The following are data handling instructions for the Public Use Dataset of the NIDDK-funded Evaluating Predictors & Interventions in Sphincter of Oddi Dysfunction: EPISOD 3 A Long Term Follow-Up Study which followed the EPISOD and EPISOD2 (NCT 00688662) patients for up to 5 years. The dataset is comprised of patients enrolled into the study.

General Notes
- Files
  - ReadMe PDF file (pdf)
  - 13 dataset SAS files: including 10 CRF datasets, 1 derived dataset, 1 data dictionary and 1 code list SAS file
  - 13 dataset CSV files: including 10 CRF datasets, 1 derived dataset, 1 data dictionary and 1 code list SAS file
  - Contents file – pdf of proc contents by form
  - Data Collection Schedule (pdf)
  - 10 Annotated Case Report Forms pdf file
- Datasets
  - A complete listing of all variables, their description and their codes may be found in the data dictionary file and code list file.
  - There is a separate dataset for each included form.
  - Forms that were collected only once during the study will have one row per subject.
  - Forms that were collected multiple times during the study will have more than one row per subject.
  - Forms that contain logs may have more than one row per subject.
  - A complete listing of which forms were collected at each study visit can be found in "EPISOD3 Data Collection Schedule".
  - Four CRF datasets were excluded from this public use dataset:
    - Forms 34 captured patient contact information
    - Form 47 captured eligibility and informed consent information at baseline as was never used for data analysis or reporting
    - Forms 48 and 49 captured information on the concomitant medication and treatments during follow up. Due to the complexity of this form, the information that was used for the EPISOD3 analysis was derived and included in the sas and CSV file titled, 'Derived'
- Subject and Site IDs
  - Subject and Site IDs are unique identifiers for each subject and are not the original ID used by the study team and match the IDs used in EPISOD PUDs.
- Variable names
  - Raw data are named in the format F##Q##, where F## is the form number and Q## is the question number. For multiple choice questions (ex. Question01 on Form55), the data are named in the format F##Q##M#, where M# is the option number. The data dictionary provides the variable names. For the date variables, the date in CRFs are named in the format F##Q##_d and the date of form collected are named in the format Formdate_F##.
  - Derived variables are included in some datasets, indicated below in Dataset Specific Notes.

- Dates and times
  - All dates are replaced with calculated days from the start date (the randomization date for EPISOD patients or the initial treatment date for EPISOD2 patients). These variables are expressed as the number of days from the start date.
  - Negative times indicate the event occurred before the start date.
  - If the day of a date is missing, the date is treated as missing since the scenario is rare.
- Text fields, including clinical narratives and general comments, are not included in the public use dataset in order to avoid releasing potential identifying patient information.
- Visit names were excluded in the datasets due to the wide ranges of the visits happened. The days of form assessment from the start date were provided.

Dataset Specific Notes

- Derived: dataset including Subject ID, Site ID, the study number(EPISOD OR EPOSOD2), the treatment arm randomized ('ztreatmentnm'), the real treatment received ('real_trt'), the re-intervention indicator('rtrt_v') and the medication violation indicator ('meds_v'), the indictor for success by RAPID after 1st re-intervention ('aft_reint_S'), the indictor for success by PGIC after 1st re-intervention ('aft_reint_pgic_S'), the indictor for success by PGIC at 1 year and final visit ('epi15_pgic_s'), the indictor for success by RAPID at 1 year and final visit ('epi15_s'), the indictor for success by RAPID at 1 year only ('status'), the indictor for success by PGIC at final visit only ('epi5_LV_pgic_s'), the indictor for success by RAPID at final visit only ('epi5_LV_s').
- Form 12: RAPID
- Form 14: HADS
- Form 15: SF-36 (v2)
  - Questions 23-31 used the official version1 SF-36 form.
- Form 16: Rome III Diagnostic Module
- Form 50: Follow Up Resource Utilization Questionnaire
- Form 51: Adverse Events
- Form 52: End of Study
- Form 53: Coping
- Form 54: Patients' Global Impression of Change (PGIC)
- Form 55: Pain Questionnaire

NOTE: These SAS data files are based on EPISOD3+2 study database frozen on Dec 19th, 2016. All potential identifying fields have been removed.