CRUK Cambridge Centre Clinical PhD project

PhD Project Title	**Computational pathology of cytosponge samples for the early detection of oesophaegeal cancer**
Head of Laboratory (PI) Name	Florian Markowetz
Second supervisor if applicable	Rebecca Fitzgerald
Programme	Early Detection
Supervisor's Email	Florian.Markowetz@cruk.cam.ac.uk
Laboratory Location	CRUK Cambridge Institute MRC Cancer Unit

Project Outline	**Aims and objectives** Oesophageal adenocarcinoma has increased 7-fold in the western world over the past three decades. Most patients present with advanced disease leading to an overall 5-year survival of <20%. However, if the cancer can be diagnosed early, survival improves markedly, such that >80% patients with superficial (stage T1) disease survive beyond 5 years. The disease has a clear pre-cancer stage, Barrett's oesophagus making early detection feasible. Currently diagnosis of Barrett's depends on endoscopy and random biopsies with the result that most patients at risk are not diagnosed and assessment of risk hinges on a subjective diagnosis of dysplasia. The genomic architecture of this cancer is now being delineated and data suggests that mutations occur very early in the disease with a large degree of inter- and interpatient diversity (Nature Genetics 2014, 2016). The Fitzgerald laboratory have taken a completely novel approach to tackle early diagnosis of this disease and have developed a non-invasive, nonendoscopic diagnostic test for diagnosing Barrett's oesophagus using a cell collection device called the Cytosponge™ which is coupled with a TFF3 immunohistochemical assay to diagnose Barrett's (PLOS Med 2015). Cytological atypia and p53 over expression are features predicting cancer risk which can be diagnosed on the same sample (Nature Genetics 2015, 2016, Lancet Gastro & Hepatol 2017). **Aims of the Project:** 1. Develop and test a computational algorithm for assessing the cytological and tissue features of Cytosponge samples in order to diagnose and risk stratify Barrett's oesophagus; 2. Evaluate the procedure on existing cohort and then apply to new cases; 3. Automate the analysis in a standalone software.

Experimental plan	**Clinical resource** There is a large archive of H&E and immunostained slides (TFF3 and TP53) on over 3,000 cases from the BEST Cytosponge trials. There are also matched biopsy samples and DNA sequencing data available from many of these cases. Additionally, the Cytosponge is currently being tested in a very large cluster randomised trial of 9,000 patients in the UK and so this project has the opportunity for direct clinical impact. **Computational methods** We will use segmentation methods to identify cell nuclei and borders. We will train supervised classification methods to identify different cell types and tissue structures. These approaches will provide quantitative summaries of the Cytosponge samples, which we will use in predictive analyses on the existing 3,000 cases to select of high-quality biomarker panel (using cross-validation and resampling techniques). This panel can then be applied to the 9,000 new cases. We aim to provide the biomarker evaluation through a standalone software (tentatively called 'CytoSponge SquarePants') to make it easy to use in clinical settings.
Main Techniques	• quantitative image analysis of cytosponge samples (Sci Trans Med 2012) • combined Image and genomic analysis (Genome Biol 2014) • application of statistical learning methods to integrate clinical, imaging and genomics data
Key References	**1.Weaver JMJ et al.** *Ordering of mutations in preinvasive disease stages of esophageal carcinogenesis. Nat Genet. 2014 Aug;46(8):837-843.* **2.Ross-Innes CS et al.** *Whole-genome sequencing provides new insights into the clonal architecture of Barrett's esophagus and esophageal adenocarcinoma. Nat Genet. 2015 Sep;47(9):1038-1046.* **3.Secrier M et al.** *Mutational signatures in esophageal adenocarcinoma define etiologically distinct subgroups with therapeutic relevance. Nat Genet. 2016 Oct;48(10):1131-41.* **4.Elliott DR et al**. *A non-endoscopic device to sample the oesophageal microbiota: a case-control study. Lancet Gastroenterol Hepatol. 2017 Jan;2(1):32-42.* **5.Yuan Y et al.** *Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. Sci Transl Med. 2012 Oct 24;4(157):157ra143.* **6.Martins FC et al.** *Combined image and genomic analysis of high-grade serous ovarian cancer reveals PTEN loss as a common driver event and prognostic classifier. Genome Biol. 2014 Dec 17;15(12):526.*