available.

Examples below: Contingency table for nalidixic acid and ciprofloxacin values, 3D view of contingency table, Chi square test for contingency table.

	CIP	IP	
Nalidixic acid (antibio-ZD_SIR)	24	3	0
Ciprofloxacin (antibio-ZD_SIR)	0	0	19



Chi square: 48.000 (2 degree of freedom)
P value: 0.000000
Significance: 100.000%

Contingency table:

	CIP	IP	Total count
Nalidixic acid	24	3	27
Ciprofloxacin	0	19	19
Total count	24	22	46

Average cell count: 7.67

Marginal column counts:

CIP	24	0	24
IP	3	19	22

Marginal row counts:

Nalidixic acid	24	3	27
Ciprofloxacin	0	19	19