OBJECTIVES

We describe an ontology-based method for representation and integration of neuroscience data, with an application to event-related brain potentials (ERPs). ERP datasets from different labs are highly heterogeneous, in that they are acquired in different experiment contexts and may be analyzed and described using different methods. To address this issue, we use data mining technique coupled with ontology as a semantic bridge to discover mappings across different spatial and temporal metrics characterizing datasets from different laboratories. A novel application of sequence similarity search techniques has also been proposed in the process of mapping discovery. The mappings and ontologies are being used to support data sharing and cross-laboratory analysis of ERP experiment results within our NEMO (Neural Electromagnetic Ontologies) consortium [2].

METHODS I: SIMULATED ERP DATASETS

Simulated ERP data were designed to model ERPs acquired from 128 EEG channels, 40 “subjects,” and two “conditions.” The data were derived from a 7-dipole, 3-shell spherical model using Dipole Simulator. Dipoles were located in occipital, posterior-temporal, and mid-cuneus regions (Fig. 1A), representing 4 patterns: 1) P100, 2) N100, 3) N1/N2, & 4) P2 (Fig. 1B). The data were projected to 129 scalp locations, with time series distributed over ~600 ms. (Fig. 1C). Individual “subject” ERPs were designed by adding random offsets to the base intensities. Noise was superimposed on the simulated ERP data for each of the 80 individual datasets (courtesy of P. Berg).

METHODS II: ERP PATTERN ANALYSIS

The data were decomposed using spatial Independent Components Analysis (sICA), which yielded a small set of discrete temporally independent latent patterns (Fig. 2). Patterns were automatically classified and labeled (e.g., “P100” or “N100”) using data-driven methods described in [1].

METHODS III: EXTRACTION OF ALTERNATIVE METRICS

For each latent pattern, we extracted an alternative set of summary metrics that provided an equivalent characterization of its spatiotemporal dimensions. The alternative measures simulate the case where research groups use different, yet comparable, measures to quantify the temporal and spatial attributes of their data.

SUMMARY & CONCLUSIONS

We describe a framework for mapping cross-laboratory ERP data. Contributions include:

- Using ontology annotation as a semantic bridge between heterogeneous ERP data.
- Recasting the mapping problem to the sequence similarity search problem
- Results suggest that the proposed method is accurate and robust.
- The ensemble method achieved an performance of 76% precision and 100% recall.
- The mapping result is consistent throughout multiple samples of random ordering.

REFERENCES


ACKNOWLEDGMENTS

This work was supported by an R01 grant to the University of Oregon by the National Institute of Biomedical Imaging and Bioengineering (NIBIB) at the National Institutes of Health (NIH), Award #R01 EB007684-01. We thank our NEMO consortium for helpful comments and ongoing support of this project.