Modeling and Simulation of Disease Progression in Spinal Muscular Atrophy Pediatric Populations

Jason H. Williams and Jeffrey S. Barrett
Laboratory for Applied PK/PD, Division of Clinical Pharmacology and Therapeutics, The Children's Hospital of Philadelphia, Philadelphia, PA

BACKGROUND

Spinal Muscular Atrophy (SMA)
- Most frequent genetic cause of infant death
- Incidence is 1:10,000; carrier frequency is 1:55

Pathology of Disease
- Rapid Loss of Spinal Motor Neurons
- Clinical Manifestation
- Proximal muscle weakness, respiratory dysfunction, scoliosis, loss of ambulation, feeding tube, death, or continuous ventilation

OBJECTIVES

1. Develop a model to predict clinical outcomes in SMA
2. Combine SNM2 distribution with model learned from data to generate a prior distribution useful for future disease progression modeling of longitudinal data.

METHODS

3. Expectation-Maximization (EM) algorithm implemented in Netica (Norsys Corp) for calculating Conditional Probability Distributions (CPDs). Efficient computation of the Joint Probability Distribution (JPD) across all model parameters was possible using the following three conditional independence relationships:

RESULTS

Incorporating SMN2 copy number distribution into Bayesian Network model learned from SMA dataset

Goal of Simulation: To verify the Bayesian Network calculates correct distribution in causal direction encountered during reversal of edges.

Step 1: Create CPD table using data obtained from Meta-analysis of 541 SMA patients (top-left, Ogino et al. 2004). Set uniform distribution of SMA subtype.

Step 2: Simulate 100,000 SMA patients with equal type I, II, III distribution (N=3335/3340/33249 of Type VIII) to generate a dataset of SMN2 copy number variation to be randomly sampled.

Step 3: Random sample (Sphas) from 100,000 patients and equal number of copy number 1x, 2x, 3x, 4x (1000 each).

Step 4: Calculate distribution of SMA subtype conditional on SMN2 copy number.

Results: Histogram (top right) shows very similar distribution as the BN visually shows thus edge reversal is efficient for generating model in causal direction (EXP-ML19, bottom-left).

CONCLUSIONS

1. Efficient computation of CPDs from a sparse SMA dataset is possible using Bayesian Network approach.
2. Incorporating Expert domain knowledge from the SMA literature or panel of clinicians can be incorporated into an SMA model learned from data.
3. Limitations remain when manually assembling a Bayesian Network structure and thus structure learning algorithms need to be evaluated.

FUTURE DIRECTIONS

Temporal (aka Dynamic) Bayesian Networks
- Modeling Longitudinal Outcomes from Natural History Study of SMA (Children's Hospital of Philadelphia; PNCR study group)

ACKNOWLEDGEMENTS

We thank Connie Garland for help with the data collection from the SMA patient registry.

REFERENCES