EPAD
European Prevention of
Alzheimer's Dementia Consortium

EPAD data and sample access quarterly bulletin

This bulletin covers applications made between May 2019 and January 2021.

Authors: Samuel O Danso (UEDIN), Cindy Birck (Alzheimer Europe), Alicia Gibson (Aridhia), Jean Manson (UEDIN) and Craig W Ritchie (UEDIN)

For any further details, please contact Samuel Danso (samuel.danso@ed.ac.uk)

TABLE OF CONTENTS

This bulletin presents an analysis of applications made to access EPAD data and samples. Data access application process is described here. All applications are logged with date and timestamps. This maiden bulletin covers applications made between May 2019 and January 2021.

1. Application outcome

A total of 125 applications was processed between the reporting period. Figure 1 shows the breakdown of decisions made on applications.

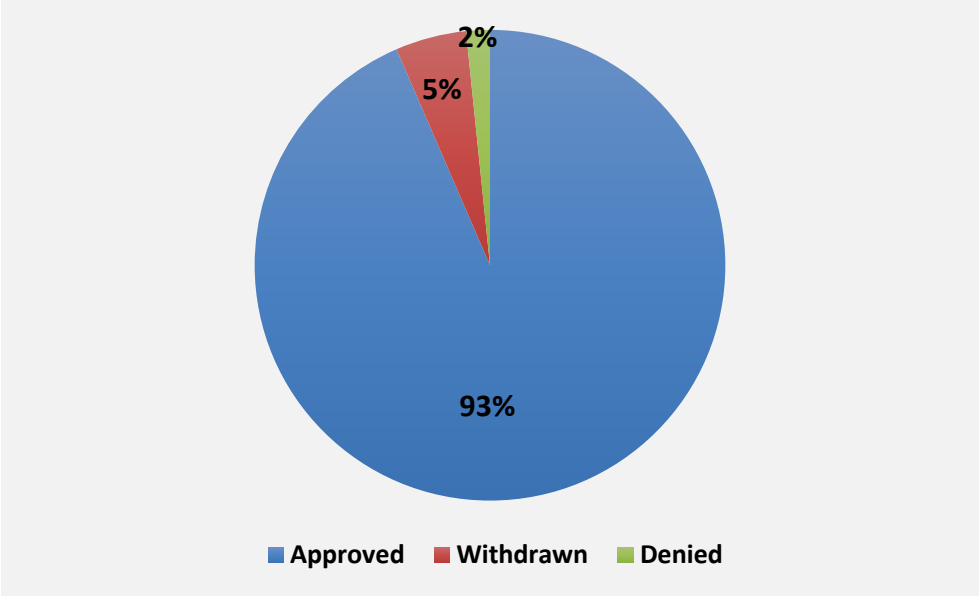

Figure 1: Application outcomes

As the figure 1 shows, 93% of applications was approved, while 5% was withdrawn and 2% was denied. Reasons for an application being denied include the identity of an applicant cannot be established or the applicant is named as a co-applicant on an already approved application.

2. Type of data requested

There are three types of requests:

- Data Only - With "Data Only type of application", only the quantitative data is made available to the researchers.
- Data + Images - "Data + Images application type" means the quantitative data as well as MRI imaging scans are made available to the researchers.
- Data + Samples - Similarly, the "Data + Samples application" would involve making the quantitative as well as the biological samples available to the researchers once application is approved. Figure 2 shows the distribution of type of applications received.

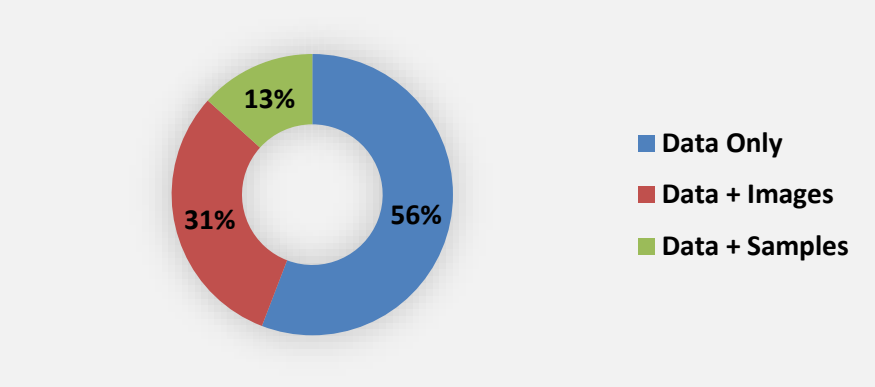

Figure 2: Type of data requested

As Figure 2 shows, Data Only accounts for about 56% of the total applications received. About 31% of the researchers requested for the MRI scans whereas 13% of the researchers were interested in biological samples.

3. Researcher affiliation

A look into the affiliation of researchers accessing the EPAD data suggests that researchers affiliated to the EPAD project form the majority of applicants. Figure 3 shows the distribution of applications by researcher affiliation.

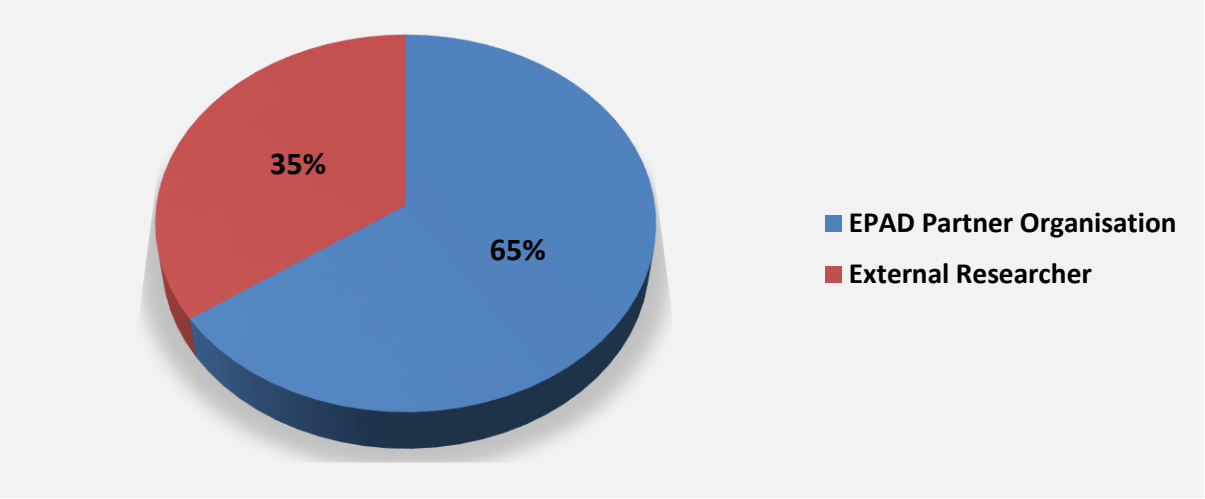

Figure 3: Distribution of researchers

As Figure 3 shows, about 65% of applications received came from researchers within EPAD partner organisations while 35% of the applications were from researchers outside of the EPAD partner organisations. The EPAD partner organisations include academic and industrial institutions across Europe and the United Kingdom.

4. Distribution of requests by period and affiliation

A detailed look into the trend of applications made during the period under review and grouped by periods highlights a positive trend and also correlates with the operational policy that guided the release of the datasets. Figure 4 shows the periodic analysis of applications and the affiliation of researchers.

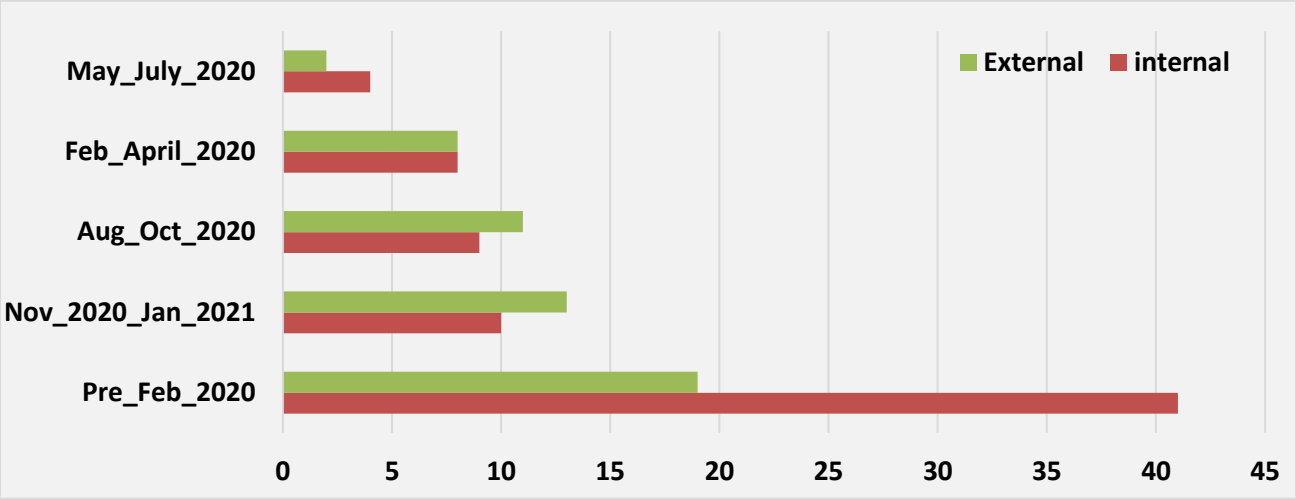

Figure 4: Distribution of applications by period and researcher affiliation

As Figure 4 shows, the period prior to February 2020 accounts for the most of applications with applicants from EPAD partner organisations being the majority. However, this trend began to change after July 2020. The interest for the EPAD data has steadily increased with applications from external institutions outside of the EPAD partner organisations, showing a gradual increase over the internal applications across the periods. The figure also shows the months between May and July 2020 as the period with the least number of the applications. It is unclear at the moment whether this is due to the COVID-19 pandemic or summer holidays. This will become clearer in subsequent analysis.

5. Internal researcher affiliation

Exploring how EPAD partner organisations engaged with the EPAD datasets as released by the EPAD consortium, Figure 5 provides a breakdown of the number of applications received from researchers affiliated to EPAD partner organisations. The figure 5 shows that researchers affiliated to the University of Edinburgh engaged most with the dataset for the period under review. This is followed by researchers at the VU University Medical Centre Amsterdam engaged with the data.

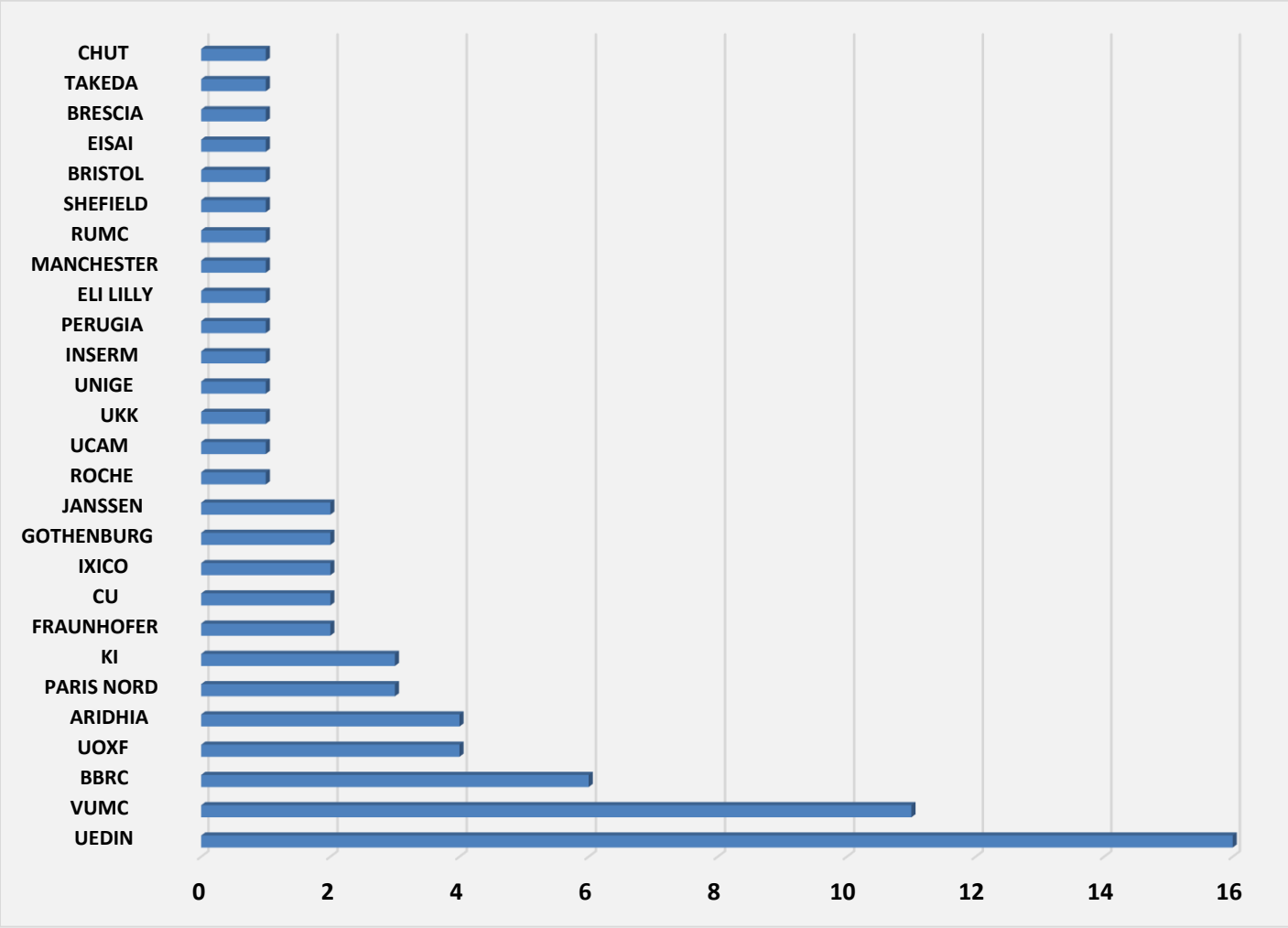

Figure 5: Distribution of applications made by EPAD partner organisations

6. Country of external applications

Applications were received from external researchers across the globe. Figure 6 shows the distribution of external applications by country.

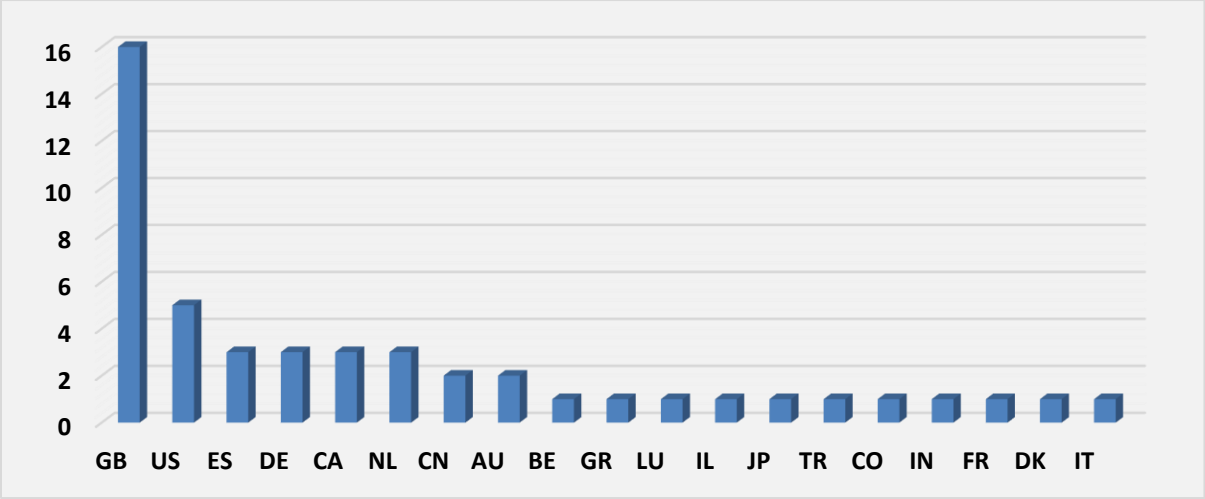

Figure 6: Country of external applicant

As figure 6 shows, researchers from the United Kingdom engaged with the EPAD dataset most, which is then followed by researchers from the United States. There is a marginal increase in the number of researchers from Australia compared to those in Japan, France and Denmark.

7. City of external researchers

Exploring which city the external researchers are based, Figure 7 shows the distribution of external researcher applications by city.

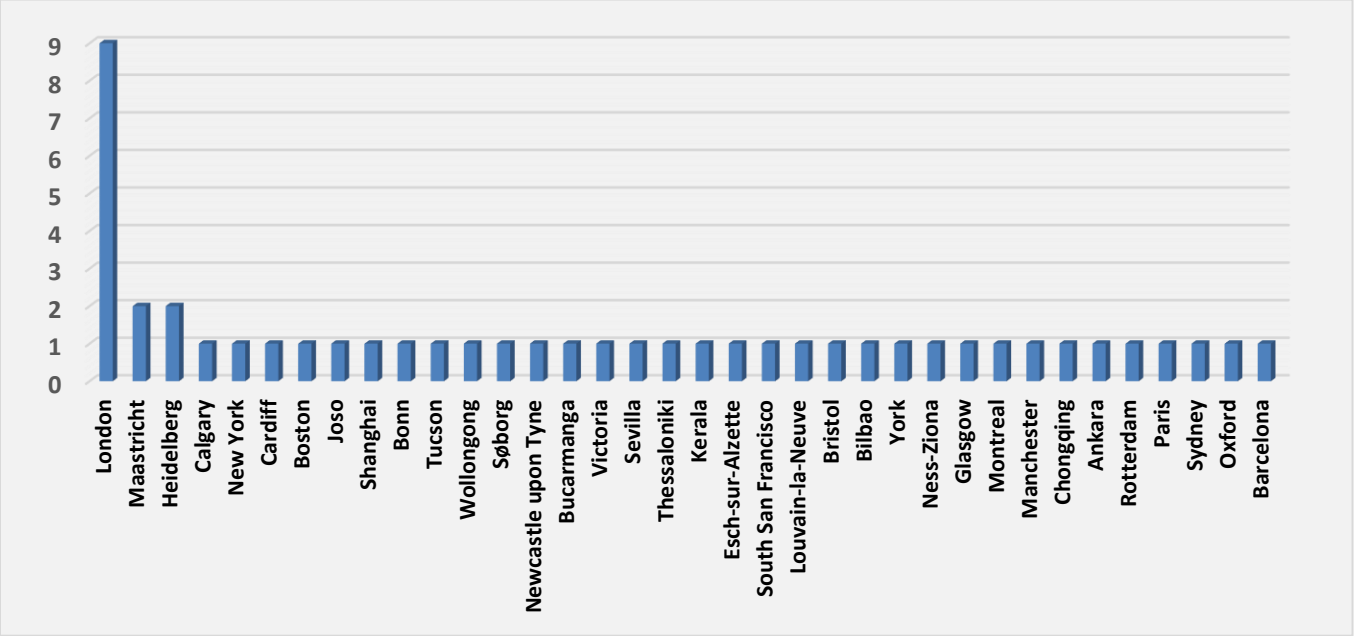

Figure 7: Distribution of external researchers by city

As the figure shows, the researchers are evenly spread across most cities internationally. London however appear to be an outlier as it stands out from the rest of the cities, being the city with most external researchers accessing the EPAD data.

8. Keywords on application in word cloud

Keywords analysis provides a 'window' to the type of analysis to be carried by researchers. Using wordcloud provides an abstract view and a snapshot of the type of analysis as outlined by researchers. Figure 8 and 9 show wordclouds generated from the keywords entered by researchers as part of the application.

The figure 8 shows the wordcloud for all applications while figure 10 shows wordcloud generated based on the applications for biological samples only. As expected keywords such as "Alzheimer's", "amyloid", "biomarkers", "risk", "CSF", "tau", "cognitive", "prediction", "data" and "learning" are some of the frequent keywords entered as part of the application. This suggests that most of the applications are looking some risk prediction models for Alzheimer's disease.

Similarly, figure 9 shows "biomarkers", "antibody", "beta-amyloid", "phago", "receptor", "repertoire", "amyloid", "Apoe", "elisa", "b-cell" and "blood" as some of the most frequent keywords, suggesting the sample access are heavily focused around biological analysis of Alzheimer's disease.

Figure 8: Wordcloud of keywords entered on all application

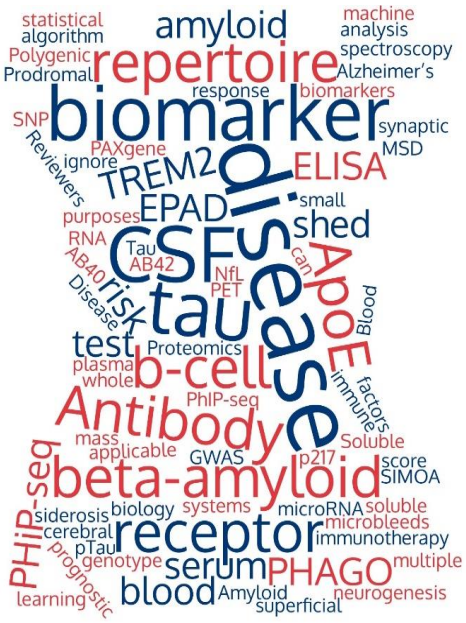

Figure 9: Wordcloud of keywords entered on sample access applications