

How To Add Data to an Existing Study in calIntegrator - Step 2.0

2.0. Now we're ready to load additional subject annotation data into the 'Demo Study for ICR Folks'. As mentioned before, you'll need the data in the form of a CSV file containing at least one field with a unique ID for each subject in the study. The CSV file we'll use in this tutorial is called 'subject_annotations_tutorial.CSV'. A partial screenshot of the file appears below as viewed in a Microsoft Excel 2007 window.

	A	B	C	D	E	F	G	H	I	J	K
1	PATIENT_ID	Stratagene	SITE	GENDER	Survival in Months						
2	5000	1	MI	Female	64.8						
3	5001	1	MI	Female	33.9						
4	5002	1	MI	Female	5.8						
5	5003	1	MI	Male	50.2						
6	5004	1	MI	Female	34.8						
7	5005	1	MI	Male	71						
8	5006	1	MI	Female	9						
9	5007	1	MI	Female	44.94						
10	5008	1	MI	Male	34.76						
11	5009	1	MI	Female	15						
12	5010	0	MI	Female	64.8						

This data came from a fictional multi-site study that compared gene expression between lung adenocarcinoma patients and healthy controls. The nature of the data itself is irrelevant to our purpose here. The relevant aspect is that the data is categorized into five fields, which are represented by columns in the spreadsheet.

Each field defines a different subject characteristic such as 'PATIENT_ID', which uniquely identifies each of the 100 subjects in this study (note that the screenshot above only displays data for the first 11 subjects). Once we've loaded the data into the study, we'll be able to query it by any of the fields.