THE EDEN STUDY | Early Detection and surveillance of Esophageal cancer with the Cytosponge and Nucleic acid biomarkers

Study Synopsis:

Study Title	**Early Detection and surveillance of Esophageal cancer with the Cytosponge™ and Nucleic acid biomarkers (the EDEN study)**
Principal Investigator	Wladyslaw Januszewicz, MD Department of Gastroenterology, Hepatology and Clinical Oncology, Centre of Postgraduate Education, Poland e-mail: wjanuszewicz@cmkp.edu.pl mobile: +48 502-569-503
Research Units:	1) Centre of Postgraduate Education, Warsaw, Poland (CMKP) 2) National Institute of Oncology, Warsaw, Poland (NIO) 3) The Medical Research Council Cancer Unit, Cambridge University, Wielka Brytania (CAM)
Funding	The National Science Centre grant OPUS-19 (2020/37/B/NZ5/04003)
Clinical Trial Identifier (clinicaltrials.gov)	NCT04192695
Study Design	Multi-center cohort study with a retrospective and a prospective arm
Primary Objective	**Retrospective arm:** • Identification of molecular abnormalities at each developmental stage of the esophageal squamous cell carcinoma (LG-IEN, HG-IEN and ESCC) to establish new molecular biomarkers with potential for early detection and surveillance using the minimally invasive Cytosponge™ cell collection device. **Prospective arm:** • Evaluation of the diagnostic assay established during the retrospective arm (sensitivity and specificity) in patients with ESCC, or at high risk for ESCC, using prospectively collected tissue samples from the Cytosponge™ and upper endoscopy with biopsies.
Secondary Objectives	• Assessment of safety and acceptability of the Cytosponge™ device in patients with ESCC and at high-risk for ESCC through active monitoring of adverse events and the procedure acceptability questionnaire with a Visual Analogue Scale (VAS).

Sample Size	No formal sample size calculations are possible due to the exploratory nature of the study
Inclusion Criteria	**Retrospective arm:** • Tissue specimens from patients ≥18 years old after endoscopic treatment (EMR/ESD) • Written informed consent given for samples to be used in research • Diagnosis: LG-IEN, HG-IEN, ESCC pT1a/pT1b in TNM classification • Lesion size: ≥1cm • Single tissue fragment (*en-bloc*) <u>Prospective Arm:</u> **Patients with ESCC** • Patients ≥18 years old clinically fit for an endoscopy procedure • New diagnosis of ESCC during qualification for endoscopic or other oncological treatment [surgery ± radiotherapy (Rth) / chemotherapy(Chth)] • Patients with ESCC currently undergoing oncological treatment (Rth/Chth) • Written informed consent to provide tissue samples for the study **Patients at high-risk for ESCC** • Patients ≥18 years old clinically fit for an endoscopy procedure • Previous definitive treatment for HNC (cancer of the oral cavity, hypopharyngeal cancer, or • laryngeal carcinoma) at least 12 months after completion of the treatment (Rth / Chth) OR • Prior definitive endoscopic treatment for early ESCC (at least 6 months since completion of the • treatment) OR • Active smokers with weekly intake of alcohol of at least 26 units • Consent to provide tissue samples for the study
Study Start	01 Jan 2021
Study Duration	3 years

Abbreviations

BLI; blue laser imaging, **EMR**; endoscopic mucosal resection, **ESCC**; esophageal squamous cell carcinoma, **ESD**; endoscopic submucosal dissection, **FDA**; Food and Drug Administration, **HG-IEN**; high-grade intraepithelial neoplasia, **HPRA**; Healthcare Products Regulatory Agency, **LG-IEN**; low-grade intraepithelial neoplasia, **NBI**; narrow-band imaging, **TCGA**; The Cancer Genome Atlas, **VAS**; Visual Analog Scale

1. Research Project Objectives

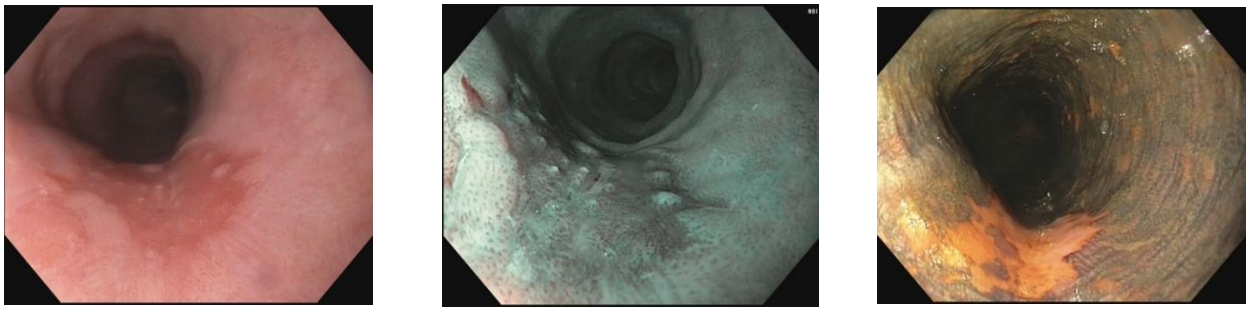
The objective of the study is to identify molecular abnormalities at each developmental stage of the esophageal squamous cell carcinoma (low-grade and high-grade intra-epithelial neoplasia and early esophageal squamous cell cancer) in order to develop a clinical assay for early detection and surveillance of the disease using the minimally-invasive Cytosponge™ cell collection device.

2. Significance of the project

Esophageal squamous cell carcinoma (ESCC) is the most common type of esophageal cancer worldwide, accounting for nearly 90% of the 456,000 incident cases of esophageal cancer each year[1]. Overall, it is the seventh most common malignancy and the sixth most common cause of cancer-related mortality[2] with high incidence rate in Eastern to Central Asia, and eastern and southern Africa[1]. In Poland, over 1,000 men and almost 300 women are diagnosed with esophageal cancer annually, and ESCC accounts for the majority of cases[3]. This cancer is more common in men (~70%) and the main risk factors include cigarette smoking, alcohol consumption, poor oral hygiene, ingestion of caustic agents, and nutritional deficiencies[4]. Additionally, an increased risk of ESCC following curative treatment of head and neck cancer (HNC) has been well-documented in the literature, with a lifetime incidence ranging between 3.8% and 14.9% in the prospective observational studies[5].

The carcinogenesis of ESCC is sequential and preceded by several precancerous stages, including low-grade intraepithelial neoplasia (LG-IEN), and subsequently, high-grade intraepithelial neoplasia (HG-IEN). Both premalignant stages are detectable in endoscopy. However, due to highly inconspicuous presentation, accurate recognition of these subtle lesions can be challenging. Improved detection is possible with advanced imaging techniques, such as narrow-band imaging (NBI), magnifying endoscopy, or Lugol's iodine dye-enhanced endoscopy (**Figure 1**).

Figure 1. Endoscopic images of high-grade intraepithelial neoplasia in different imaging modalities (source: National Institute of Oncology, Warsaw, Poland)



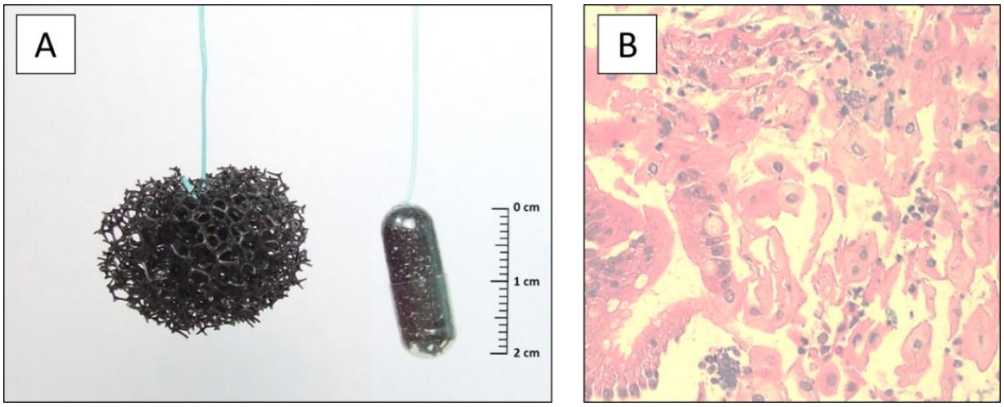
White light endoscopy (standard)	Narrow-band imaging (NBI)	Imaging with 2% Lugol's iodine solution

Although the prognosis of ESCC is extremely poor with 5-year survival below 20%, it dramatically improves if the disease is detected at an early stage[6]. Consequently, mass screening in high-incidence regions is being widely debated[7]. However, population-wide screening presents a large challenge in terms of cost-effectiveness and man-power, as currently, a potential screening regime for ESCC would rely on endoscopic examination with biopsies. Furthermore, since around 80% of all ESCCs occur in

economically less-developed regions, newer, cheaper, and less-invasive diagnostic tools are being highly warranted.

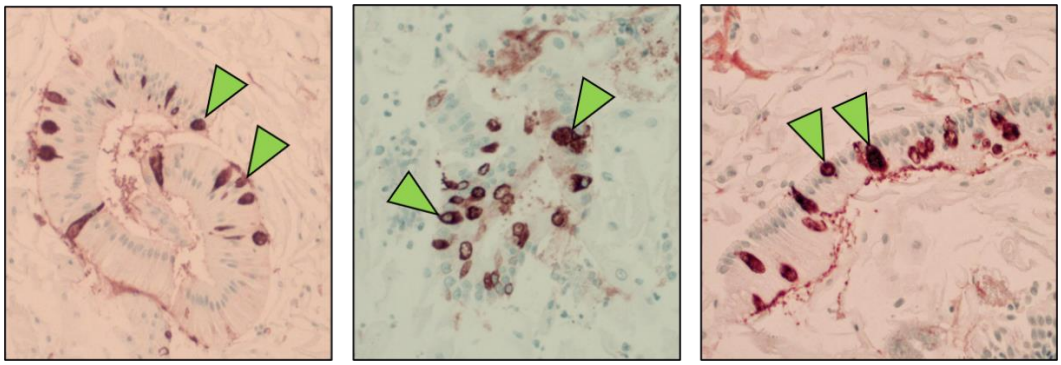
The Cytosponge™ is a novel, non-endoscopic cell collection device which consists of a capsule, containing a 30-mm polyurethane sponge, attached to a string (**Figure 2A**). The patient swallows the capsule which then dissolves in the stomach releasing the sponge. When withdrawn, the device collects esophageal cells for analysis. The procedure requires minimal training and can be safely administered by a nurse in a primary care setting[8,9]. The cells retrieved from the sponge are spun down and embedded to produce a pseudo-biopsy suitable for routine laboratory analysis (**Figure 2B**). So far, the diagnostic efficacy of the Cytosponge™ has been evaluated primarily in patients with gastroesophageal reflux disease (GERD) in the context of early diagnosis of Barrett's esophagus – a precancerous condition for esophageal adenocarcinoma.

Figure 2. Cytosponge™ cell collection device in the capsule and expanded (A), and sample cell material collected with Cytosponge™ (B); source: Cambridge University



In the previous two large, prospective studies (BEST1 and BEST2), Cytosponge™ coupled with a trefoil factor-3 biomarker (**Figure 3**)[10] showed a high accuracy in detecting Barrett's metaplasia [sensitivity of 73.3% and 79.5%, and specificity of 93.8% and 92.4%, respectively[11,12]]. The Cytosponge™ has an excellent safety profile (<1:2,000 adverse incidence rate), and patients' tolerability (average visual analog scale (VAS) of 6.0, IQR 5.0-8.0)[8], and has been approved by both the Food and Drug Administration (FDA) and the Healthcare Products Regulatory Agency (HPRA). Furthermore, a pilot study confirmed the feasibility of Cytosponge™ for diagnosis of HG-IEN[13]. However, the diagnosis of HG-IEN was based only on the identification of atypical cell types which was a limiting factor. Analogous to Barrett's esophagus, molecular alterations characterizing ESCC could be used to develop clinically relevant diagnostic biomarkers to better diagnose squamous intraepithelial neoplasia (IEN).

Figure 3. Examples of trefoil-factor 3 (TFF3) staining (20x magnification) in Cytosponge™ samples with Barrett's esophagus tissue showing columnar lined epithelium with highlighted goblet cells (arrowheads); Source: Cambridge University



Taken together, the present study aims to expand the current knowledge on molecular abnormalities related to each developmental stage of the esophageal squamous cell carcinoma to develop novel, robust and affordable clinical diagnostic assays, which will then be coupled with the minimally-invasive Cytosponge™ cell collection device. Progress made in this field could improve early detection of esophageal cancer and extend the long-term survival of patients worldwide.

3. Work plan

Study Overview

This is a two-stage multi-center cohort study with a 1. retrospective arm, and a 2. prospective arm. The study involves two host institutions from Poland and one collaborating unit from the United Kingdom with specific responsibilities regarding the study:

Host Units	Role:
Centre of Postgraduate Medical Education, Warsaw, Poland (**CMKP**)	• Supervision and coordination of the study • Legal and ethical aspects • Communication between Units and transportation of samples • Prospective arm: sequencing library preparation using method selected in the retrospective arm (Postdoctoral researcher) • Joint computational analysis with CAM (Postdoctoral researcher)
National Institute of Oncology, Warsaw, Poland **(NIO)**	Clinical aspects, including: • Recruitment of patients • Diagnostic procedures (blood withdrawal, Cytosponge™, endoscopy) • Tissue and blood sample storage Laboratory analyses: • Histopathological assessment of specimens
International collaboration:	**Role:**

MRC Cancer Unit, University of Cambridge Cambridge, United Kingdom (**CAM**)	• Training of administration and processing of Cytosponge samples
	• Expert, independent review and scoring of Cytosponge histopathology and IHC biomarkers
	Laboratory preparation and analyses, including:
	• Next Generation Sequencing based technologies (retrospective arm):
	  ○ Shallow Whole genome sequencing
	  ○ mutREAD
	• Computational analysis including clinical stratification (joint with CMKP)

Primary aims:

Retrospective arm

- Identification of molecular abnormalities at each developmental stage of the esophageal squamous cell carcinoma (LG-IEN, HG-IEN and ESCC) to establish new molecular biomarkers with potential for early detection and surveillance using the minimally invasive Cytosponge™ cell collection device.

Prospective arm

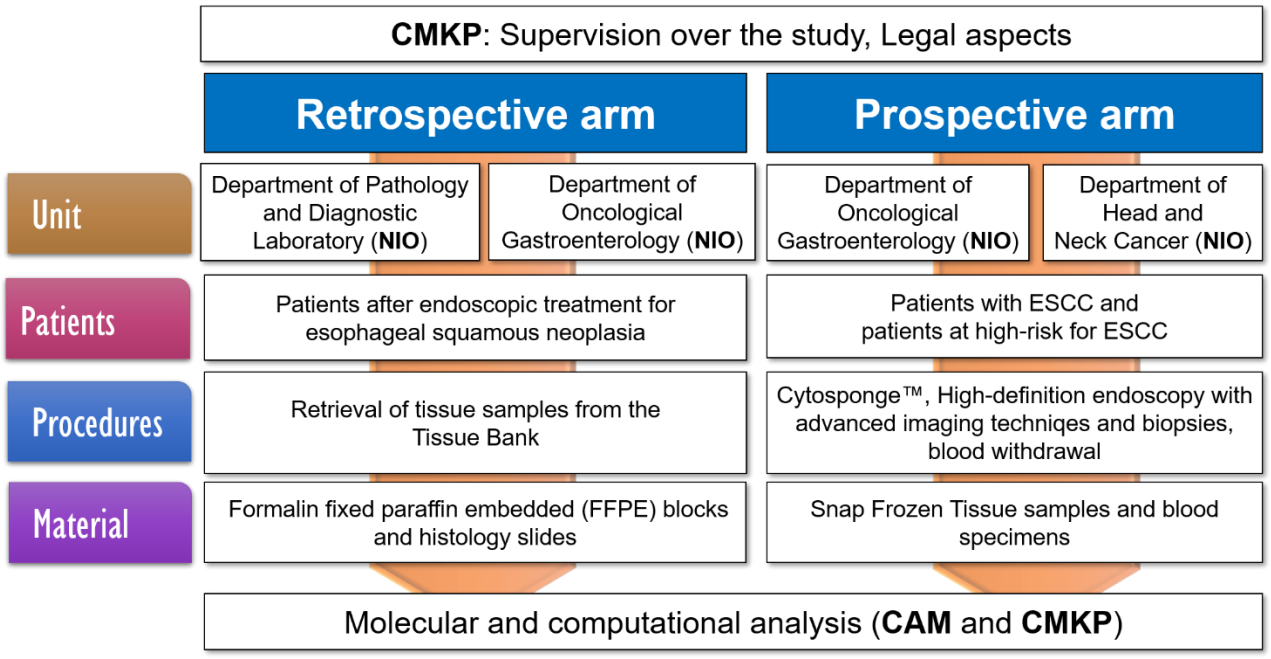
- Evaluation of the diagnostic assay established during the retrospective arm (sensitivity and specificity) in patients with ESCC, or at high risk for ESCC, using prospectively collected tissue samples from the Cytosponge™ and upper endoscopy with biopsies.

Secondary aims:

- Assessment of safety and acceptability of the Cytosponge™ device in patients with ESCC and at high-risk for ESCC through active monitoring of adverse events and the procedure acceptability questionnaire with a Visual Analogue Scale (VAS).
- Knowledge and expertise transfer between CAM and CMKP / NIO. CMKP aims to recruit postdoctoral researcher who will undertake training at CAM during retrospective arm and will independently perform sample analysis for prospective arm of the study.
- Establishment of a biobank of frozen and FFPE samples from ESCC and control patients for future studies.

The study will follow procedures described in **Figure 4**.

Figure 4. Study overview.



Ethics and trial registration

The study has received a positive opinion of the Bioethics Committee at the CMKP (opinion number 109/PB/2019). A Collaboration Agreement between all three institutions (CMKP, NIO, CAM) regulating the Legal aspects and transportation of tissue material has been obtained (17 March 2020). The study has been registered at the Clinicaltrials.gov (NCT04192695).

4. Research methodology

Detailed study characteristics

Retrospective arm

Procedures:

- Retrieval of tissue material from the Tissue Bank and transportation to CAM
- Assessment of histological features of individual samples
- Isolation of nucleic acids from the samples
- Genetic analysis of samples using shallow Whole Genome Sequencing (sWGS) and mutREAD (both detailed in section: I*dentification of Biomarkers*)
- Computational analysis of the samples

Patients:

- Adult patients who previously underwent endoscopic treatment, including both endoscopic mucosal resection (EMR) and submucosal dissection (ESD) for early esophageal squamous neoplasia (LG-IEN, HG-IEN, and early ESCC)

Inclusion Criteria:

- Tissue specimens from patients ≥18 years old after endoscopic treatment (EMR/ESD)
- Written informed consent given for samples to be used in research
- Diagnosis: LG-IEN, HG-IEN, ESCC pT1a/pT1b in TNM classification[14]
- Lesion size: ≥1cm
- Single tissue fragment (*en-bloc*)

Sample size:

- Due to the exploratory nature of the study a formal sample size calculation was not possible. After reviewing our clinical and histological databases at NIO, we estimate to analyze approximately 80-100 tissue samples collected from 20-30 individual patients.

Identification of Biomarkers

Despite well-established demographic and clinical risk factors for ESCC, molecular abnormalities associated with the development of this malignancy remain largely unknown. Over the past decade, the focus of molecular research on this disease has been primarily confined to ESCC, with scarce data on the pre-cancerous stages (LG-IEN and HG-IEN). Currently, IEN is diagnosed by histological assessment of esophageal samples. Multiple studies attempted to identify molecular (either RNA or

protein) biomarkers of IEN [15–17]. However, due to sample heterogeneity, the specificity and sensitivity of these biomarkers were poor.

On the genetic level, the Cancer Genome Atlas (TCGA) program has confirmed the histological classification of esophageal carcinoma into its squamous cell and adenocarcinoma subtypes[18]. *TP53* gene mutations occur in the majority of patients with ESCC (approx. 85-90%)[19,20]. Mutant variants have also been found in *NFE2L2*, *MLL2*, *ZNF750*, *NOTCH1*, and *TGFBR2* genes. However, the feasibility of these mutations as potential biomarkers of progression was questioned by observation that *TP53* and *NOTCH1* mutations are widespread in the normal aging epithelium [21]. Additionally, ESCC is affected by large scale DNA copy number aberrations (CNA) that seem to be early events of its development and are present in the majority of IEN cases[4]. Finally, exome-wide studies identified widespread contribution of APOBEC-related mutational processes to the genesis of ESCC (COSMIC Signatures 2 and 13, C>T and C>G transitions at TpCpW trinucleotide sites)[19].

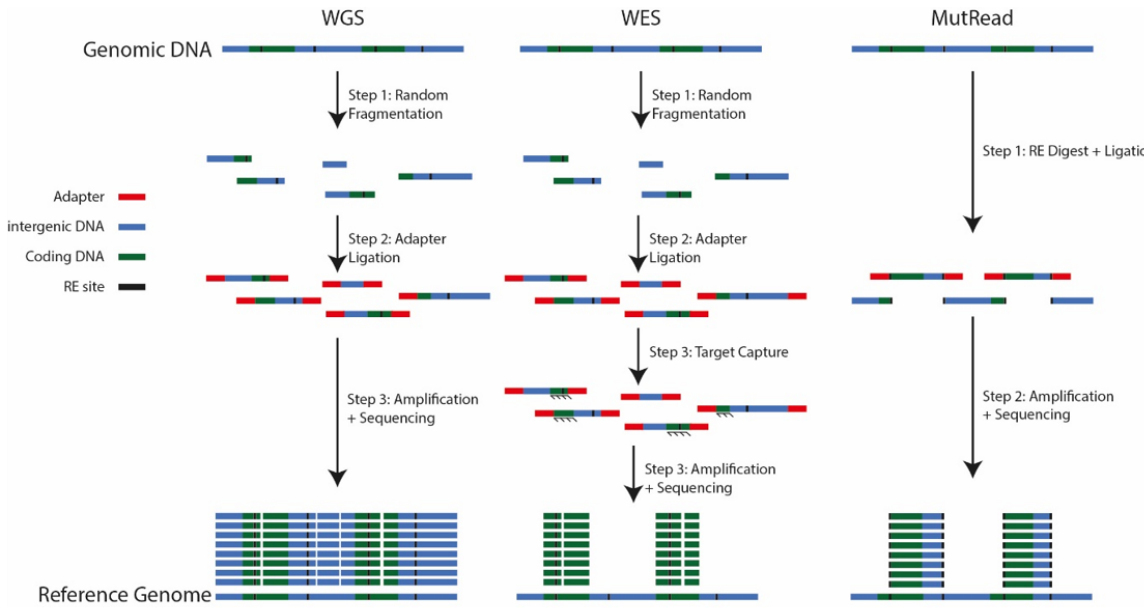
The retrospective arm of the study aims to identify molecular biomarkers of IEN. Past studies have shown difficulties in identifying protein and RNA-based biomarkers for this condition. In the current study, we will take advantage of recently developed technologies at CAM – shallow whole genome sequencing and Mutational Signature Detection by Restriction Enzyme-Associated DNA Sequencing (mutREAD, Pernier et al., **Figure 5**) to test the potential of mutational signatures and the CNA as potential biomarkers of IEN. Both methods are cost-effective alternatives to Whole Genome Sequencing (at least 10-fold decrease in cost) and can be coupled with plethora of samples including degraded formalin-fixed paraffin-embedded (FFPE) specimens. Furthermore, an RNA sequencing (RNA-seq) analysis of samples matched to samples used for CNA and mutational signatures analysis will be performed at CAM with a post-doc from CMKP. In order to maximize the likelihood of success and take full advantage of the available cohort, previous successful CAM methodologies for identification of esophageal biomarkers will be performed[10,22].

The retrospective arm of the project will comprise of ~30 patients diagnosed with LG-IEN, HG-IEN or early ESCC. It is common, that in addition to the highest grade of the condition, specimen will contain regions with lower grades of the disease (e.g. HG-IEN would be adjacent to LG-IEN) and normal tissue margins. We will take an advantage of this diversity to maximize the number of individual samples studied per grade. The samples will be process as follows:

1. Preparation of histopathological slides (4 μm for H&E and 2 x 20 μm for DNA/RNA extraction)
2. Identification of regions of interest (ROI) by a trained pathologist
3. Extraction of DNA / RNA from microscopically micro-dissected pathological tissue (LG-IEN, HG-IEN, ESCC) and regions with normal mucosa (up to 5 regions per specimen for a total of up to 150 samples)
    a. If available, reference DNA will be isolated from the blood samples, otherwise, normal stromal/muscle tissue surrounding pathological samples will be used as reference
4. After quality control, sequencing libraries will be prepared following manufacturer's protocols (sWGS, RNA-seq) or custom protocols (mutREAD)
5. After quality control of sequencing libraries, they will be pooled and sequenced using Nova-Seq sequencers (Illumina Inc.) at CAM with post-doc from CMKP
6. RNA sequencing will be performed using standard commercially available methods and sequenced at CAM

7. In addition to nucleic acid analysis of the specimen, immunohistochemical (IHC) analysis of FFPE slides will be performed using previously identified biomarkers of IEN (e.g. p53)
8. Bioinformatics analysis will be performed jointly at CMKP and CAM and will be used to identify the best combination of features (CNA regions, mutational signatures, gene expression patterns, p53 staining pattern) that distinguish IEN / ESCC samples from normal squamous tissue:
    a. The CAM team has recently shown that sWGS provides data sufficient for identification of copy number changes[23] and further employed machine learning approaches (elastic-net regression) to identify prognostic biomarkers for Barrett's esophagus progression [24]. We aim to employ similar approaches to identify CNA features within sWGS data obtained from this project
    b. Moreover, the CAM team have recently shown that mutREAD provides an accurate estimate of mutational signatures exposure in individual cancer samples of varying quality (Pernier, Nat. Comms. Accepted, patent pending). Our preliminary analysis suggests that mutREAD further allows for identification of CNA within the same samples (**Figure 5**). We will use this approach to identify CNA and mutational signatures within studied samples and extend it to machine learning approaches for further identification of individual features linked with the diagnosis of IEN and ESCC.
    c. RNA-seq analysis will be performed using standard pipelines (DESEQ2, EdgeR) and differentially expressed genes will be identified. Similarly to previous strategies [10,22], the list of putative biomarkers will be narrowed down using publicly available dataset (e.g. proteinatlas.org, gepia2[25]) and the expression of the individual genes will be validated either on the protein level (immunohistochemistry) or RNA level (RT-qPCR).
9. The best (high sensitivity and specificity) biomarker will be selected and validate in the prospective arm of the study.

Figure 5. Comparison of mutREAD with whole genome sequencing (WGS) and hole exome sequening (WES).



	WGS	WES	mutREAD
Genomic regions covered	Unbiased representation of genome	Only Coding	Unbiased representation of genome
Part of genome sequenced (%)	~90%	~2%	0.1 to 10%
Cost in comparison to WGS	100%	20-40%	~10%

In order to fulfil the secondary aim of the study, i.e. knowledge transfer between CAM and CMKP / NIO, the CAM team will perform molecular analysis of the samples together with a Post-doc employed by CMKP. The post-doc will undertake sequencing library synthesis, IHC analysis of the samples and computational analysis of the samples. This knowledge will be used in the prospective arm of the study during sample analysis at CMKP / NIO.

<u>Prospective arm</u>

Procedures:

- Behavioral questionnaire
- Blood withdrawal
- Cytosponge™ procedure
- High-Definition upper endoscopy with advanced imaging modalities (narrow-band imaging, Lugol's staining, magnification) and biopsies

Inclusion Criteria:

Patients with ESCC

- Patients ≥18 years old clinically fit for an endoscopy procedure

- New diagnosis of ESCC during qualification for endoscopic or other oncological treatment [surgery ± radiotherapy (Rth) / chemotherapy(Chth)]
- Patients with ESCC currently undergoing oncological treatment (Rth/Chth)
- Written informed consent to provide tissue samples for the study

Patients at high-risk for ESCC

- Patients ≥18 years old clinically fit for an endoscopy procedure
- Previous definitive treatment for HNC (cancer of the oral cavity, hypopharyngeal cancer, or laryngeal carcinoma) at least 12 months after completion of the treatment (Rth / Chth)

  OR

- Prior definitive endoscopic treatment for early ESCC (at least 6 months since completion of the treatment)

  OR

- Active smokers with weekly intake of alcohol of at least 26 units
- Consent to provide tissue samples for the study

Exclusion Criteria (for both groups of patients):

- Significant swallowing difficulties: dysphagia grade ≥3 (able to swallow only liquid foods)
- Esophageal varices or stricture of the oesophagus requiring dilatation
- Active anticoagulation therapy (warfarin, acenocoumarol, clopidogrel) for high-risk conditions, which preclude possibility to stop the medication without bridging with heparin
- Myocardial infarction or any significant cardiac event within less than six months prior to recruitment into the study
- Patients in long-term care or institutional care (physical, psycho-social disorders, intellectual disability)

Sample size:

- Due to the exploratory nature of the study a formal sample size calculation was not possible. We estimate to recruit approximately 100 patients (~20 patients with ESCC undergoing treatment and ~80 patients at high risk for ESCC), however, this number may be changed (most likely increased) after obtaining results from the initial, retrospective phase of the study.

Study procedures

- Behavior Questionnaire

Subjects will be asked to complete a behavior questionnaire that includes:

- Demographics
- Socio-economic information and eating habits
- Risk factors (cigarette smoking, alcohol consumption)
- Clinical information (history of cancer, prior medications)

The study team will also have access to patient's medical history in order to collect any missing clinical data. All information will be collected anonymously and in a manner that ensures the safety of subjects' identifiable data. Only the members of the research group will have access to information collected using the behavior questionnaire.

- Blood samples

A blood sample will be taken as a reference for DNA analyses on the tissue. A total of approximately 10 mls of blood will be taken into EDTA tubes. The tube will be frozen and stored at -80°C until sent for further analysis.

- Cytosponge™

The Cytosponge™ procedure will be done prior the scheduled endoscopic procedure (on the same day). Patients must be fasting (no food intake for 6 hours and no liquids intake for at least 2 hours). The procedure will be performed by an investigator or qualified nurse following relevant training. The procedure involves the patient swallowing a Cytosponge™ capsule and most of the string with a small amount of water (approx. 50 mL). After approx. 5 minutes, which is the time required for the cellophane coating of the capsule to dissolve and release the polyurethane sponge in the stomach, the sponge is withdrawn through the patient's mouth. The sponge along with the cells on its surface is detached from the string and secured in a sealed container with liquid preservative.

- Endoscopy with biopsies

Patients will undergo a clinically indicated, standard-of-care endoscopy using high-definition endoscopes available at the NIO (Olympus GIF-H180 / Olympus GIF-H190 / Olympus Evis X1). A separate written informed consent will be obtained from each individual prior to the procedure. Patients will be asked to give an additional consent for the collection of research biopsies for the purposes of this study. The endoscopy will involve either topical anesthesia of the pharynx (lignocaine spray), intravenous conscious sedation (intravenous Midazolam +/- Fentanyl) or general anesthesia with anesthesiologist's support, depending on individual preference. As per standard of care, endoscopy will evaluate the entire length of the esophagus using white light imaging, as well as advanced imaging techniques, such as NBI. At the discretion of the operator, additional imaging with 2.0% Lugol's iodine solution will be performed. Biopsy specimens will be taken from suspicious areas:

- Two biopsy specimens from suspicious areas <20mm in diameter: one for routine pathology evaluation and one for research purposes
- Three or four biopsy specimens from suspicious areas ≥20mm diameter: 1-2 for routine pathology evaluation and 2 for research purposes (**Figure 6**)
- Additional biopsy specimens will be collected from normal-appearing mucosal sites (as normal control tissue) and immediately frozen in peroxynitrite (-80°C) and secured for later shipment.

In patients with suspicious lesions, a further clinically indicated oncological treatment will be arranged (EMR / ESD / surgery). The resected specimens will constitute a final diagnosis for the patient.

In the absence of any suspicious lesions, patients will be assigned to the 'healthy tissue' control group and only two biopsy specimens will be collected from normal-appearing mucosa in the middle part of the esophagus for research purposes.

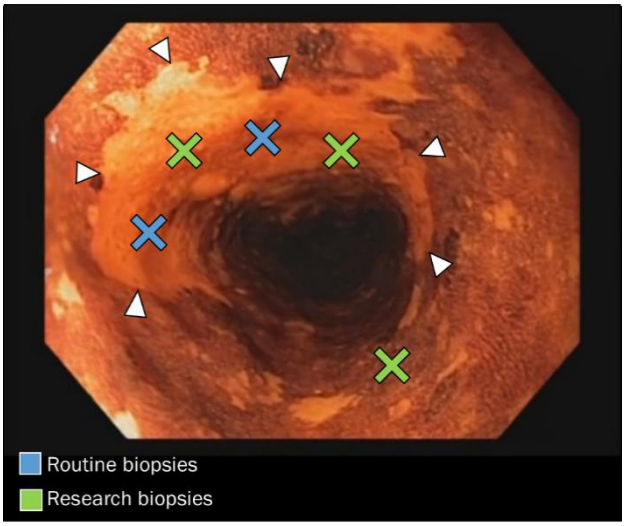


Figure 6. A schematic representation of both clinical (routine) and research biopsies taken from a Lugol-voiding suspicious lesion of >20mm in diameter (indicated by white arrowheads)

Prospective validation of the diagnostic assay

This arm of the project aims to validate biomarkers established in the retrospective arm of the project using an independent cohort of patients. The project will further extend to include samples collected using Cytosponge™ device. Since Cytosponge™ sample the entire surface of esophagus, we anticipate that, albeit at low levels, the evidence of IEN will be present within these samples. The prospective arm will be split into:

Validation of biomarkers in the individual biopsies (frozen and FFPE) collected during routine endoscopies:

1. The availability of frozen samples with increase the likelihood of successful RNA-seq analysis
   a. Additionally, for patients diagnosed with IEN / ESCC the endoscopic treatment samples (EMR / ESD) will be analyzed in line with protocol described in the retrospective arm
   b. Frozen samples will serve as a biobank of specimen for potential future studies (e.g. methylation analysis of DNA)
2. Validation of biomarkers with specimen collected using Cytosponge™ device.

Prospective arm end-points

Each approach (1/2) will be coupled with normal histopathological assessment of biopsies and Cytosponge™ samples. Research team will be blinded to the outcome of histopathological assessment until the conclusion of genetic analysis. A Cytosponge™ diagnosis will be made basing on the cytopathological assessment of the collected cells (cellular atypia etc.) and molecular profile from the biomarker assay. A diagnosis of neoplastic vs normal tissue will be made and juxtaposed with the final endoscopic diagnosis (gold standard). Sensitivity and specificity of the Cytosponge™ coupled with molecular biomarkers for the diagnosis of esophageal neoplasia will be generated.

Prospective arm methodology

For approach (1), we will use the same, selected, methodology described in the retrospective arm of the projects. Briefly, nucleic acids will be extracted from tissue scrolls immediately adjacent to the sections used for histopathological analysis. As the research is blinded to histopathological assessment, with the exception of treatment specimen, microdissections will not be performed. Subsequently, the

samples will be sequenced and genetic-based diagnosis for individual biopsies will be made based on the presence of previously selected biomarker. After genetic analysis, the results of the analysis will be compared to the histological assessment of patients' biopsies (gold standard).

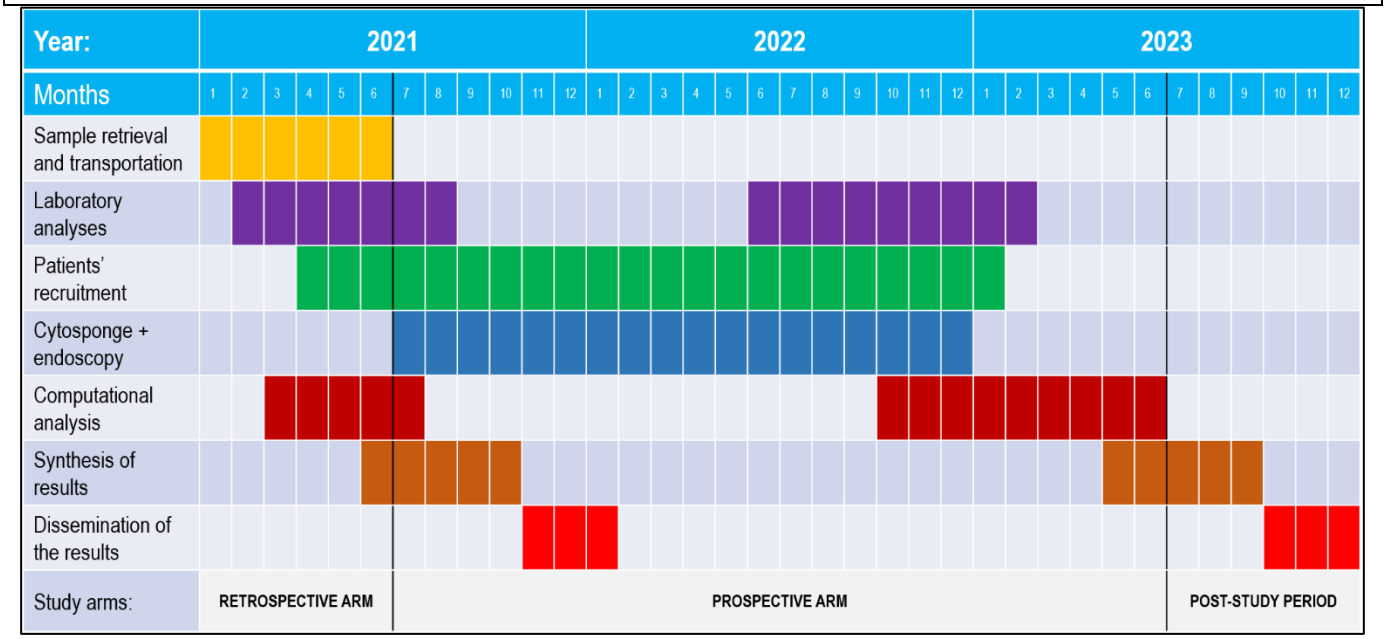
Approach (2) will follow an analogues procedure using Cytosponge™ specimen as input samples. Due to larger normal tissues contamination (as the Cytosponge™ travels through esophagus, it collects normal tissue in additional to diseased tissue), the sequencing methods will be adjusted to account for sample lower purity. After computation analysis, the biomarker-based diagnosis of Cytosponge™ specimen will be made. This diagnosis will be compared to the histological assessment of patients' biopsies (gold standard) and the sensitivity and specificity of Cytosponge™-based diagnosis will calculated. Furthermore, since Cytosponge™ samples the entire surface of esophagus, the results obtained from it will be compared with molecular analysis of individual targeted biopsies. This will allow us to establish the source of discrepancies between histopathological diagnosis and Cytosponge™-based diagnosis.

Specimen analysis will be performed jointly by CMKP / NIO and CAM teams. The postdoctoral researcher trained at the CAM site during the retrospective arm of the study will undertake, except for sequencing itself, all experimental steps (e.g. biomarker extraction, sequencing library preparation) at CMKP / NIO. Finally, CMKP / NIO and CAM teams will jointly perform the analysis of samples.

5. Study timeline and risks

The study is expected to last 36 months with specific task plan as presented in **Figure 7**.

Figure 7. Gannt graph presenting the timeline and tasks regarding the study



Potential risks

The study is susceptible for potential risk on each of its stage. The potential risks and adequate responses to those events are presented below:

Risk	Reaction
Failure to identify an accurate molecular biomarker during the retrospective arm	In these circumstances, we will limit our diagnostic assay to histopathological assessment of the Cytosponge™ samples in conjunction with IHC markers, such as the *P53*. Nevertheless; even if the molecular assay won't show a diagnostic potential, the study still will contribute a substantial knowledge on the molecular characterization of each developmental stage of the ESCC, which currently remains limited.
Cytosponge™ not safe / feasible in the studied cohort.	The clinical team of the study group will ensure safety of Cytosponge™ administration and will actively monitor adverse events and patients' tolerability. Cytosponge™ is regarded as an extremely safe procedure even in patients with an increased risk of foreign body impaction (this device was tested in the subset of patients with eosinophilic esophagitis known to have swallowing difficulties). However; in the unlikely event of poor patient's tolerability, we will limit the evaluation of the biomarker assay only to endoscopic biopsy samples and post endoscopic treatment tissue specimens.
Recruitment issues	NIO is a high-volume institution and a referral center for esophageal cancer in Poland, therefore, we don't expect issues with recruitment. However; if this becomes a problem, we will contact other units from the National Cancer Institute network for help with patient's recruitment.

6. Conclusions

ESCC is a severe disease with poor survival affecting over 1,300 people in Poland every year. Early detection of ESCC and its precursor conditions leads to a significant improvement in patients' survival. Furthermore, early detection of this disease coupled with modern endoscopic treatments, significantly improves patients' quality of life by sparing them from life-changing surgeries. Our proposal relies on the expertise of CMKP, NIO, and CAM teams and aims to identify robust, cost effective method for early detection and surveillance of ESCC. Firstly, in the retrospective arm of the study we will use our existing knowledge in establishment of molecular biomarkers to identify robust methods for measurement of molecular biomarkers in ESCC and its precursors. Secondly, using an independent prospective cohort we will validate our biomarkers and extend their measurements to Cytosponge ™ specimen which will be compared with the current gold standard, which is endoscopy with biopsies. To conclude, our proposal will bring a multi-disciplinary team with international collaboration and will combine proven concept of esophageal sampling by Cytosponge™ with novel, robust molecular biomarkers that can be effectively, and more importantly affordably, measured in the clinical setting, ultimately leading to improved patients outcomes.

References:

1. Abnet, C. C., Arnold, M. & Wei, W. Q. Epidemiology of Esophageal Squamous Cell Carcinoma. *Gastroenterology* **154**, 360–373 (2018).
2. Arnold, M. *et al.* Global Burden of 5 Major Types Of Gastrointestinal Cancer. *Gastroenterology* (2020) doi:10.1053/j.gastro.2020.02.068.
3. Wong, M. C. S. S. *et al.* Global Incidence and mortality of oesophageal cancer and their correlation with socioeconomic indicators temporal patterns and trends in 41 countries. *Sci. Rep.* **8**, 4522 (2018).
4. Rustgi, A. K. & El-Serag, H. B. Esophageal Carcinoma. *N. Engl. J. Med.* **371**, 2499–2509 (2014).
5. Chaber-Ciopinska, A., Kiprian, D., Kawecki, A. & Kaminski, M. F. Surveillance of patients at high-risk of squamous cell esophageal cancer. *Best Pract. Res. Clin. Gastroenterol.* **30**, 893–900 (2016).
6. Wang, G. Q. *et al.* Long-term results of operation for 420 patients with early squamous cell esophageal carcinoma discovered by screening. *Ann. Thorac. Surg.* **77**, 1740–1744 (2004).
7. Codipilly, D. C. *et al.* Screening for esophageal squamous cell carcinoma: recent advances. *Gastrointest. Endosc.* **88**, 413–426 (2018).
8. Januszewicz, W. *et al.* Safety and Acceptability of Esophageal Cytosponge Cell Collection Device in a Pooled Analysis of Data From Individual Patients. *Clin. Gastroenterol. Hepatol.* **17**, 647-656.e1 (2019).
9. Offman, J. *et al.* Barrett's oESophagus trial 3 (BEST3): Study protocol for a randomised controlled trial comparing the Cytosponge-TFF3 test with usual care to facilitate the diagnosis of oesophageal pre-cancer in primary care patients with chronic acid reflux. *BMC Cancer* **18**, (2018).
10. Lao-Sirieix, P. *et al.* Non-endoscopic screening biomarkers for Barrett's oesophagus: From microarray analysis to the clinic. *Gut* **58**, 1451–1459 (2009).
11. Kadri P, S. R. *et al.* Acceptability and accuracy of a non-endoscopic screening test for Barrett's oesophagus in primary care: Cohort study. *BMJ* **341**, 595 (2010).
12. Ross-Innes, C. S. *et al.* Evaluation of a Minimally Invasive Cell Sampling Device Coupled with Assessment of Trefoil Factor 3 Expression for Diagnosing1. Ross-Innes CS, Debiram-Beecham I, O'Donovan M, et al. Evaluation of a Minimally Invasive Cell Sampling Device Coupled with Asse. *PLoS Med.* **12**, 1–19 (2015).
13. Roshandel, G. *et al.* Pilot study of cytological testing for oesophageal squamous cell dysplasia in a high-risk area in Northern Iran. *Br. J. Cancer* **111**, 2235–2241 (2014).
14. Brierley, J. D., Gospodarowicz, M. K. & Wittekind, C. TNM classification of malignant tumours - 8th edition. *Union Int. Cancer Control* 241 (2017) doi:10.1002/ejoc.201200111.
15. Kashyap, M. K. *et al.* Genomewide mRNA profiling of esophageal squamous cell carcinoma for identification of cancer biomarkers. *Cancer Biol. Ther.* **8**, 34–46 (2009).
16. Xue, L. Y. *et al.* Tissue microarray analysis reveals a tight correlation between protein expression pattern and progression of esophageal squamous cell carcinoma. *BMC Cancer* **6**, 296 (2006).

17. Taylor, P. R., Abnet, C. C. & Dawsey, S. M. Squamous dysplasia-The precursor lesion for esophageal squamous cell carcinoma. *Cancer Epidemiol. Biomarkers Prev.* **22**, 540–552 (2013).

18. Kim, J. J. J. J. *et al.* Integrated genomic characterization of oesophageal carcinoma. *Nature* **541**, 169–174 (2017).

19. Chang, J. *et al.* Genomic analysis of oesophageal squamous-cell carcinoma identifies alcohol drinking-related mutation signature and genomic alterations. *Nat. Commun.* **8**, (2017).

20. Zhang, L. *et al.* Genomic analyses reveal mutational signatures and frequently altered genes in esophageal squamous cell carcinoma. *Am. J. Hum. Genet.* **96**, 597–611 (2015).

21. Martincorena, I. *et al.* Somatic mutant clones colonize the human esophagus with age. *Science (80-.).* **362**, 911–917 (2018).

22. Couch, G. *et al.* The discovery and validation of biomarkers for the diagnosis of esophageal squamous dysplasia and squamous cell carcinoma. *Cancer Prev. Res.* **9**, 558–566 (2016).

23. Macintyre, G. *et al.* Copy number signatures and mutational processes in ovarian carcinoma. *Nat. Genet.* **50**, 1262–1270 (2018).

24. Killcoyne, S. *et al.* Genomic copy number predicts oesophageal cancer years before transformation. *bioRxiv* 2020.02.27.967612 (2020) doi:10.1101/2020.02.27.967612.

25. Tang, Z., Kang, B., Li, C., Chen, T. & Zhang, Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* **47**, W556–W560 (2019).