Using MetaMap to Process a List of Terms

MetaMap by default expects input records to be separated by a blank line. Many users need to process lists of short terms, e.g.,

Written informed consent. Age 18 years or older. informed consent. Male or female. At least 18 years of age. Able to give informed consent. Signed informed consent. Healthy. Male. Negative pregnancy test. Signed written informed consent. Ability to understand document. Measurable disease. ECOG performance status 0-2. Signed informed consent. Adult patients, >/= 18 years of age. Type 2 diabetes. Other protocol-defined criteria.

To process such a list of terms without inserting a blank line between each term, call MetaMap with the --sldi (Single-Line Delimited Input) option. MetaMap will then treat each phrase as a separate input record, rather than processing all the text as one monolithic block.

To ensure that each term is processed as a single phrase, and generates the best possible results, we also recommend using these options:

- -z (--term_processing)—explained Term Processing,
- -i (--ignore_word_order)-explained Ignore Word Order, and
- -C (--relaxed_model)-explained Relaxed Model.

The last line of all MetaMap input files must end with a newline character!