**Data details** – James (jingham@liverpool.ac.uk)

Email me if there are any further problems/questions.

Janet has highlighted two images which are a good place to start, A12 and G13. The FTIR data is saved as .npy files for ease (in a 3D format) and the masking areas are simple .png’s which can be loaded via imageio (or many other ways). The masked areas are areas which have been professional labelled as tumours tissue.

A12 – A core from a tumour tissue sample of a patient who went on to live for more than 10 years after diagnosis.

G13 – A core from a tumour tissue sample of a patient who died in less than 2 years after diagnosis.

The spectral data is contained within the XXX\_FTIR.npy files. The 1st and 2nd dimension are spatial and the 3rd the spectral axis as we saw last week.

Both images were collected over the same wavenumber range, so I have only included one wavenumber set (wavenumbers.npy) which is identical for both.

I have included the H&E (Hematoxylin and eosin stained) images for reference but note these are adjacent tissue slices so may have small variations between the FTIR and the H&E image.

The samples imaged with the FITR are still paraffin wax embedded which is the reason there is features in the spectra within the areas surrounding the tissue. This wax will permeate the sample and will have a contribution in the tissue spectra to varying degrees, so the standard approach is to remove areas of the spectra associated to paraffin absorption (talked about in more detail in the paper I gave you) before the analysis. This wax has been removed in the H&E images.

Both tumours had 'poor prognosis' clinical indicators and received similar treatment:  Both were large tumours (pT4) with extranodal extension (spread outside the cervical lymph node capsule). Both received surgery followed by chemoradiation.

I have added a notable review paper discussing FTIR of tissue samples from a group we work with in Manchester, it should do a good job of providing a general background.