CHARIOT PRO Sub-Study (CPSS) Data Use Agreement (DUA)

I request access to the CPSS data for the purpose of scientific investigation, or the planning of clinical research studies.

I agree to the following terms:

1. I will receive access to pseudonymised data and I will not attempt to re-identify participants or request additional information about specific individuals.
2. I will not attempt to make direct contact with the CPSS study team to enquire about specific results of individual participants.
3. I will not further share this data beyond the uses outlined in this agreement and my data use application. I understand that redistribution of individual participant-level data in any manner is prohibited. However, I am allowed to produce reports and publish summaries of not participant-level data.
4. I will not utilise any AI tools that are not specifically and clearly defined in my Data Use application, nor any AI tools with public-facing interfaces that do not offer guarantees regarding the containment of data inputs, for CPSS data. I recognise that the use of such AI tools, including generative and analytical models, poses a potential risk of inadvertent data sharing due to the nature of how these tools process information and may lead to data being sent, saved, viewed, or used in unforeseen ways by parties not covered by the DUA, which is a direct violation of this agreement. See Appendix A for additional AI tool concerns, restrictions, and guidance.
5. I will not disclose any participant-level raw or derived datasets beyond the uses outlined in this agreement. If I create any derived datasets containing participant-level data that I anticipate will benefit the scientific community, I will submit these to the CPSS study Chief Investigator (CI), Prof. Lefkos Middleton, using the format described in Derived Data Submission Form (available via the CPSS entry on the AD Discovery Portal). The study CI will determine whether my derived data may be distributed via the AD Discovery Portal. I understand derived datasets should only be shared through the relevant catalogue entry via the AD Discovery Portal.
6. I will ensure anyone on my team utilising these data to comply with this Data Use Agreement. I understand that I am not allowed to distribute CPSS data outside of my team, even if the recipient is within my same institution/company. Access to the CPSS data may only occur via requests through the AD Discovery Portal.

7. I will accurately provide the requested information for persons who will use these data and the analyses that are planned using these data.
8. I will respond promptly and accurately to annual requests for updates on my use of the CPSS data.
9. I will comply with any rules and regulations imposed by my institution and its institutional review board/ethics committee in requesting these data.

When publishing abstracts using CPSS data:

1. I agree to cite "CHARIOT PRO Sub-Study (CPSS)" as the source of data and the CPSS funders.
2. I understand that for abstracts, it is not required to cite CPSS in the authorship line.

When publishing manuscripts using CPSS data, I agree to the following:

1. On the by-line of the manuscript, after the named authors, I will include CPSS as an author by using the phrase: *"for the CHARIOT PRO Sub-Study**" with the asterisk referring to the following statement:

    **The data used in this manuscript were obtained from the CHARIOT PRO Sub-Study (CPSS), made available via the Alzheimer's Disease Data Initiative's AD Workbench research data platform. CPSS was initiated in 2015 as an industry sponsored (Johnson & Johnson) prospective longitudinal study led by Lefkos Middleton MD, of Imperial College London. The study's primary goal is to evaluate factors of risk and/or protection for cognitive decline and dementia due to Alzheimer's disease (AD) and related neurodegenerative diseases. Towards this goal, the study employed magnetic resonance imaging (MRI), positron emission tomography (PET) and fluid-based biomarkers (cerebrospinal fluid and blood-based) as well as repetitive clinical and multiple neuropsychological assessments to track cognitive trajectories longitudinally. Additional funding from Merck, Takeda and Gates Ventures allowed, in 2022, a two- year extension of CPSS.*

    The authors would like to express their gratitude to the study participants, the CPSS investigators and study funders for their invaluable contribution to the development and processing of the extensive clinical, imaging and other biomarker data and making them available via the AD Workbench to the international scientific community to accelerate discovery in the field.

2. I will include language similar to the following in the Methods section of my manuscripts in order to accurately acknowledge data gathering by the CPSS team. Depending upon the length and focus of the article, it may be appropriate to include more or less than the example below. However, inclusion of some variation of the language shown below is mandatory.

   The Cognitive Health in Ageing Register: Investigational, Observational and Trial Studies in Dementia Research (CHARIOT): Prospective Readiness cOhort (PRO) SubStudy (denoted hereafter as CPSS) is a prospective, longitudinal, biomarker-enriched, observational study. CPSS opened in 2015 as an industry sponsored, biomarker enriched, prospective, longitudinal study led by Professor Lefkos T. Middleton, MD as Chief Investigator (CI). CPSS included two study sites, Imperial College London (ICL) and University of Edinburgh (EDI). The main objective was to evaluate whether brain imaging, in particular Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET), as well as six monthly clinical and neuropsychological assessments, could highlight any risk and/or resilience factors towards developing cognitive decline in healthy older adults. The study also aimed to track the participants' patterns of cognitive trajectories and, in cases where this occurred, the emergence of the earliest stage of Mild Cognitive Impairment (MCI) and subsequent dementia due to Alzheimer's Disease and related neurodegenerative diseases. To this end, an equal number of participants with above- and below-threshold amyloid brain load, determined through brain imaging, were followed up longitudinally over at least four years. A total of 519 participants consented to undergo screening, 410 at ICL and 109 at EDI. The Edinburgh site was closed in early 2020, with CPSS continuing as a single-site study at ICL. A two-year follow-up study was conducted in a selected 100 participants (the majority of whom had shown to have high amyloid burden) who underwent an MRI and up to three Tau-PET scans. CPSS was completed in July 2024. I will acknowledge funding for CPSS in the support acknowledgement section of the manuscript using language similar to the following:

   Data collection and sharing for the CPSS was made possible through the financial support of Johnson & Johnson Pharmaceutical Company and, more recently, through additional financial support from Gates Ventures, Merck and Takeda Pharmaceuticals.

3. I will submit all manuscripts to the CHARIOT PRO team for review prior to submitting them to a journal. This will not be a scientific review but rather aimed at ensuring that the terms above are correctly implemented. CPSS will maintain confidentiality of the

manuscript and will complete its review within 3 weeks. CPSS can be reached by email to dementia.prevention@imperial.ac.uk.

Note: If a journal requires a Conflict of Interest declaration because CPSS is listed as an author of the manuscript, the CPSS team will provide one upon request.

4. I will ensure that all Investigators who utilise CPSS data use appropriate administrative, physical, and technical measures to prevent the use or disclosure of the data other than as outlined in this Agreement.
5. I will report to the CPSS team any use or disclosure of the data breaching the terms of this Agreement within 15 days of becoming aware of such use or disclosure.

CPSS maintains the right to modify terms of this Agreement and may do so by posting notice of such modifications on this page: https://www.imperial.ac.uk/school-public-health/ageing/ Any modification made is effective immediately upon posting the modification (unless otherwise stated). You should visit this page periodically to review the current Data Use Agreement terms.

APPENDIX A

CPSS DUA: Use and Requirements of AI tools

Prohibited use of AI Tools: CPSS participant-level data may only be processed and analysed using computational systems and AI tools that remain under the direct administrative, contractual, and technical control of the AD Workbench and that do not permit secondary use, model training, or access to CPSS data by any third party not bound by this Data Use Agreement. For a variety of reasons, including the important need to protect the privacy of research participants, all data access requests must originate using the CPSS entry in the AD Discovery Portal and the data must be directly released to permissioned individuals via the AD Workbench.

The use of AI tools, including certain types of generative and analytical models, pose a genuine risk of inadvertent data sharing due to the nature of how AI tools process, store, and regenerate information. AI tools, including but not limited to those with public-facing interfaces, such as OpenAI's free-version of ChatGPT, usually do not offer guarantees regarding the containment of data inputs. Any data input into a prompt or output generated from a prompt could result with proprietary data infringed, sent, saved, viewed, misinterpreted, or used in unforeseen public ways. As a result, the use of such AI tools which may release CPSS participant level data to any third party not bound by this Data Use Agreement, is in direct violation of CPSS Data Use Agreement policies. However, an important exception to the above stated prohibition would be AI tools, including generative and analytical models, which explicitly prevent sharing data with others. Such AI tools may be developed and internally trained by individual research groups or academic institutions and strictly used for the individual research group or academic institution provided that "sharing" data beyond the tool in the AD Workbench is strictly prohibited. Therefore, CPSS data may be analysed using AI tools which provide guarantees that all data is safeguarded, contained, and will not be released to others.

Note that this Appendix does not impose any new restrictions on the use of CPSS data but is intended for clarification on term #3 in the DUA (above): "I will not further share these data beyond the uses outlined in this agreement and my data use application and understand that redistribution of data in any manner is prohibited."

As of the writing of this Appendix the term "Artificial Intelligence" or "AI" means a broad range of technologies and tools capable of generating, transforming, predicting, or analysing content based on input data, including generative AI, machine learning systems, statistical software, and cloud-based analytics platforms. The recent commercialization of these AI tools represents a significant challenge in maintaining the privacy of CPSS participants, and

of participants in human subjects' research in general. Uploading CPSS data to a third-party platform that is not explicitly in compliance with the prohibition and restrictions on redistribution of CPSS data set forth herein is a clear violation of this Agreement.

CPSS requires investigators to adhere to the following terms when deciding whether a project utilizing these tools is in compliance with the terms of this Agreement:

- Computational resources
    - The application of any statistical or other analytic methods to CPSS data - whether branded as 'AI' and "AI Tool" or otherwise - should not be performed using computational or AI resources owned by third parties who engage in the long-term retention of user content that is either shared with other parties or used in the training of public-facing models, such activity could include the contents of CPSS data. This includes the use of third-party platforms for inference (e.g. the inclusion of CPSS data as part of an input or 'prompt' for a generative language model), training (including both the training of foundation models and the fine-tuning of pre-trained models via transfer learning or some other process), prediction, or any other task that could result in CPSS data being retained, shared, regenerated or otherwise by entities that are not in compliance with the DUA.
    - Some commercial entities may offer an 'opt-out' clause in the terms of use for their models, allowing users to decline the right of the entity to retain the data for training and other purposes. This does not constitute sufficient protection, and under no circumstances should this be viewed as an acceptable safeguard for sharing CPSS data with these third parties.
- Public release of models trained using CPSS data
    - The use of CPSS data in training models is both permitted and encouraged, provided that the process is carried out in compliance with the DUA as outlined above. However, investigators who intend on making the weights of a trained model publicly available should consider whether the model could be used to easily reconstruct parts of the CPSS data set. CPSS prohibits release of participant level data, even if it is sufficiently processed or altered and lack CPSS subject codes. See Term (3) above "I understand that redistribution of individual participant-level data in any manner is prohibited".
    - This precaution is particularly important in the case of large language models and other highly overparameterized generative models. Numerous studies have demonstrated that it is possible to extract training examples from these large and complex models, and investigators should keep this fact in mind.

I understand that failure to abide by these guidelines will result in the termination of my privileges to access CPSS data. Furthermore, CPSS reserves the right to pursue damages for actions which violate CPSS guidelines.