ODELIA

Open Consortium for Decentralized Medical Artificial Intelligence

HORIZON-HLTH-2021-CARE-05-02

Deliverable D2.7

MIDTERM RECRUITMENT REPORT

Lead beneficiary	CAM
Author(s)	Dr Lorena Escudero Sanchez, Dr Nicholas Payne
Dissemination level	PU
Type	R
Delivery date	22/12/2023

ODELIA is funded by the European Union's Horizon Europe Framework under Grant Agreement 101057091

Funded by the European Union

ODELIA

TABLE OF CONTENTS

Executive summary 3

Introduction 3

Current Statistics 3

Issues faced 4

Preparation for Public Challenge 4

Conclusion 4

DISCLAIMER

EXECUTIVE SUMMARY

This report updates on overall recruitment achieved by sites by the end of month 12 of ODELIA (31 December 2023) and any issues faced – a target of 50% of the expected cases (i.e. 25,000 cases) were due to be available at this time point. Each centre completed a comprehensive spreadsheet to gather this information, detailing cases available per site and providing details of cancer vs non-cancer case numbers, population type, and annotations available if applicable. A total of 37,647 cases were found to be available, exceeding the expectation the collaboration sought to report at this point of the grant.

INTRODUCTION

This report aims to collect statistics on the current status of data availability per site, and to identify any potential issues early on. To that effect, at the beginning of December 2023, we asked each centre to fill a comprehensive shared spreadsheet to help us understand the different populations of patients available per site, how many cases are to be expected in total, how many were already available for Swarm Learning, and their expectations of how many could be made publicly available in the following year for a public challenge (D2.3). We also collected information on how many of the available and expected scans are cancer cases, how many of those cases are annotated, and the level of annotation.

The expectation is to have achieved 50% recruitment of study population at the time of this report. This means, from a total of 50,000 exams that will be included in ODELIA, approximately 25,000 cases to be available (not shared) at clinical sites by 31 December 2023.

In addition to the current statistics, we also asked each centre to provide details of any delays or problems that might have happened, and we summarized the findings in this report.

CURRENT STATISTICS

Recruitment of patients' data has started and all centres bar one have images available. These are the available scans per site at the time of writing this report:

- CAM: **149 (of total of 153, all cancers) and 177 (of total of 2,500 with some tens of cancers)** from two populations: High-risk screening (cohort of high-risk individuals) and BRAID trial (a dense breast trial with abbreviated MRI). **Currently minimum level of annotation is available for these scans.**
- UKA: **10,000 (of total of 10,000 with 500 cancers)**. Pre-operative planning / Problem solving / Follow-up after Breast Cancer / other. **Annotations are available for 427 cases.**
- UMCU: **11,521 (of total of 11,521 with 178 cancers)**. DENSE – a randomized-controlled population screening trial, incorporating prevalent and incidence rounds. **Annotations are available, spreadsheet to be filled.**
- USZ: **313 (of a total of 3,000 with >80 cancers)**. Pre-operative planning / Problem solving / Follow-up after Breast Cancer / other. **Annotations are available, spreadsheet to be filled.**
- MHA: **1,600 (of total of 1,600 with 380 cancers).** Screening / Preoperative Local Staging / Problem Solving. **Annotations are available, spreadsheet to be filled.**
- RUMC: **13,772 (of total of 13,772 with 252 cancers)**. High-risk screening database including prevalent and incident rounds

- VHIO: **0 (of total 200 with 75 cancers)**. High risk screening. Notice that this dataset had to change with respect to the one expected at the time of the proposal due to sharing agreements, explaining the delay.
- RSH: **25 (of total 2,640 with >1200 cancers).** High-risk screening, follow-up of B3 lesions and problem solving and surveillance for breast cancer. **Currently minimum level of annotation is available for these scans.**

ISSUES FACED

Four centres have 100% of the expected cases available: UKA, UMCU, MHA, RUMC

Other centres have ongoing data acquisition. Delays are variously due to:

- Delays in staff recruitment for the project during 2023 (e.g. UCAM staff started in November 2023, USZ in September 2023).
- Discussions on data format across the consortium have only happened in the last quarter of 2023.
- Getting the appropriate hardware (procurement) available on site.
- Sharing agreements with sites. For example, some datasets have images coming from different sites and it has been difficult/slow to get access to them. One centre had to change the dataset from the one originally intended (VHIO).

The centres are already working on alleviating the sources of delays and it is expected that they will have the rest of cases available during 2024.

PREPARATION FOR PUBLIC CHALLENGE

Discussions have already started in preparation for the public challenge deliverable of the grant. It has been decided to align the public challenge with a major conference, and MICCAI 2025 was chosen. This sets the timescale for the data to be public and the challenge to take place in Spring-Summer 2025. These are the current expectations regarding the dataset to be made public:

- CAM expected to contribute with cases from two different datasets (50+ cancers, 200+ non-cancer cases)
- UKA expected to contribute with 200 cases (100 cancer, 100 non-cancer)
- RSH has already shared 70 cases (60 cancers, 10 non-cancer) - they need to be checked
- MHA expected to contribute with 100 cases (40 cancer, 60 non-cancer)
- RUMC potentially, not sure, would need to get approval

CONCLUSION

A total of 37,647 scans are currently available in the different sites across all centres. This is over the expected 25,000 at the time of this report. Reasons for some data not being available at centres have been mentioned in this report, and the centres affected are working on mitigating issues and acquiring the remaining cases during 2024.