

miniTUBA Tutorial

miniTUBA is a web-based modeling system that allows clinical and biomedical researchers to perform complex medical/clinical inference and prediction using dynamic Bayesian network analysis with temporal datasets. The software allows users to continuously update their data and refine their results. miniTUBA can make temporal predictions to suggest interventions based on an automated learning process pipeline using all data provided. A detailed step-by-step walk through is provided as follows.

1. Get started

Point your browser to <http://www.minituba.org> as shown in Figure 1. To start your own research project, click on “Research Projects”. To test all the features in miniTUBA using our demo account, click on "Sandbox Demo".

Differences between “Demo” & “Research Projects”:

- (1) “Research Projects” need individual account.
- (2) “Research Projects” need approval.
- (3) “Demo” projects are public, no privacy.

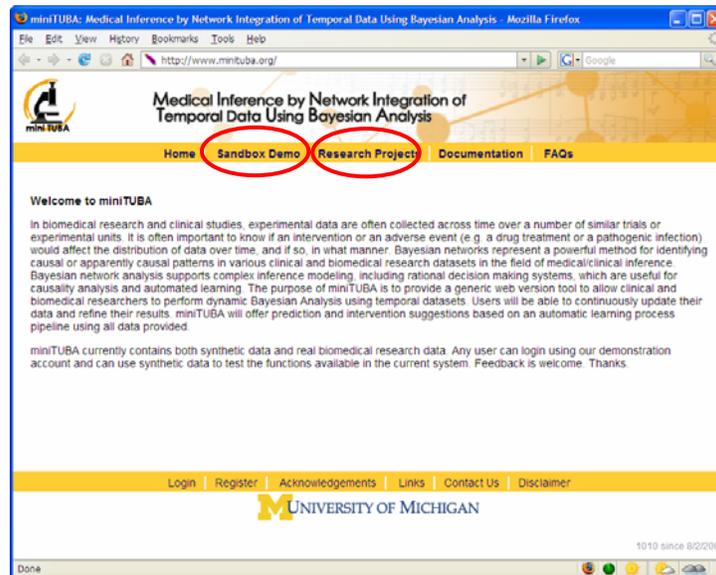


Figure 1

The "Sandbox Demo" (Figure 2) or “Research Projects” (Figure 3) web page lists many projects publicly viewable:

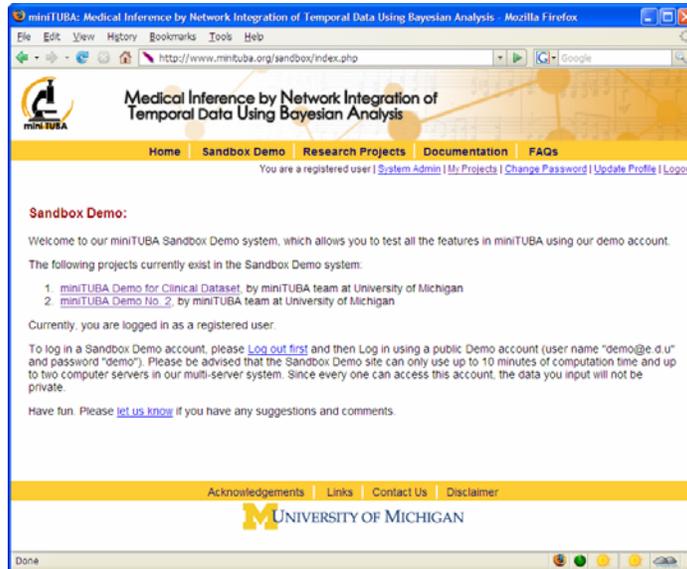


Figure 2

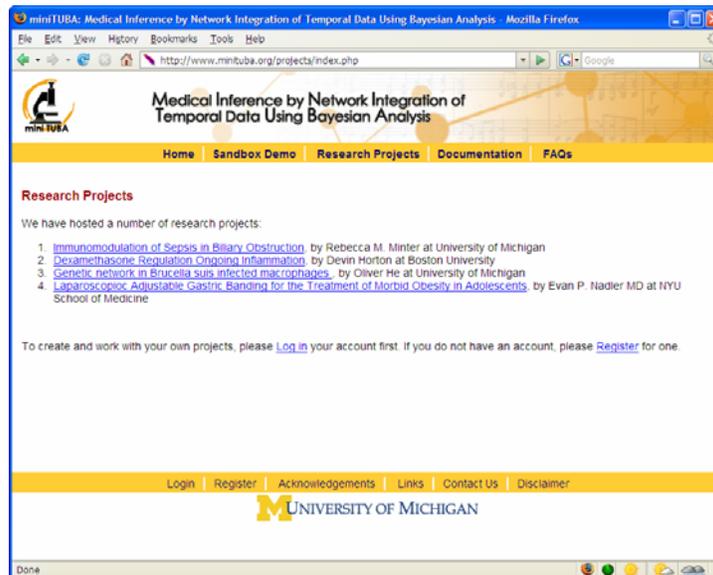


Figure 3

Each project shown in the "Sandbox Demo" or "Research Projects" web page is described (Figure 4). However, the project owner may decide not to show all the operations.

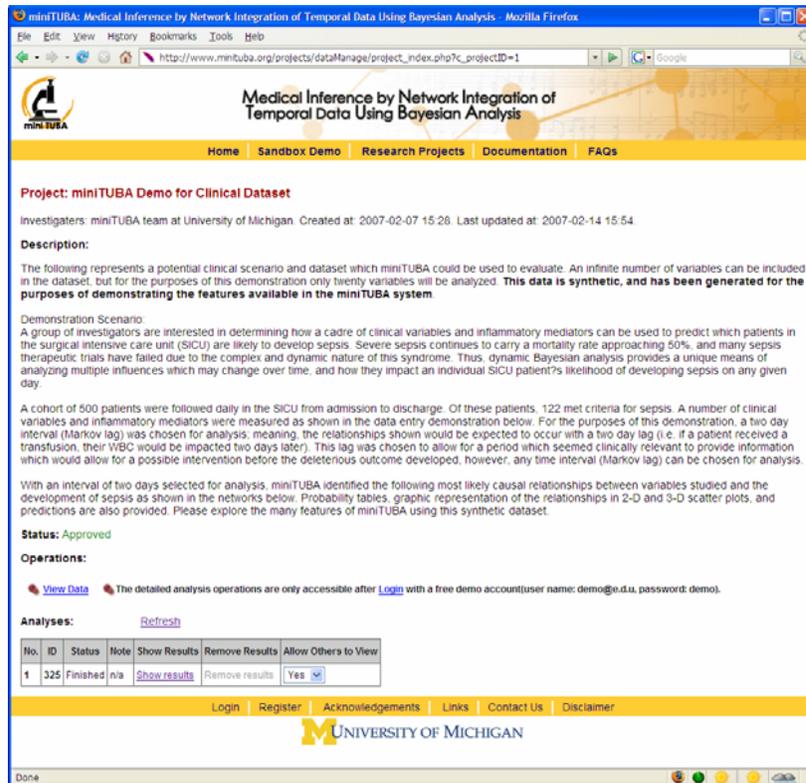


Figure 4

2. Register your account:

Before setting up your own research project, you need to register an account in miniTUBA and login to our system. To register, click "Register" and fill out the form shown in Figure 5. Once your registration is finished, you can log in and start to create your new project.

NOTE: You don't need to register for an account if you just want to test the features using our "Sandbox Demo". Instead, you can use login with user name "demo@e.d.u" and password "demo".

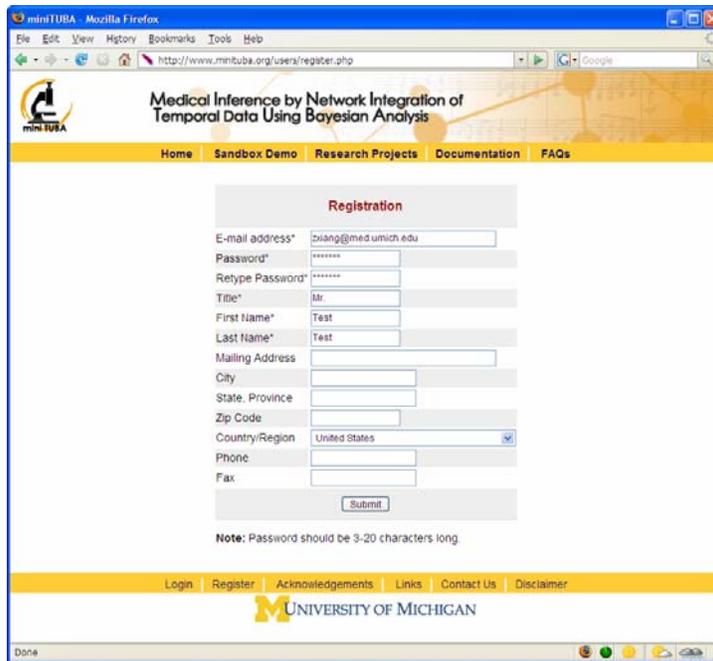


Figure 5

3. Create your own project

Once logged in, you will see a page listing all of your projects if you have any. You can create a new project as instructed (Figure 6).

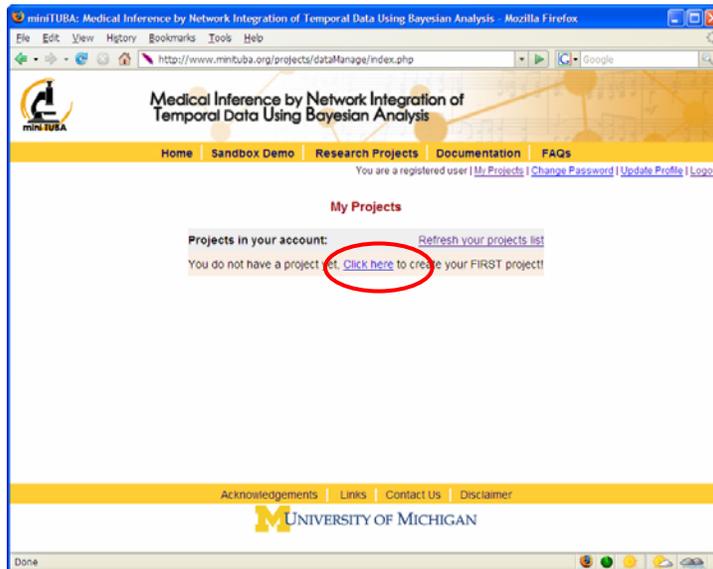


Figure 6

To create a new project, please fill out the form seen in Figure 7.

Medical Inference by Network Integration of Temporal Data Using Bayesian Analysis

Home | Sandbox Demo | Research Projects | Documentation | FAQs

You are a registered user | [My Projects](#) | [Change Password](#) | [Update Profile](#) | [Logout](#)

Create a New Project

Please fill out the following form (all fields are required) and click submit. We will review your project and give you the permission to upload/analysis your data if your project is proper for our system. You be notified by email once your project is authorized.

Title of your project*
Testing project for tutorial

Institution(s)*
University of Michigan

Investigator(s) *
Allen Xiang

Grant Support

Description*
Please enter a description about your project.
This is a testing project for tutorial.

If you do not wish to have your project listed on miniTUBA, check here. Data will be viewable to you alone unless you indicate otherwise.

How many variables are measured for each experimental unit?*
(i.e. genes, proteins, cytokines, clinical variables)
20

How many experimental units are there within your dataset?*
(i.e. patients, animals)
200

On average, how many time points are measured in each experimental unit?*
20

*All the fields are required.

Acknowledgements | Links | Contact Us | Disclaimer

UNIVERSITY OF MICHIGAN

Figure 7

We will review your request and contact you within a few days (Figure 8). NOTE: If you are creating a project in “Sandbox Demo”, your project will be approved automatically.

miniTUBA: Medical Inference by Network Integration of Temporal Data Using Bayesian Analysis

Home | Sandbox Demo | Research Projects | Documentation | FAQs

You are a registered user | [My Projects](#) | [Change Password](#) | [Update Profile](#) | [Logout](#)

My Projects

Project Testing project for tutorial created! We will notify you by email once your project got approved.

Projects in your account: [Refresh your projects list](#)

No.	Title	Creation Time	Operations
1	Testing project for tutorial	2007-03-20 12:04	Select/Open Project Remove Project

[Click here](#) to create another new project.

Acknowledgements | Links | Contact Us | Disclaimer

UNIVERSITY OF MICHIGAN

Figure 8

Once your project is approved, you can select and open your project by click “Select/Open project” (Figure 8).

4. Upload data.

To upload project data, click “Load/Update Data” in Figure 9.

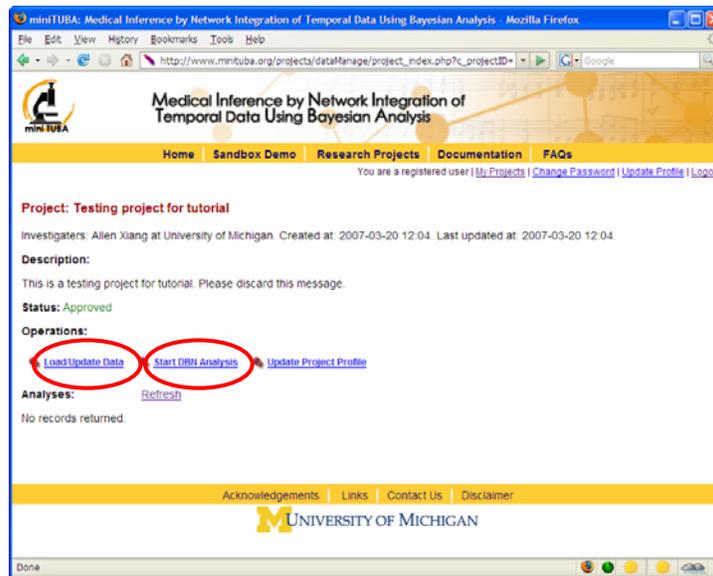


Figure 9

A form will appear and allow you to paste your data into the text box or select the data file you wish to upload (Figure 10).

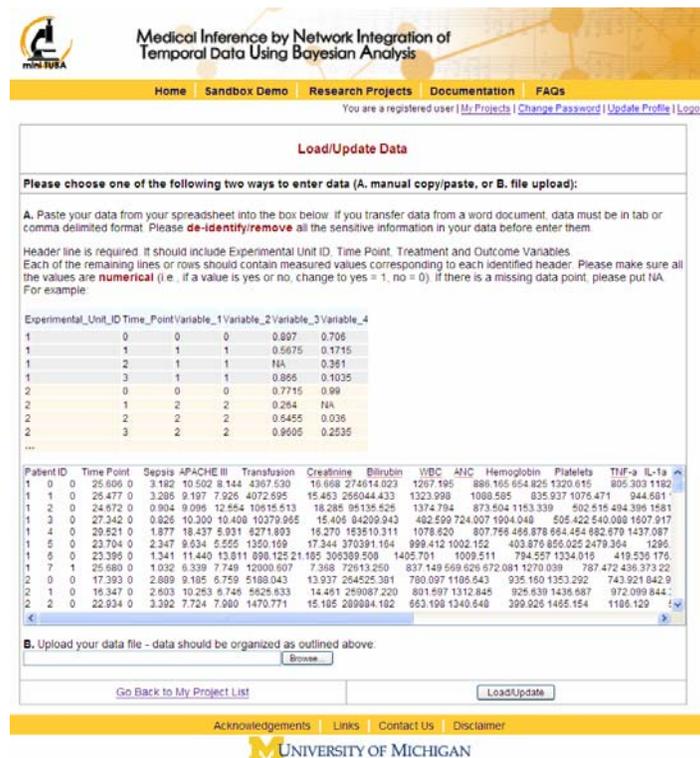


Figure 10

5. Setting up dynamic Bayesian network analysis parameters

- Once you have your data uploaded, you can click “Start DBN Analysis” (Figure 9). A form will appear (Figure 11), where you can set up dynamic Bayesian network (DBN) analysis parameters.

Medical Inference by Network Integration of Temporal Data Using Bayesian Analysis

Home | Sandbox Demo | Research Projects | Documentation | FAQs

You are a registered user | My Projects | Change Password | Update Profile | Logout

Dynamic Bayesian network (DBN) Analysis

Instruction: To run your DBN analysis, please first set up DBN analysis parameters, then click "Run DBN Analysis"...

Select a Setting:

Current setting of DBN analysis parameters:

(1) Which experimental units do you wish to include in the analysis? [Select all](#) [Deselect all](#)

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

(2) Inclusion and characteristics of experimental variables:

Variables to include?	Allow parents or children?	Spline Fitting	Discretization Policy
<input checked="" type="checkbox"/> Septa	Both	No Fitting	No Discretization
<input checked="" type="checkbox"/> APACHE II	Both	No Fitting	Quantile: 3 bins
<input checked="" type="checkbox"/> Transfusion	Both	No Fitting	No Discretization
<input checked="" type="checkbox"/> Creatinine	Both	No Fitting	Quantile: 3 bins
<input checked="" type="checkbox"/> Bilirubin	Both	No Fitting	Quantile: 3 bins
<input checked="" type="checkbox"/> WBC	Both	No Fitting	Quantile: 3 bins
<input checked="" type="checkbox"/> ANC	Both	No Fitting	Quantile: 3 bins
<input checked="" type="checkbox"/> Hemoglobin	Both	No Fitting	Quantile: 3 bins
<input checked="" type="checkbox"/> Platelets	Both	No Fitting	Quantile: 3 bins
<input checked="" type="checkbox"/> TNF- α	Both	No Fitting	Quantile: 3 bins
<input checked="" type="checkbox"/> IL-1 α	Both	No Fitting	Quantile: 3 bins
<input checked="" type="checkbox"/> IL-1 β	Both	No Fitting	Quantile: 3 bins
<input checked="" type="checkbox"/> IL-5	Both	No Fitting	Quantile: 3 bins
<input checked="" type="checkbox"/> IL-6	Both	No Fitting	Quantile: 3 bins
<input checked="" type="checkbox"/> INF- γ	Both	No Fitting	Quantile: 3 bins
<input checked="" type="checkbox"/> IL-10	Both	No Fitting	Quantile: 3 bins
<input checked="" type="checkbox"/> IL-1 α	Both	No Fitting	Quantile: 3 bins
<input checked="" type="checkbox"/> sTNFi	Both	No Fitting	Quantile: 3 bins
<input checked="" type="checkbox"/> sTNFii	Both	No Fitting	Quantile: 3 bins

(3) Network structure properties

Settings	Edges in the Setting	Option
Must Be Excluded Edges	Please list any relationships which should NOT be allowed in the proposed network. For example: Variable2-->Variable1 Variable3-->Variable1 Variable4-->Variable1	<input type="radio"/> Use this option <input checked="" type="radio"/> Ignore this option
Must Be Present Edges	Please list any relationships which must be present in the proposed network. For example: Variable2-->Variable3 Variable1-->Variable4	<input type="radio"/> Use this option <input checked="" type="radio"/> Ignore this option
Initial Structure	Do you want to include the previous computed network structure in the current analysis? This will save the computation time.	<input type="radio"/> Use this option <input checked="" type="radio"/> Ignore this option

(4) Searcher Choice: Simulated annealing

(5) Number of Instances: 1

(6) Markov Lags: 1

(7) Maximum Query Time: 0 Hours 0 Minutes

(8) Note

[Go back to the front page of current project](#)

Acknowledgements | Links | Contact Us | Disclaimer

UNIVERSITY OF MICHIGAN

Figure 11

NOTES:

- **Select a Setting:** All settings for each analysis will be stored in our database. You can reload settings from any previous analysis by selecting an analysis and click "Reload selected setting".
- **Variables to include:** choose variables to include in the analysis.
- **Allow parents or children:** You can exclude all the relationships start from a variable by select "Parents only", or exclude all the relationships point to a variable by select "Children only"
- **Spline Fitting:** If you have missing data points for continuous data, spline fitting is required to interpolating those missing data. Spline fitting does NOT work for nominal data.
- **Discretization Policy:** If the measured values for a variable are continuous, discretization is required. For interval discretization the data range of each bin is equal. The number of data values in each bin varies according to the bin range. For quantile discretization each bin receives an equal number of data values. The data range of each bin varies according to the data values it contains. For customized binning, please enter the cutting points separated by semicolon. 3 is the suggested number of bins.
- **Must be Excluded Edges:** list any relationships that should NOT be allowed.
- **Must Be Present Edges:** List any relationships that must be included.
- **Initial Structure:** include any relationships as prior knowledge.
- **Searcher Choice:** Simulated Annealing and Greedy approaches are two approaches for searching networks. The Simulated Annealing approach accepts networks based on a stochastic program implementing Metropolis-Hastings. The Greedy approach always searches for networks with better scores. Simulated annealing is the default selection.
- **Number of Instances:** Number of computer nodes you wish to use.
- **Markov Lags:** units of time between two time points to be analyzed in this project. For example, for a project with hourly data sets, Markov lag 1 means 1 hour, and Markov lag 2 means 2 hours, etc.
- **Maximum Query time:** Depending on the number of variables, the total number of time points, and the discretization policy you select, the maximum query time varies. The more variables, time points or bins you have, the longer time you need to specify here.

- Note: You can put some extra note for this analysis.

After all parameters are set, click “Run Bayesian Analysis” (Figure 11).

6. Run DBN analysis

While the dynamic Bayesian network engine is running, you can periodically check the progress by clicking “Refresh this page” (Figure 12).

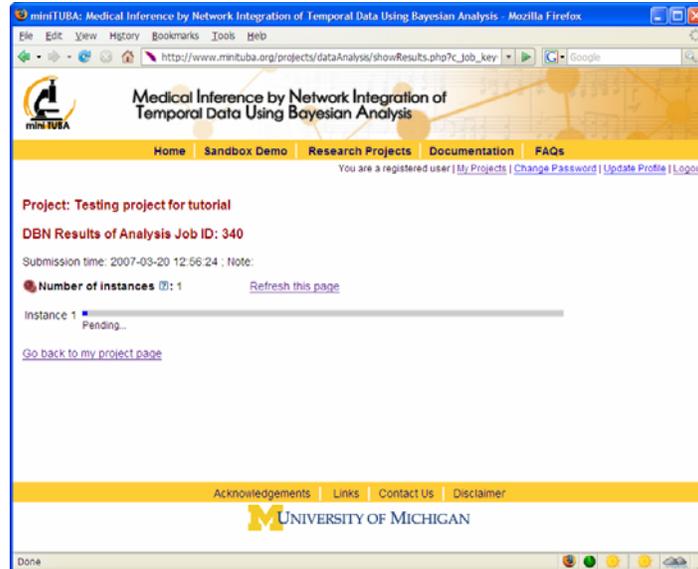


Figure 12

7. Check DBN analysis results

After the DBN analysis is finished, you can check the results (Figure 13).


Medical Inference by Network Integration of Temporal Data Using Bayesian Analysis

[Home](#) | [Sandbox Demo](#) | [Research Projects](#) | [Documentation](#) | [FAQs](#)

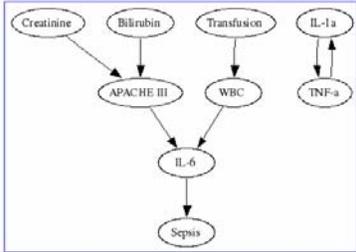
Project: miniTUBA Demo for Clinical Dataset
DBN Results of Analysis Job ID: 325

Submission time: 2007-02-12 14:33:31 | Finish time: 2007-02-12 14:36:12 | Note:

Networks Generated:

Top 10 Networks: [No 1](#) | [No 2](#) | [No 3](#) | [No 4](#) | [No 5](#) | [No 6](#) | [No 7](#) | [No 8](#) | [No 9](#) | [No 10](#)

[Score Distribution of Top 10 Networks](#) | [Conserved Subnetworks among Top 10 Networks](#)



Probability Table and Possible Scatter Plot:
 Please click a variable in above network to check its probability table and possible scatter plot.

Generate Subnetwork: please select a subset from available variables:

in

Note: To select more than one nodes, press "Ctrl" or "Shift" button in the keyboard, hold it and click on specific nodes.
[Go back to my project page](#)

[Login](#) | [Register](#) | [Acknowledgements](#) | [Links](#) | [Contact Us](#) | [Disclaimer](#)



Figure 13

- Top 10 networks: show the top 10 DBN network. Click each one for network display.
- Probability Table and Possible Scatter Plot: Click a variable in above network to check its probability table and possible scatter plot. See Figure 14. For the scatter plot, if it includes three variables, a 3-D image will appear. You can rotate the 3-D image to explore the details.
- Generate Subnetwork: Select a subset of available variables to generate a subset of the above network.

Project: miniTUBA Demo for Clinical Dataset

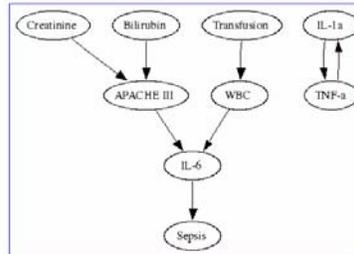
DBN Results of Analysis Job ID: 325

Submission time: 2007-02-12 14:33:31 | Finish time: 2007-02-12 14:36:12 | Note:

Networks Generated:

Top 10 Networks: [No 1](#) | [No 2](#) | [No 3](#) | [No 4](#) | [No 5](#) | [No 6](#) | [No 7](#) | [No 8](#) | [No 9](#) | [No 10](#)

[Score Distribution of Top 10 Networks](#) | [Conserved Subnetworks among Top 10 Networks](#)

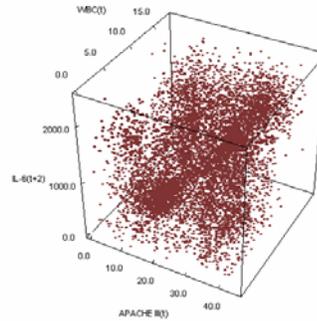


Probability Table and Possible Scatter Plot:

Conditional probability table of IL-6

Time n		Time n+2		
APACHE III	WBC	< 501.000	501.000 - 1500.000	> 1500.000
< 16.000	< 4.000	0.1854	0.4870	0.3276
16.000 - 30.000	< 4.000	0.1978	0.3822	0.4200
> 30.000	< 4.000	0.2290	0.2900	0.5010
< 16.000	4.000 - 11.000	0.8935	0.9522	0.0543
16.000 - 30.000	4.000 - 11.000	0.3220	0.2920	0.3951
> 30.000	4.000 - 11.000	0.2607	0.2546	0.4848
< 16.000	> 11.000	0.2018	0.4183	0.3799
16.000 - 30.000	> 11.000	0.1706	0.2243	0.6051
> 30.000	> 11.000	0.0481	0.0552	0.8997

Scatter plot of IL-6 and it's parent(s). Please drag to rotate (Powered by [LiveGraphics3D](#)).



Generate Subnetwork: please select a subset from available variables:

Sepsis
 APACHE III
 Transfusion
 Creatinine

in [No 1 scoring network](#)
 Show sub network of selected variables

Note: To select more than one nodes, press "Ctrl" or "Shift" button in the keyboard, hold it and click on specific nodes.

[Go back to my project page](#)

Figure 14

8. Run prediction

You can also run prediction. You will be prompted to set up the prediction parameters first. See Figure 15.

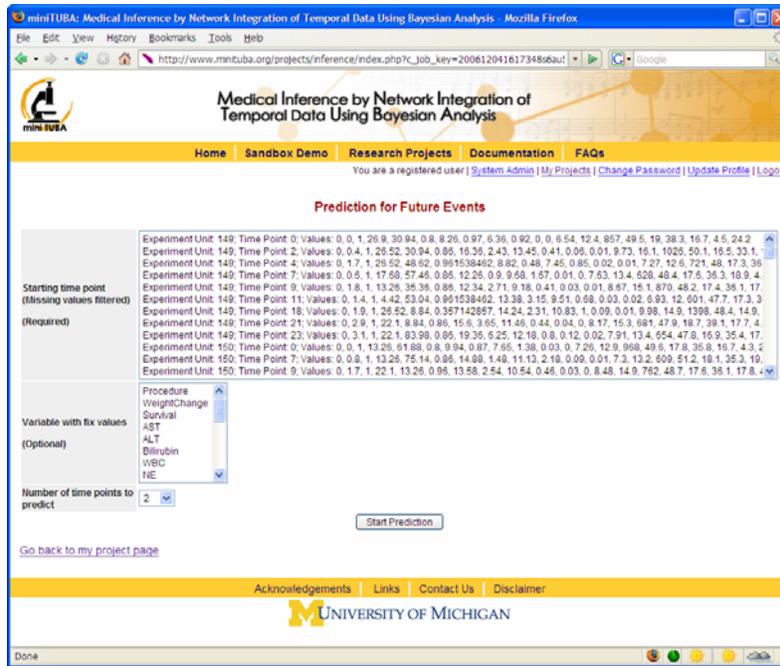


Figure 15

- Starting time point: It is required to set up a starting time point for future prediction. Any time point with missing data can not be selected.
- Variable with fix values: Select variables with fixed values. The miniTUBA prediction engine will not predict any value for these variables. This is optional.
- Number of time points to predict: select the number of time points to be predicted.

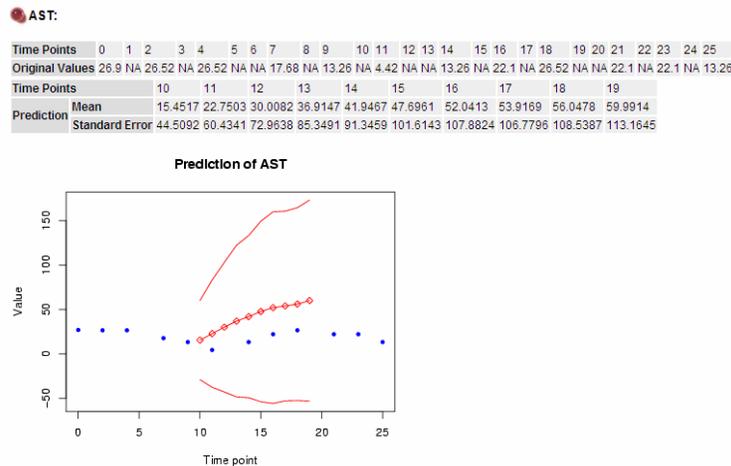


Figure 16

One example of prediction results is shown in Figure 16.

Suggestions and comments are welcome. Thank you!