

Neuroimaging Feature Terminology: A Controlled Terminology for the Annotation of Brain Imaging Features

Anandhi Iyappan^{a,b}, Erfan Younesi^a, Alberto Redolfi^c, Henri Vrooman^d, Shashank Khanna^{a,b}, Giovanni B. Frisoni^{c,e} and Martin Hofmann-Apitius^{a,b,*}, for the Alzheimer's Disease Neuroimaging Initiative¹

^a*Department of Bioinformatics, Fraunhofer Institute for Algorithms and Scientific Computing (SCAI), Schloss Birlinghoven, Sankt Augustin, Germany*

^b*Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn-Aachen International Center for Information Technology, Bonn, Germany*

^c*Laboratory of Epidemiology and Neuroimaging, IRCCS San Giovanni di Dio Fatebenefratelli, Brescia, Italy*

^d*Departments of Radiology and Medical Informatics, Biomedical Imaging Group Rotterdam, Erasmus MC University Medical Center, The Netherlands*

^e*Memory Clinic and Laboratoire de Neuroimagerie du Vieillissement (LANVIE), University Hospitals and University of Geneva, Geneva, Switzerland*

Accepted 9 June 2017

Abstract. Ontologies and terminologies are used for interoperability of knowledge and data in a standard manner among interdisciplinary research groups. Existing imaging ontologies capture general aspects of the imaging domain as a whole such as methodological concepts or calibrations of imaging instruments. However, none of the existing ontologies covers the diagnostic features measured by imaging technologies in the context of neurodegenerative diseases. Therefore, the Neuro-Imaging Feature Terminology (NIFT) was developed to organize the knowledge domain of measured brain features in association with neurodegenerative diseases by imaging technologies. The purpose is to identify quantitative imaging biomarkers that can be extracted from multi-modal brain imaging data. This terminology attempts to cover measured features and parameters in brain scans relevant to disease progression. In this paper, we demonstrate the systematic retrieval of measured indices from literature and how the extracted knowledge can be further used for disease modeling that integrates neuroimaging features with molecular processes.

Keywords: Alzheimer's disease, annotation, brain, neuroimaging, terminology

¹Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.usc.edu>). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

*Correspondence to: Dr. Martin Hofmann-Apitius, Department of Bioinformatics, Fraunhofer Institute of Algorithms and Scientific Computing (SCAI), Schloss Birlinghoven, Sankt Augustin, 53754, Germany. E-mail: martin.hofmann-apitius@scai.fraunhofer.de

INTRODUCTION

Brain imaging technologies have revolutionized the way that neurodegenerative diseases, such as Alzheimer's disease (AD), are diagnosed and tracked. Since the human brain is largely inaccessible for direct sampling, neuroimaging provides an alternative for measuring *in vivo* structural and functional features that can be used as biomarkers of disease onset and progression. The quantitative nature of imaging biomarkers, their potential to assess disease-modifying effects, and the ability to monitor the

safety of candidate drugs all make neuroimaging readouts an extensively used, measurable endpoint in clinical trials for neurodegenerative diseases [1]. Parameters and features that are clinically measured using neuroimaging biomarkers reflect biological or pathological changes underlying disease in the brain of patients; for example, positron emission tomography (PET) imaging measures the load of amyloid- β ($A\beta$) neuritic plaques through the uptake and binding of a particular radioligand in the living brain [2]. But such measurements at the clinical level are often disconnected from their underlying mechanistic causes, on one hand, and from their corresponding patient clinical tests, on the other hand.

The importance of neuroimaging in the new era of systems neurology is highlighted by its pivotal role in linking clinical readouts to underlying mechanistic changes [3]. Thus, there is a need for integrative approaches that enable multiscale modeling of both biological and clinical data with the aim of bridging the translational gap [4]. The first step toward this goal is, however, the collection and standardization of disparate and scattered data and knowledge across many resources available for research. Among several efforts in this direction, ApiNATOMY correlates brain imaging diagnostics to affected anatomical regions of the brain through the Foundational Model of Anatomy ontology (FMA) [5]; the OntoNeuroLOG ontology covers the domain of imaging datasets and their processing methods [6]; RadLex provides a lexicon of terms relevant to diagnostic and interventional radiology [7]; and the quantitative imaging biomarker ontology (QIBO) represents a series of heterogeneous concepts across several fields including imaging physics, contrast agents, biology, and quantitation techniques [8]. Neuroimaging Data Model and Taskforce (NIDM) facilitates the exchange of large publication corpus and other relevant metadata such as provenance information related to the neuroimaging research for establishing the reproducibility of research experiments as well as overcoming the challenge of data sharing (<http://nidm.nidash.org/specs/nidm-overview.html>).

In parallel to the generation of imaging datasets, an increasing amount of imaging information is published within literature articles which often report on measured features in patients with AD [9, 10]. Such studies typically try to correlate neuroimaging readouts with defined disease stages or subtypes. As an example, Whitwell and co-workers utilized magnetic resonance imaging (MRI) in patients with atypical variants of AD and were able to categorize

these patient groups, based on measuring patterns of atrophy in medial temporal and cortical grey matter, into hippocampal sparing AD, limbic predominant AD, and typical AD subtypes [11]. This example clearly shows the importance of harvesting neuroimaging feature information from literature not only for monitoring critical imaging findings but also for stratification of patients based on their diagnostic status.

To this end, UMLS metathesaurus vocabularies were used to annotate and index radiology journal figure captions from more than 9000 articles for image information retrieval [12]. Similarly, RadLex was applied to the biomedical imaging literature and annotated more than 385,000 figures with RadLex terms [13]. However, when the National Cancer Institute Thesaurus (NCIT), Radiology Lexicon (RadLex; <http://www.radlex.org/>), Systemized Nomenclature of Medicine (SNOMED-CT), and International Classification of Diseases (ICD-9-CM) were evaluated for retrieval of radiology reports containing critical imaging findings, it was found that no single terminology is optimal for retrieving radiology reports with critical findings [14].

Biomedical terminologies and ontologies have proven their role in namespace harmonization and mediation of semantic interoperability in numerous examples [15]. One of the main application domains of shared semantics (ontologies and terminologies) lies in metadata annotation as well as data integration and knowledge retrieval [16]. The neuroimaging community has not yet come up with a consensus for commonly used and shared metadata. However, over the past decades, many initiatives have made their primary data publicly available [17]. Out of those, the Alzheimer Disease Neuroimaging Initiative (ADNI) (<http://adni.loni.usc.edu/>) and Parkinson's Progression Markers Initiative (PPMI) (<http://www.ppmi-info.org/>) have gained increasing momentum for creating an impact on data sharing across the scientific community. Despite the ongoing efforts, the significant lack of structural and semantic interoperability impedes the momentum of data sharing [18] and lack of an established framework hampers the merging of imaging data from other resources [8]. The heterogeneous data such as cortical thickness or neuropsychological assessments of individual patients that are stored in ADNI/PPMI datasets do not follow a standard nomenclature, which makes them difficult to interpret or use for validation.

Motivated by the obvious need for a terminology that enables a systematic representation and retrieval

of features derived from neuroimaging techniques, we aimed at developing a Neuro-Imaging Feature Terminology (NIFT) to capture and organize the knowledge domain of structural and functional brain features as measured and represented by neuroimaging technologies in the domain of neurodegeneration. In this study, we demonstrate the value of NIFT for the identification and extraction of neuroimaging features in both Medline abstracts and full-text publications in the context of neurodegenerative disease pathology. We also demonstrate the applicability of NIFT for the annotation of imaging readouts in MRI and CT scans. Furthermore, we go beyond retrieval and annotation of imaging concepts by providing an example of how extracted neuroimaging features can be utilized for mechanistic modeling of disease pathology.

MATERIALS AND METHODS

The NIFT terminology is built based on a hierarchical knowledge representation system by organizing higher level concepts as root nodes followed by specific sub-classes organized under them; however, it is not an ontology as it uses simple hierarchical relationships but is capable of being leveraged to an ontology in the future. However, to be leveraged to an ontology, NIFT should undergo major changes in the current hierarchical structure based on ontology formalism definitions. NIFT in its current form provides a first substrate for the scientific community to elaborate its conceptual complexity and structure. The Protégé OWL editor was used to build this hierarchical terminology (<http://protege.stanford.edu/>). This terminology was constructed using the OWL language for two reasons: firstly, the hierarchical edition and annotation of concepts in OWL language facilitates creation of such a granular terminology: and secondly, the OWL format of NIFT ensures the interoperability of the terminology file. The concepts that are included under this terminology were examined by experts from the clinical research domain.

Generation of NIFT

The NIFT terminology concepts were gathered by collecting and reading relevant publications, e-books, websites, and medical blogs related to imaging in neurodegeneration. Following the initial literature search, we also adapted some concepts from already published, highly relevant ontologies such as QIBO [19] and Radlex. Ontologies such as QIBO and

Radlex had well-structured concepts such as Imaging Techniques and Imaging Agents, which were contextually relevant for the development of NIFT. Essential entities used in the ADNI (<http://adni.loni.usc.edu/>) were also included in our terminology system.

Consequently, we enriched the NIFT with measured biomarkers obtained from the Biomedical Imaging Group Rotterdam (BIGR) pipeline, UMC Rotterdam [20] and neuGRID platform (<http://www.neugrid4you.eu>). The BIGR pipeline consists of six image processing pipelines such as FreeSurfer (<http://freesurfer.net/>), BIGR Tissue Segmentation [21], BIGR hippocampus segmentation [22], BIGR SAM-SCO [23], BIGR diffusion imaging pipeline [24], and Human Connectome Mapper [25]. The neuGRID platform consists of three image processing pipelines including: FreeSurfer, Adaboost, and SPM-grid. This platform was used to extract measured imaging indices to be added to the terminology.

For the sake of covering brain-specific anatomical structures in NIFT, we made use of the Brain Region & Cell Type Terminology (BRCT) which was initially developed to capture a wide range of key concepts representing human brain neuroanatomical structures and integrate their corresponding cell types (<http://bioportal.bioontology.org/ontologies/BRCT>). Alzheimer Disease Ontology [26] was also re-used to enrich NIFT. Pathway concepts were derived from the pathway Terminology System, that was developed with the intention to support the extraction of pathway information specific to the neurodegenerative disease domain [27].

Upon completion, this terminology system was reviewed by a clinical imaging expert team (Professor Frisoni's team at the University Hospitals Geneva) which further improved the quality and relevance of the classification.

Natural language processing (NLP)-based assessment of NIFT performance

In order to assess the relevancy of the NIFT terminology, we compared the performance of our terminology with the two already well-established imaging ontologies, QIBO and Biomedical Image Ontology (BIM) [28]. This comparison was performed at the level of terminologies found in those ontologies as they claim to capture the knowledge of image annotations and imaging biomarkers, respectively. To perform this, we used an NLP-based approach and ran the PDF tagger over the previously selected full-text publications (PMC1, PMC2,

PMC3, and PMC4) using all the three ontology/terminology systems and validated the retrieval of the maximum annotation of terms specific to neuroimaging domain. The validation of the terminology using PDF tagger was performed using the formula:

$$m_{jk} = \sum_{i=1}^{nk} a_i$$

where a_i , frequency of single term; j , document number; m_j , overall frequency in document j ; nk , number of items in dictionary k ; k , dictionary number.

This index sums up the recall of relevant terms captured using the relevant terminology over all the terms found in the document. This sum gives an overall count of different concepts and terms captured from the given document.

This analysis was done to demonstrate the usability of NIFT in extracting relevant context from publications of interest.

Correlating clinical diagnosis with imaging features for staging AD

For bridging the clinical indices with imaging readouts, we systematically harvested relevant publications using the query “((((([Neuroimaging Feature]) AND [MeSH Disease “Alzheimer Disease”]) AND [Alzheimer Ontology: “Cognitive tests”]) AND [Organism: “Homo sapiens”]) AND [BRCT”] in SCAIView.

Retrieval and mining figure captions and full-text from PubMed

Following the curation and further refinement of the terminology, NIFT was integrated into our in-house literature mining environment SCAIView [29]. SCAIView enables the users to efficiently retrieve context specific articles from the literature using standardized terminologies and ontologies. NIFT in its SCAIView integrated form can be freely accessed using this link (<http://academia.scaview.com/academia/>).

We performed an overall coverage analysis of NIFT by running it over figure captions and full-text articles using SCAIView. For this, we converted the OWL file into a dictionary (.syn) file using a java program. The resulting dictionary was incorporated in ProMiner, which is a rule-based entity recognition system [30]. The hierarchical structure of the OWL file was converted into an XML tree so that NIFT can be navigated within the SCAIView environment

and faceted search becomes feasible. The ProMiner program was subsequently run over the five figure captions which were enriched with imaging indices from PubMed articles and four full-text publications from PubMed Central (PMC), which generated an abstract with markup of the terms specific for NIFT.

We also performed an analysis of full-text publications using a special PDF tagger (http://publica.fraunhofer.de/eprints/urn_nbn_de_0011-n-936860.pdf). In order to perform this task, we chose four full-text publications, which were relevant to the neurodegenerative context as well as reported imaging findings namely PMC1 [31], PMC2 [32], PMC3 [33], and PMC4 [34]. The PDF tagger was run over these publications for validation of the coverage of NIFT and results were stored in a dedicated directory. The PDF tagger first makes use of the documents in the directory as an input to create a term list from all the PDF files. Following the complete annotation of the PDF files, an output file was automatically generated with the original PDF file containing additional annotations highlighted through markup of terms.

Annotation of image scans using NIFT

In order to annotate brain scans with NIFT, we chose three groups of patients with different diagnostic features, namely: ADNI_016_S_4952 Control (CN), ADNI_002_S_4171 Mild Cognitive Impairment (MCI), and ADNI_003_S_4136 AD from the ADNI dataset (<http://adni.loni.usc.edu/>). ADNI is a large-scale, multicenter study which has been structured to develop molecular, clinical, and biochemical biomarkers from longitudinal patient data for early detection of AD. We processed PET (F18-AV-45 and FDG [18]) and T13D MPRAGE scans using the neuGRID platform with different pipelines, such as: SPMgrid to detect hypo-metabolism as well as amyloid burden; Freesurfer to highlight cortical thickness measurements and subcortical morphological differences; and Adaboost to quantify the hippocampal differences among the three diagnostic groups. The morphological changes observed from the patient scans were further annotated manually using the NIFT terminology.

Mechanistic modeling of image-derived indices in the context of AD

Yet another important aspect of this paper is to identify the role of molecular mechanisms, which bring about clear diagnostic outcomes captured by

imaging techniques. For this purpose, we generated the following query in SCAIView: “((([Neuroimaging Feature]) AND [MeSH Disease: “Alzheimer Disease”]) AND [Organism: “Homo sapiens”])” and filtered for Human Genes/Proteins. Next, we developed a “global map” of brain-region image-derived features along with molecular readouts, such as genes linked to a neuroimaging feature. We studied mechanisms of hippocampal atrophy in detail at molecular and cellular level. This knowledge was transformed into a cause-and-effect model using Biological Expression Language (BEL) platform (<http://openbel.org/>). BEL is a platform for representing causal and correlative relationships from biological context in a computer readable form. Then, we performed a high-resolution modelling of the mechanism underlying hippocampal atrophy. The outcome of this analysis will be further discussed in the Results section.

RESULTS

Often, literature resources misclassify an imaging technique as a biomarker while many others denote the derived indices as a biomarker. Owing to this, NIFT was constructed to represent, integrate, and harmonize heterogeneous knowledge across the domain of imaging biomarkers in the context of neurodegeneration.

Structure and content

NIFT comprises of 7 major classes namely Algorithms, Brain Region, Clinical indices, Clinical trial information, Imaging technique, Measured Feature, and finally Radiopharmaceutical compound. There are in total 1,221 terms in NIFT. The root concepts of NIFT include

- (i) Algorithms which contains 4 children nodes namely: Image acquisition, MR-image analysis, PET-image analysis, and Post-processing algorithm. This concept contains all the brain imaging features that are automatically detected using various imaging pipelines such as FreeSurfer.
- (ii) The second root concept in NIFT is Clinical Indices which has two children concepts namely AD and Parkinson’s disease. This includes all the genetic, proteomic biomarkers mentioned in the literature for AD and Parkinson’s disease.
- (iii) The third root concept is Clinical trial information which contains three children concepts namely adverse effects observed in patients with neurodegeneration, neuropsychological assessments and scores such as Mini-Mental Status Examination score, Alzheimer’s Disease Assessment Scale-Cog test, and clock draw test to name a few.
- (iv) The fourth root concept of NIFT is Imaging Technique. This contains 7 children concepts, each of them represents the different imaging techniques used to study the various structural and functional dimensionality of the brain.
- (v) The fifth root concept consists of measured features. This concept covers a wide range of “observable indicators” that determine the state of the brain and disease progression observed using various imaging techniques. This concept includes structural features such as cortical thickness, cerebral atrophy and functional features such as glucose metabolism, blood oxygenation level dependent signal.
- (vi) The last root concept consists of radiopharmaceutical compounds. This concept contains all the radioactive tracers that are induced in the brain to diagnose dysfunction.

NIFT is available in OWL format and can be accessed from the following link (<https://www.scai.fraunhofer.de/en/business-research-areas/bioinformatics/downloads.html>). The hierarchical structure of NIFT is illustrated in Fig. 1.

NIFT evaluation

The content evaluation of NIFT in comparison to QIBO and BIM ontologies showed that NIFT performed comparatively better than QIBO and significantly better than BIM in capturing relevant terminology (see Fig. 2). For the first document (PMC1), we found 97 relevant terms annotated by BIM, 204 with NIFT, and 113 with QIBO. The second document (PMC2) was annotated with 308 relevant terms by BIM, 1334 terms by NIFT, and 1056 terms by QIBO. The third document (PMC3) retrieved 153 terms for BIM, 552 terms for NIFT, and 495 terms for QIBO. The fourth and final document (PMC4) retrieved 87 terms for BIM, 303 terms for NIFT, and 217 terms for QIBO. The PMC documents can be found in the Supplementary File 1.

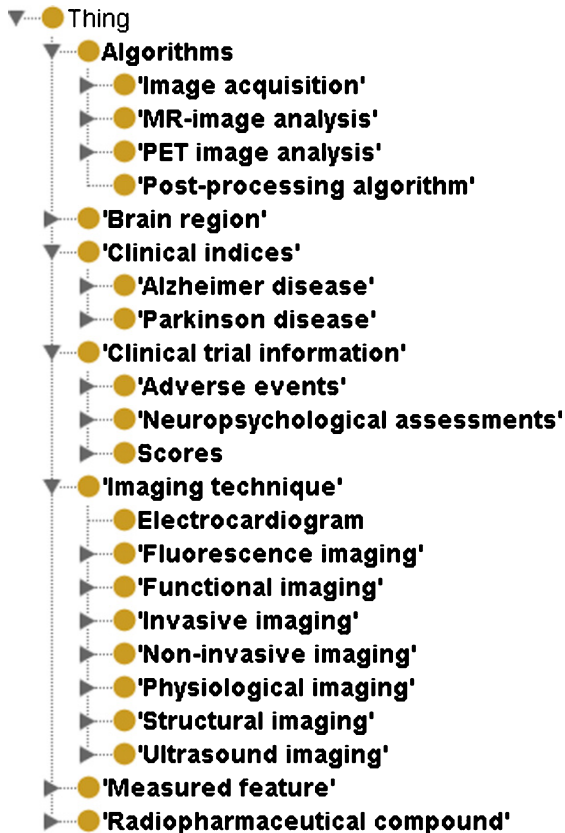


Fig. 1. Hierarchical structure of NIFT as visualized in the Protégé OWL Editor. This figure depicts the higher level concepts the terminology namely Algorithms, Brain Region, Clinical Indices, Clinical trial information, Imaging Technique, Measured Feature, and Radiopharmaceutical compound.

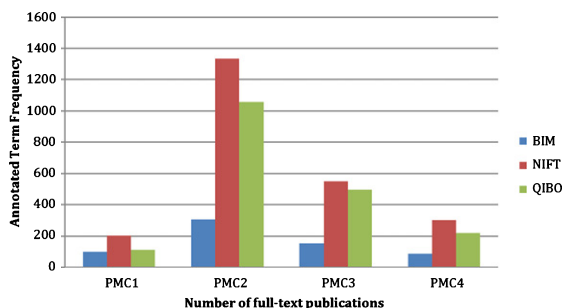


Fig. 2. Cross-validation of NIFT terminology against QIBO and BIM. The figure illustrates the evaluation of NIFT by comparing the term relevancy from NIFT, QIBO, and BIM against four full-text PubMed Central articles (PMC1, PMC2, PMC3, and PMC4).

The usability of well-annotated terminology systems can only be considered useful if they are applicable to relevant research. To assess the applicability of NIFT, we have studied the role of image

derived indices for diagnosis and how they complement the clinical assessments for better disease prognosis. One of our aims was to establish a (plausible) bridge between clinical, imaging, and cognitive tests which is not only multi-modal, but should enable disease sub-type identification and classification. Our hope is that linking imaging with anatomical as well as diagnostic readouts in AD can help to gain better insight into disease progression and thereby provide more accurate diagnoses.

However, imaging-derived indices information is often scattered throughout the largely unstructured scientific literature, which needs to be analyzed in a systematic manner. Using NIFT, we retrieved 4,029 publications. Out of the 4,029 publications, we filtered 1000 documents that contained at least one quantitative imaging feature, one neuropsychological test at clinic, and a diagnosed stage of AD in corresponding patients (see the query in the Methods section). To exclude false positive documents, we manually curated all the 1000 publications and we found 101 articles that were relevant. For manual curation of documents, we followed a 3-step procedure which are as following:

- (i) Only those articles that had informative relationship between neuropsychological assessment and radiological finding were considered for further analysis.
- (ii) Articles that only had information about either neuropsychological assessment or radiological findings were not considered for further analysis
- (iii) Articles that contained both neuropsychological assessment as well as radiological finding but did not have any meaningful relationship between them, were not considered for further analysis.

The resulting overview shows a pattern based on which imaging technologies and measured features derived from these technologies can be used to categorize the underlying clinical manifestations of patients and thereby links clinical and imaging readouts with the stage of the patient (see Table 1). The overall relation between the quantitative imaging feature, psychological feature, and diagnosis can be found in Supplementary File 2.

We also conducted a systematical analysis of the heterogeneous imaging techniques and readouts and combining them with the anatomical correlates and clinical endpoints. According to our analysis (as seen in Table 1), the medial temporal atrophy (MTA) as

Table 1
Mapping imaging derived features to clinical diagnosis and stage in Alzheimer's disease

PMID	IMAGING TECHNIQUE	IMAGING FEATURE	RELATION	NEUROPSYCHOLOGICAL ASSESSMENTS	BRAIN REGION	DIAGNOSIS
18468705	MRI	Medial Temporal Atrophy	Positive correlation	Mean visual rating scale (5.39)	Temporal lobe	Late AD
18468705	MRI	Medial Temporal Atrophy	Positive correlation	Mean visual rating scale (2.16)	Temporal lobe	Vascular dementia
18468705	MRI	Medial Temporal Atrophy	Positive correlation	Mean visual rating scale (0.56)	Temporal lobe	Healthy
11594917	MRI	Hippocampal atrophy	Correlation	Neuropsychological test ($r > 0.5$)	Hippocampus	Late AD
2061508	DTI	White matter integrity, brain atrophy	Associated	Clinical dementia rating (0.5)	Brain, white matter	Mild AD, EOAD
23880336	MRI	Cortical thinning	Associated	Montgomery Asberg Depression Rating Scale	Cortex	Mild AD, LBD
12834197	MRI	Hippocampal atrophy, corpus callosum atrophy	Associated	CDR (1 and 2)	Hippocampus	Probable AD
12834197	MRI	Hippocampal atrophy	Associated	CDR (0.5)	Corpus callosum	Mild AD
18400396	MRI	Cortical thinning	Associated	CDR (0.5)	Hippocampus Cortex	Pre-dementia

MRI, magnetic resonance imaging; DTI, diffusion tensor imaging; CDR, Clinical Dementia Rating; AD, Alzheimer's disease; EOAD, early onset Alzheimer's disease; LBD, Lewy body disease.

such is a common phenomenon observed in all the three diagnostic classes; however, they could be better distinguished as AD when the MTA score is the highest (5.39). Similarly, the atrophy could be classified as vascular dementia when the MTA score is 2.16 and in case of healthy patients, MTA could still occur, but with a very minimal score (0.56).

Analysis of literature for neuroimaging features

Mining image captions from literature

Although a fair amount of information on the image-derived findings is usually reported in the abstract of publications, specific features and interpretations gained from brain imaging experiments are often described in the caption of imaging figures that accompany the abstract text in PubMed. We, therefore, tested the relevance and performance of NIFT by applying it to a text mining scenario for analysis of figure legends extracted from publications. A typical example of figure captions annotated using NIFT terms that were extracted from PubMed abstracts is shown in Fig. 3. This figure highlights important quantitative biomarkers such as cortical ribbon which occurs due to the hyperintensity of the cortex observed in patients with early MCI and AD. This radiological sign can be detected using a diffusion tensor imaging technique and fractional anisotropy which is an important measure that demonstrates the connectivity of the brain as well as the tissue characteristics such as myelination and fiber density.

Mining full-text publications

In a separate experiment, we annotated a large corpus of full-text publications in order to examine the coverage of NIFT. A typical example for the automated annotation of a section of a full-text publication is shown in Fig. 4. This figure highlights the coverage of the NIFT terms from the full-text publication which includes neuropsychological assessments, brain regions, imaging technique as well as imaging biomarkers. This application demonstrates the usability of NIFT in mining context-specific, full-text publications in the field of neurodegeneration. Retrieval of context-specific, full-text publications can further be used to build a gold-standard corpus in the neurodegeneration domain for generation of novel hypotheses.

Annotation of image scans using NIFT

In a separate experiment, we tested, to what extent NIFT terms are suitable for the annotation of

Sentence healthy control brain imaging <1> **DTI**

Sentence Figure 4: A, B. Cortical <2> **gray matter** and pericortical tissue properties quantified along the cortical mantle in control subjects. Mean values (A) and variability (B, coefficient of variation) for <3> **cortical thickness**, cortical <4> **mean diffusivity** and <5> **fractional anisotropy** and <6> **mean diffusivity** of pericortical <7> **white matter** (2 mm below the gray-white boundary) for control subjects are shown for each point on the surfaces of the two hemispheres. The <8> **corpus callosum** is cut out, and the regions with very low values (e.g., thickness < 1 mm, coefficient of variation < 1%) appear in gray.

Sentence Figure 1: Cortical dysplasia marked by cortical thickening and gray-white boundary blurring in Patient 3.A. Significantly increased thickening of the <9> **cortical ribbon** (in red/yellow) detected by the automated quantitative approach falls within the expert-delineated area of focal cortical dysplasia (green tracing on the inflated pial surface). The area with highest z-score shows <10> **cortical thickness** measures of up to 6.5 mm between <11> **white matter** and pial surfaces. B. Calculating the T1-signal contrast at .5 mm above vs below the gray/white interface shows an area of significant blurring within the lesion. The green circle on the left shows the area of maximum blurring on a coronal <12> **volumetric MRI** slice.

Fig. 3. Annotation of an assembly of figure captions with NIFT terminology. This figure showcases the figure captions extracted from publications using NIFT terminology. The red box indicates the NIFT terms present in the figure captions.

atrophy scores (Korf et al. 2004). More recently, automated techniques to extract volumes of interest and cortical thickness values for numerous neocortical regions (Dale et al. 1999, Fischl & Dale 2000), as well as semiautomated whole-brain morphometry techniques, such as voxel-based morphometry (Ashburner & Friston 2000) and other techniques (Fan et al. 2008, Hua et al. 2008) that determine the density or volume of gray matter (GM), white matter (WM), and CSF on a voxel-by-voxel basis, have been developed and utilized in studies of brain aging and AD. More advanced structural imaging techniques, including diffusion weighted imaging and diffusion tensor imaging (DWI/DTI), magnetic resonance spectroscopy (MRS), and perfusion imaging have also been used to evaluate changes in patients with AD and those in prodromal stages. DWI/DTI techniques measure the integrity of WM tracts using two types of measures: (a) fractional anisotropy (FA), which reflects the diffusion direction of water in the fiber tracts and is thought to be a general measure of axonal integrity; and (b) mean diffusivity (MD) or apparent diffusion coefficient (ADC), which measures the overall diffusivity. Reduced FA and increased MD/ADC are considered to be markers of neuronal fiber loss and WM atrophy. MRS is a noninvasive technique allowing the measurement of biological metabolites in target tissue that has been used in studies of brain aging and AD. Two major metabolites that consistently show alterations in patients with AD and MCI include N-acetylaspartate (NAA) and myo-inositol. Finally, two MRI methods have traditionally been employed to measure cerebral perfusion in studies of AD, including dynamic susceptibility contrast enhanced MRI, which involves the injection of contrast agents to measure CBF and regional cerebral blood volume, and a relatively new technique called arterial spin labeling, which measures CBF without any external contrast agents by using magnetic pulses that label blood entering the brain (Alsop et al. 2010, Wang et al. 2011).

MRI can also be used to measure brain function. Functional MRI (fMRI) measures brain activity during a cognitive, sensory, or motor task or at rest by measuring blood flow and blood oxygen levels. The primary outcome measured in fMRI studies is a blood-oxygenation-level-dependent (BOLD) signal in which regional brain activity is assessed via changes in local blood flow and oxygenation (Ogawa et al. 1992). Given that activity-related brain metabolism is tightly coupled to regional blood oxygenation and flow (i.e., blood flow increases to keep the regional blood oxygen level high during brain activation and associated increases in metabolic demand), the BOLD signal is a useful measure for brain activation (Logothetis et al. 2001). However, altered coupling of neuronal metabolism and blood flow due to brain atrophy and/or hypoperfusion may cause alterations in the BOLD signal. Therefore, studies in older patient populations with brain atrophy and hypoperfusion, such as MCI and AD, should be evaluated and interpreted with these considerations in mind.

Fig. 4. Annotation of a section of a full-text article using the NIFT terminology. The ProMiner tagger was used to identify NIFT terms in full text; matching terms are marked up in red.

primary neuroimaging data (brain scans). Figure 5 depicts the comparison between control, MCI, and AD patient brain scans based on: the amyloid burden through [18] AV45-PET, the regional pattern of hypometabolism through FDG-PET, and hippocampal

volumetry as well as cortical thickness through T13D MP-RAGE brain scans. The top part of the figure shows the amyloid burden and the hypometabolic clusters across the different brain regions. As it can be seen in Fig. 5, the control does not have any

amyloid deposit cluster and no hypometabolism detected, while in MCI, hypometabolic pattern starts to appear in the left hemisphere and more extensively in AD. The expected hypometabolic topography spread across the temporo-parietal regions, precuneus, and posterior cingulate cortex. All the patient-derived image scans can be found in Supplementary File 3.

Mechanistic modeling of imaging features in the context of AD pathology

Generating links between molecular entities and imaging modalities, even if very demanding and complex, could provide interesting insights into the disease progression as well as help to raise our understanding of the underlying pathology. On that note, we tried to establish that link by querying our in-house SCAIView tool for genes/proteins relevant to imaging features (See Methods section). We retrieved 1,853 gene/protein entities, out of which we identified the top 20 entities confined to interesting brain regions such as cortex, hippocampus, temporal lobe, and cerebrum. Using these entities, we produced the ‘global map’ of genes and imaging features (See Fig. 6). We also inferred from this model that these top ranking genes play a role in cortical thickness, hippocampal atrophy, temporal lobe atrophy, grey matter atrophy and cerebral atrophy, as follows.

Cortical thickness

Our systematic analysis of the literature revealed that many key players contributed to thinning of the cortex, which is a strong indicator of AD progression. In the following, we demonstrate lines of evidence about factors causally involved in or correlated with cortical thinning and exemplify their corresponding BEL codes:

- Increased expression of APP results in the accumulation of A β , which affects the thinning of the cortex [35, 36].

```
p(HGNC:APP) -> a("Amyloid beta-Peptides")
a("Amyloid beta-Peptides") - a(NIFT: "Cortical thinning")
a(NIFT: "Cortical thinning") -> path(MESHID: "Alzheimer Disease")
```

- Increased expression of CHI3L1, a gene responsible for inflammatory response [37, 38], was found to be correlated with cortical thickness [39].

```
p(HGNC:CHI3L1) -> bp(GO:"inflammatory response")
bp(GO:"inflammatory response")
negativeCorrelation a(NIFT: "Cortical thickness")
```

- PSEN1 was found to cause neuronal loss [40, 41], which results in the shrinkage of the cortex due to neuronal injury [42, 43].

```
p(HGNC:PSEN1) -> bp(GO:"neuronal loss")
bp(GO:"neuronal loss") -> a(NIFT: "Cortical thinning")
```

- Well-known genes such as APOE4 along with APOE4 and BCHE carriers contributed to the structural alteration of the cortex, resulting in cortical thinning [44–47].

```
p(HGNC:APOE) -> a(NIFT: "Cortical thinning")
p(HGNC:BCHE) -> a(NIFT: "Cortical thinning")
```

- Some genes can be linked to cortical thinning through genetics approaches: genes such as FJ10357 [48], TOMM40 [49], and BDNF [50, 51] play a protective role in preserving the structure of the cortex, however, their genetic alteration results in cortical thinning—rs3748348, rs10524523 and rs6265, respectively.

```
p(HGNC: FJ10357) -> a(NIFT: "Cortical thickness")
g(dbSNP: rs3748348) - a(NIFT: "Cortical thinning")
p(HGNC: TOMM40) -| a(NIFT: "Cortical thickness")
g(dbSNP: rs10524523) - a(NIFT: "Cortical thinning")
p(HGNC: BDNF) -> a(NIFT: "Cortical thickness")
g(dbSNP: rs6265) - a(NIFT: "Cortical thinning")
```

Temporal lobe atrophy

We investigated two genes, APOE ϵ 4 and TREM2, which mainly contribute to the atrophy of temporal lobes. TREM2 is an inflammatory response gene predominantly found in microglia [52, 53]. They are known to enhance phagocytosis as well

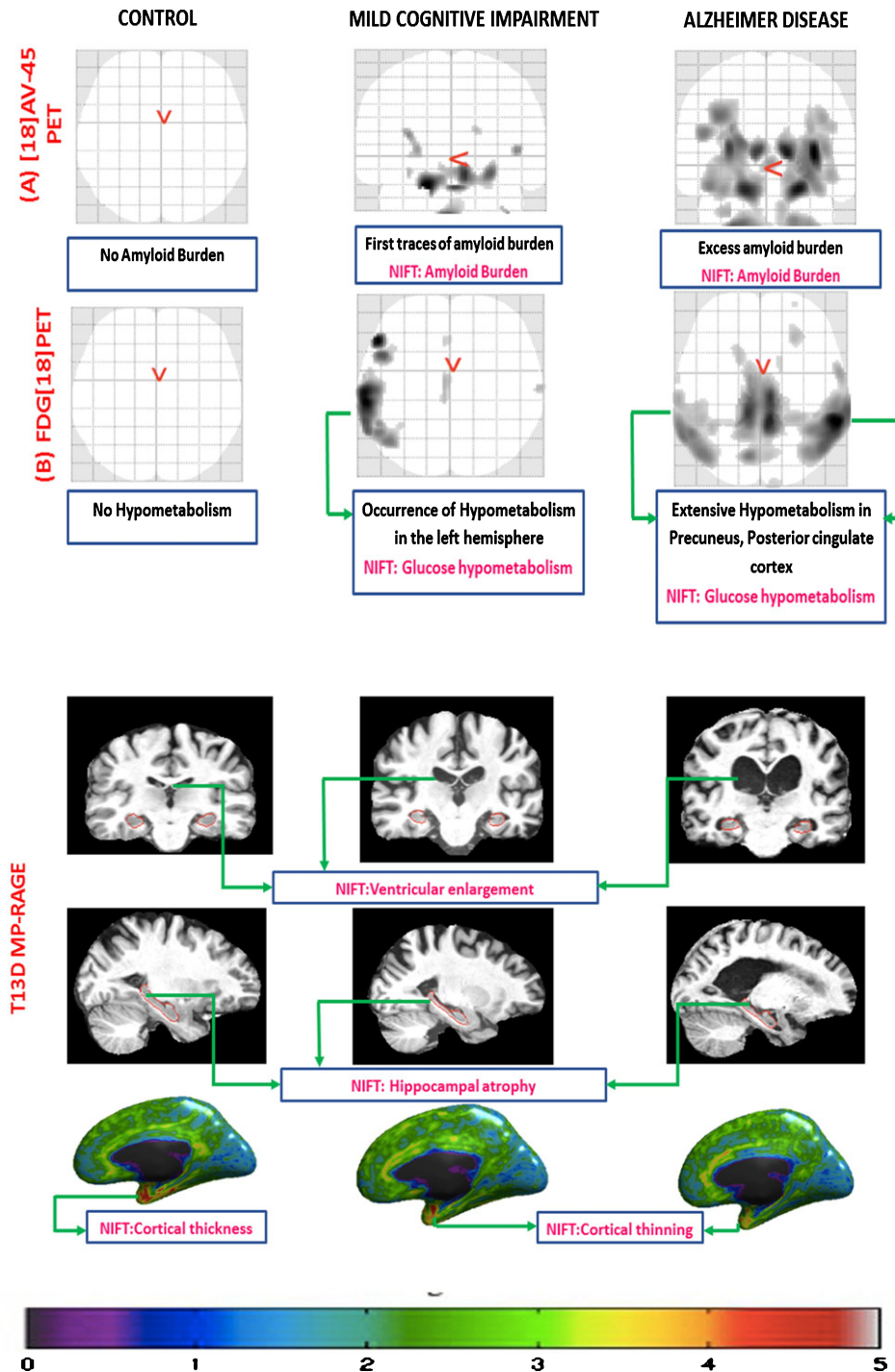


Fig. 5. Manual annotation of brain image scans using NIFT. This figure represents different biomarkers captured using three different imaging techniques in control, mild cognitive impairment (MCI), and AD respectively. A) [18] AV-45 PET scan: this figure captures the increased amount of amyloid burden (p -value threshold 0.001; voxel extend 10; smoothing kernel [8-8-8]) during the disease progression across CN, MCI, and AD, respectively. B) FDG [18] PET: this figure captures no hypometabolism in control, increased hypometabolic pattern in case of MCI, and extensive hypometabolic topography in the temporo-parietal regions, precuneus, and posterior cingulate cortex (p -value threshold 0.001; voxel extend 10; smoothing kernel [8-8-8]). C) T13D MP-RAGE: the first row of the figure demonstrates the progressive ventricular enlargement among control, MCI, and AD respectively. The second row represents progressive hippocampal atrophy across control, MCI, and AD. The third row represents progressive cortical shrinkage in the temporal-parietal lobe, posterior cingulate and precuneus area.

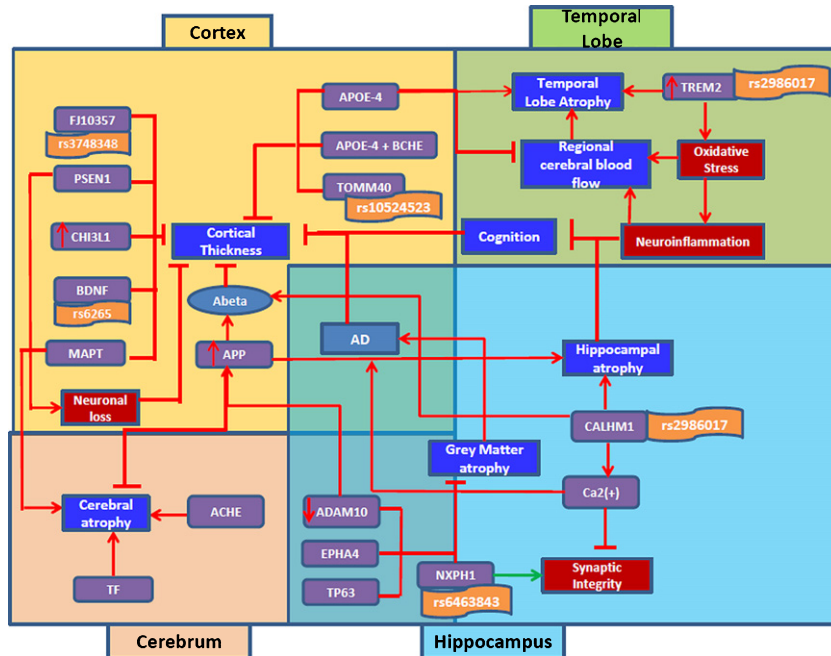


Fig. 6. Integrative view of literature-derived associations between molecular and clinical indices in AD through image-derived features. This figure illustrates the complex interaction of genetic players playing a causative/protective role in underlying disease pathology through neuroimaging indices. Top left part of the figure key genetic factors that play a role in shrinking of the cortex eventually leading to AD; top right part of the figure consists of genes involved in neuro-inflammation and temporal lobe atrophy; Bottom left part of the figure displays genes involved in cerebral atrophy; bottom right part consists of genes playing a role in hippocampal and gray matter atrophy. The red color symbol (-) indicates perturbation of a gene. The red color arrow indicates the function of a gene in disease condition. The green arrow represents the normal process.

as maintaining cytokine production so that inflammatory responses can be triggered by TREM-1, a novel receptor expressed on neutrophils and monocytes [54, 55]. However, the genetic mutation of TREM2, rs75932628, causes the atrophy of temporal lobes through enhancing oxidative stress [56], which in turn causes the reduction of cerebral blood flow leading to reduction in regular supply of oxygen, glucose, and other nutrients to the temporal lobe, and finally the shrinkage of the temporal lobe [57, 58]. On the other hand, increased expression of APOE ϵ 4 allele affects the flow of cerebral blood, further contributing to atrophy [59, 60].

Hippocampal atrophy

Through our work, we identified an interesting gene, CALHM1 which was known to regulate A β clearance [61] through the activation of insulin-degrading enzyme [62]. However, a genetic mutation by rs2986017 results in (i) loss of hippocampal neurons further causing atrophy as well as (ii) increased A β levels and altered calcium homeostasis which

could result in reduced synaptic integrity and mitochondrial dysfunction [63].

Grey matter atrophy

Here, we identified a gene namely NXPH1, which was found to play a role in adhesion of dendrites and axons and maintaining synaptic integrity. However, the mutation of the gene, rs6463843, affects the synaptic integrity and results in loss of grey matter density leading to atrophy [64]. Apart from that, EPHA4 was also found to play a protective role in the glial glutamate transport that ultimately regulates hippocampal function as well as the maintenance of grey matter density [65, 66].

Cerebral atrophy

Cerebral atrophy was found to be regulated by two key players—Transferrin as well as ACHE. Transferrin was found to play a significant role in iron homeostasis [67]. However, the alteration of the gene could result in iron overload which causes damage to the cerebral structure, ultimately leading to cerebral

atrophy [68, 69]. On the other hand, the altered function of ACHE in the cholinergic system could result in the loss of cerebral neurons leading to cerebral atrophy [70]. The computer readable BEL version of this model is found in Supplementary File 4.

DISCUSSION

A “biomarker” is an accurately measured medical sign that indicates the medical state of the patient. However, in the field of imaging, this term is often misinterpreted due to the lack of standardization of terminology and concepts. Often, literature resources misclassify an imaging technique as a biomarker while many others denote the derived indices as a biomarker [71, 72]. Owing to this, NIFT was constructed to represent, integrate, and harmonize heterogeneous knowledge across the domain of imaging biomarkers in the context of neurodegeneration. NIFT serves as a single resource of the standard terminology that describes the domain of neuroimaging biomarkers in a hierarchical manner and has been designed to capture relevant image-derived features (“indices”) with high specificity and granularity. As shown by our analysis, what distinguishes NIFT from other existing resources is the inclusion of various concepts ranging from algorithms that automate the process of measuring features to radiological tracers that help in revealing functional alterations of the brain. Such a standard reference terminology has the potential not only to support organization and exchange of imaging information among neurologists and clinical researchers but also to provide a useful tool for annotation of brain scan metadata as detection of meta-information in brain scans helps inferring neuroanatomical relationships present in imaging data [73]. With such an inventory, it is indeed possible to automatically extend the annotation of scans by incorporating NIFT in image annotation tools. Since NIFT combines specificity and granularity of imaging features in the context of neurology knowledge domain, users can intuitively navigate through different levels of concept granularity within a search engine and for instance, perform faceted search in literature mining environments.

With respect to the contextual specificity, as benchmark analysis of NIFT against two other highly domain-specific, relevant terminologies showed the overall granularity of medical relevant terms and cognitive tests in NIFT was comparably high, making NIFT a reference terminology resource specific to

neuroimaging. The applicability of NIFT could be extended toward information retrieval and extraction. As demonstrated earlier, using NIFT for literature mining improves retrieval of the relevant, informative neuroimaging publications and supports curation and extraction of captured information from unstructured text. In the presence of other terminology sources, powerful filtering for faceted searches can be implemented. For instance, we can combine NIFT with HypothesisFinder [74] to systematically harvest speculative statements linked to imaging features; or combination of NIFT terms with ADO terms will allow us to systematically harvest factual statements that link imaging readouts to aspects of AD progression in literature; and finally, we also have the possibility of mining “shared imaging features” amongst other diseases by making use of the already integrated Parkinson Disease Ontology [75] and Multiple Sclerosis Ontology [76]. This could lead to domain specific imaging feature identification across disease scales.

Importantly, the usage of NIFT is not limited to information retrieval and extraction. Since the major mission of the neuroscience community currently is to bridge the gap between molecular mechanisms and imaging readouts, NIFT can be used to address this challenge by bringing context to computational modeling efforts. To demonstrate this possibility, we showed how NIFT serves as a valuable resource to support mechanistic modeling of complex AD pathomechanisms. As highlighted in Fig. 6, this high resolution mechanistic model captures novel genetic players such as CALHM1, NXP1, and ADAM10, which cause hippocampal atrophy through neuronal loss. Here, we identified the various roles played by CALHM1 in AD pathology, ranging from controlling cytosolic Ca(2+) concentrations and A β levels to increased oxidative stress through glutamatergic neurotransmission inhibition [77, 78]. Similarly, another two novel genetic biomarkers were CHI3L1 and CAND1. CHI3L1, a protein that encodes YKL-40, was found to be associated with cortical thinning and was found to play a role in neuroinflammatory response. They were found to play a role in cell morphology and behavior; however, their association with susceptibility to AD has only been recently studied [79].

The current neuropathological studies on AD suggest that the clinical onset of the disease goes decades before the formation of neurofibrillary tangles and A β plaques [80, 81]. This brings up the need for heterogeneous measurable indicators that can aid

systematic tracing of alternative patterns of disease progression. ADNI have positioned themselves as pioneers in assembling patient records with cognitive and longitudinal assessments along with genetic and fluid sample measures. This interesting combination of measured metadata could provide unique insights into measurable signs before the expected onset of the disease. However, the challenge still remains to identify those patterns at an earlier stage through the use of combinatorial features. Furthermore, we foresee the option to perform systematic association studies in the literature between SNPs and mutations on one side and imaging features on the other side (Iyappan et al., in preparation). The multi-level association between genetic factors and clinical readouts can be directly used for modeling and mining across scales in the neurology and psychiatry field. A first attempt of demonstrating such systematic harvesting approach is the association of imaging features with cognition readouts (refer to Table 1). Such associations lead to comprehensive analysis of imaging features correlating with cognition.

To the best state of our knowledge, NIFT is the first reference compendium, which apprehends the various aspects of the derived quantitative measures from neuroimaging scans. We invite the scientific community to contribute to edition and enrichment of NIFT so that it can be leveraged to the level of a formal ontology in future.

Conclusion

To our knowledge, there have been little efforts invested so far in the direction of standardizing and capturing observable clinical imaging features, particularly in the neurology domain. Through this work, we attempted to bridge “omics” and imaging/clinical level data. This type of integration across scales is often regarded as the “holy grail” of integrative modeling and mining. Future approaches should be able to represent and model the disease progression in a longitudinal model by integrating molecular processes and imaging features over time, provided that longitudinal data capture would be extended to other omics data types beside imaging. For this purpose, we obviously need trajectories. Currently, the BEL modeling framework does not deliver this time dimension. However, we are working towards the extension of BEL by a time dimension. A long term perspective of this extension is the vision of a virtual patient cohort that comprises several such longitudinal “trajectories” representing the dynamics of important imaging

features. The link between imaging and genetics will be a cornerstone for the construction of the virtual cohort; the generation of a “virtual dementia cohort” has recently been made a task in IMI-project AETIONOMY and we will see first results of the simulation of entire trials based on a “virtual dementia cohort” in the near future. The imaging derived features captured through NIFT will have a major role in that “virtual dementia cohort”. We believe it would be desirable to generate a “metadata atlas” of the brain populated with NIFT concepts. Such an atlas could serve as a template for qualitative models that integrate imaging features from different, heterogeneous studies.

ACKNOWLEDGMENTS

The authors wish to acknowledge Frisoni’s group for curation; Bernard de Bono for constructive discussions and advice; and students from the Life Science Informatics curriculum at Bonn-Aachen International Center for Information Technology (BIT) for their support with curation and annotation tasks.

Data collection and sharing for this project was funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer’s Association; Alzheimer’s Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the

National Institutes of Health (<http://www.fnih.org>). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under AETIONOMY grant agreement n°115568, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution.

Authors' disclosures available online (<http://j-alz.com/manuscript-disclosures/16-1148r2>).

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: <http://dx.doi.org/10.3233/JAD-161148>.

REFERENCES

- Cash DM, Rohrer JD, Ryan NS, Ourselin S, Fox NC (2014) Imaging endpoints for clinical trials in Alzheimer's disease. *Alzheimers Res Ther* **6**, 87.
- Ishii K (2014) PET approaches for diagnosis of dementia. *AJNR Am J Neuroradiol* **35**, 2030-2038.
- Younesi E, Hofmann-Apitius M (2013) Biomarker-guided translation of brain imaging into disease pathway models. *Sci Rep* **3**, 3375.
- Kassab GS, An G, Sander EA, Miga MI, Guccione JM, Ji S, Vodovotz Y (2016) Augmenting surgery via multi-scale modeling and translational systems biology in the era of precision medicine: A multidisciplinary perspective. *Ann Biomed Eng* **44**, 2611-2625.
- De Bono B, Safaei S, Grenon P, Nickerson DP, Alexander S, Helvensteijn M, Kok JN, Kokash N, Wu A, Yu T, Hunter P, Baldock RA (2015) The open physiology workflow: Modeling processes over physiology circuitboards of interoperable tissue units. *Front Physiol* **6**, 24.
- Gibaud B, Kassel G, Dojat M, Batrancourt B, Michel F, Gaignard A, Montagnat J (2011) NeuroLOG: Sharing neuroimaging data using an ontology-based federated approach. *AMIA Annu Symp Proc* **2011**, 472-480.
- Langlotz CP (2006) RadLex: A new method for indexing online educational materials. *Radiographics* **26**, 1595-1597.
- Buckler AJ, Ouellette M, Danagoulian J, Wernsing G, Liu TT, Savig E, Suzek BE, Rubin DL, Paik D (2013) Quantitative imaging biomarker ontology (QIBO) for knowledge representation of biomedical imaging biomarkers. *J Digit Imaging* **26**, 630-641.
- Soucy J-P, Bartha R, Bocti C, Borrie M, Burhan AM, Laforce R, Rosa-Neto P (2013) Clinical applications of neuroimaging in patients with Alzheimer's disease: A review from the Fourth Canadian Consensus Conference on the Diagnosis and Treatment of Dementia 2012. *Alzheimers Res Ther* **5**, S3.
- Gispert JD, Rami L, Sánchez-Benavides G, Falcon C, Tucholka A, Rojas S, Molinuevo JL (2015) Nonlinear cerebral atrophy patterns across the Alzheimer's disease continuum: Impact of APOE4 genotype. *Neurobiol Aging* **36**, 2687-2701.
- Whitwell JL, Jack CR, Knopman DS, Boeve BF, Petersen RC, Josephs KA, Parisi JE, Dickson DW, Murray ME, Whitwell JL, Dickson DW, Murray ME, Weigand SD, Tosakulwong N, Senjem ML, Knopman DS, Boeve BF, Parisi JE, Petersen RC, Ord C, Jack R, Josephs KA (2012) Neuroimaging correlates of pathologically defined subtypes of Alzheimer's disease: A case-control study. *Lancet Neurol* **11**, 868-877.
- Kahn CE, Rubin DL (2009) Automated semantic indexing of figure captions to improve radiology image retrieval. *J Am Med Inform Assoc* **16**, 380-386.
- Kahn CE (2014) Annotation of figures from the biomedical imaging literature: A comparative analysis of radlex and other standardized vocabularies. *Acad Radiol* **21**, 384-392.
- Warden GI, Lacson R, Khorasani R (2011) Leveraging terminologies for retrieval of radiology reports with critical imaging findings. *AMIA Annu Symp Proc* **2011**, 1481-2588.
- Tao C, Pathak J, Solbrig HR, Wei W-Q, Chute CG (2013) Terminology representation guidelines for biomedical ontologies in the semantic web notations. *J Biomed Inform* **46**, 128-138.
- Hofmann-Apitius M (2015) Is dementia research ready for big data approaches? *BMC Med* **13** 145.
- Poldrack RA, Gorgolewski KJ (2014) Making big data open: Data sharing in neuroimaging. *Nat Neurosci* **17**, 1510-1517.
- Breeze JL, Poline J-B, Kennedy DN (2012) Data sharing and publishing in the field of neuroimaging. *Gigascience* **1**, 9.
- Buckler AJ, Ouellette M, Danagoulian J, Wernsing G, Liu TT, Savig E, Suzek BE, Rubin DL, Paik D (2013) Quantitative imaging biomarker ontology (QIBO) for knowledge representation of biomedical imaging biomarkers. *J Digit Imaging* **26**, 630-641.
- Vrooman HA, Cocosco CA, van der Lijn F, Stokking R, Ikram MA, Vernooij MW, Breteler MMB, Niessen WJ (2007) Multi-spectral brain tissue segmentation using automatically trained k-Nearest-Neighbor classification. *Neuroimage* **37**, 71-81.
- Fortunati V, Verhaart RF, van der Lijn F, Niessen WJ, Veenland JF, Paulides MM, van Walsum T (2013) Tissue segmentation of head and neck CT images for treatment planning: A multiatlas approach combined with intensity modeling. *Med Phys* **40**, 71905.
- van der Lijn F, den Heijer T, Breteler MMB, Niessen WJ (2008) Hippocampus segmentation in MR images using atlas registration, voxel classification, and graph cuts. *Neuroimage* **43**, 708-720.
- de Boer R, Schaap M, van der Lijn F, Vrooman HA, de Groot M, van der Lugt A, Ikram MA, Vernooij MW, Breteler MMB, Niessen WJ (2011) Statistical analysis of minimum cost path based structural brain connectivity. *Neuroimage* **55**, 557-565.
- Guyader JM, Bernardin L, Douglas NH, Poot DH, Niessen WJ, Klein S (2015) Influence of image registration on apparent diffusion coefficient images computed from free-breathing diffusion MR images of the abdomen. *J Magn Reson Imaging* **42**, 315-330.

- [25] Daducci A, Gerhard S, Griffa A, Lemkaddem A, Cammoun L, Gigandet X, Meuli R, Hagmann P, Thiran JP (2012) The Connectome Mapper: An open-source processing pipeline to map connectomes with MRI. *PLoS One* **7**, e48121.
- [26] Malhotra A, Younesi E, Gündel M, Müller B, Heneka MT, Hofmann-Apitius M (2014) ADO: A disease ontology representing the domain knowledge specific to Alzheimer's disease. *Alzheimers Dement* **10**, 238-246.
- [27] Iyappan A, Gündel M, Shahid M, Wang J, Li H, Mevissen HT, Müller B, Fluck J, Jirsa V, Domide L, Younesi E, Hofmann-Apitius M (2016) Towards a pathway inventory of the human brain for modeling disease mechanisms underlying neurodegeneration. *J Alzheimers Dis* **52**, 1343-1360.
- [28] Bukhari AC, Nagy ML, Krauthammer M, Ciccarese P, Baker CJO (2015) BIM: An open ontology for the annotation of biomedical images. *CEUR Workshop Proc* **1515**, 1-5.
- [29] Hofmann-Apitius M, Fluck J, Furlong L, Fornes O, Kolárik C, Hanser S, Boeker M, Schulz S, Sanz F, Klinger R, Mevissen T, Gattermayer T, Oliva B, Friedrich CM (2008) Knowledge environments representing molecular entities for the virtual physiological human. *Philos Trans A Math Phys Eng Sci* **366**, 3091-3110.
- [30] Hanisch D, Fundel K, Mevissen H-T, Zimmer R, Fluck J (2005) ProMiner: Rule-based protein and gene entity recognition. *BMC Bioinformatics* **6**(Suppl 1), S14.
- [31] Trojanowski JQ1, Vandeersticbele H, Korecka M, Clark CM, Aisen PS, Petersen RC, Blennow K, Soares H, Simon A, Lewczuk P, Dean R, Siemers E, Potter WZ, Weiner MW, Jack CR Jr, Jagust W, Toga AW, Lee VM, Shaw LM; Alzheimer's Disease Neuroimaging Initiative (2010) Update on the biomarker core of the Alzheimer's Disease Neuroimaging Initiative subjects. *Alzheimers Dement* **6**, 230-238.
- [32] Risacher SL, Saykin AJ (2013) Neuroimaging and other biomarkers for Alzheimer's disease: The changing landscape of early detection. *Annu Rev Clin Psychol* **9**, 621-648.
- [33] Risacher SL, Saykin AJ (2013) Neuroimaging biomarkers of neurodegenerative diseases and dementia. *Semin Neurol* **33**, 386-416.
- [34] Shokouhi S, Claassen D, Riddle W (2014) Imaging brain metabolism and pathology in Alzheimer's disease with positron emission tomography. *J Alzheimers Dis Parkinsonism* **4**, pii: 143.
- [35] Becker JA, Hedden T, Carmasin J, Maye J, Rentz DM, Putcha D, Fischl B, Greve DN, Marshall GA, Salloway S, Marks D, Buckner RL, Sperling RA, Johnson KA (2011) Amyloid- β associated cortical thinning in clinically normal elderly. *Ann Neurol* **69**, 1032-1042.
- [36] Kim CM, Hwang J, Lee JM, Roh JH, Lee JH, Koh JY; Alzheimer's Disease Neuroimaging Initiative (ADNI) (2015) Amyloid beta-weighted cortical thickness: A new imaging biomarker in Alzheimer's disease. *Curr Alzheimer Res* **12**, 563-571.
- [37] Bhardwaj R, Yester JW, Singh SK, Biswas DD, Surace MJ, Waters MR, Hauser KF, Yao Z, Boyce BF, Kordula T (2015) RelB/p50 complexes regulate cytokine-induced YKL-40 expression. *J Immunol* **194**, 2862-2870.
- [38] Melah KE, Lu SY-F, Hoscheidt SM, Alexander AL, Adluru N, Destiche DJ, Carlsson CM, Zetterberg H, Blennow K, Okonkwo OC, Gleason CE, Dowling NM, Bratzke LC, Rowley HA, Sager MA, Asthana S, Johnson SC, Bendlin BB (2016) Cerebrospinal fluid markers of Alzheimer's disease pathology and microglial activation are associated with altered white matter microstructure in asymptomatic adults at risk for Alzheimer's disease. *J Alzheimers Dis* **50**, 873-886.
- [39] Antonell A, Mansilla A, Rami L, Lladó A, Iranzo A, Olives J, Balasa M, Sánchez-Valle R, Molinuevo JL (2014) Cerebrospinal fluid level of YKL-40 protein in preclinical and prodromal Alzheimer's disease. *J Alzheimers Dis* **42**, 901-908.
- [40] Verdile G, Gnjec A, Miklossy J, Fonte J, Veurink G, Bates K, Kakulas B, Mehta PD, Milward EA, Tan N, Lareu R, Lim D, Dharmarajan A, Martins RN (2004) Protein markers for Alzheimer disease in the frontal cortex and cerebellum. *Neurology* **63**, 1385-1392.
- [41] Scahill RI, Ridgway GR, Bartlett JW, Barnes J, Ryan NS, Mead S, Beck J, Clarkson MJ, Crutch SJ, Schott JM, Ourselin S, Warren JD, Hardy J, Rossor MN, Fox NC (2013) Genetic influences on atrophy patterns in familial alzheimer's disease: A comparison of APP and PSEN1 mutations. *J Alzheimers Dis* **35**, 199-212.
- [42] Knight WD, Kim LG, Douiri A, Frost C, Rossor MN, Fox NC (2011) Acceleration of cortical thinning in familial Alzheimer's disease. *Neurobiol Aging* **32**, 1765-1773.
- [43] Balasa M, Vidal-Piñero D, Lladó A, Antonell A, Bosch B, Castellanos F, Bargalló N, Bartres-Faz D, Molinuevo JL, Sánchez-Valle R (2012) PSEN1 mutation carriers present lower cerebrospinal fluid amyloid- β 42 levels than sporadic early-onset Alzheimer's disease patients but no differences in neuronal injury biomarkers. *J Alzheimers Dis* **30**, 605-616.
- [44] Gutiérrez-Galve L, Lehmann M, Hobbs NZ, Clarkson MJ, Ridgway GR, Crutch S, Ourselin S, Schott JM, Fox NC, Barnes J (2009) Patterns of cortical thickness according to APOE genotype in Alzheimer's disease. *Dement Geriatr Cogn Disord* **28**, 476-485.
- [45] Donix M, Burggren AC, Scharf M, Marschner K, Suthana NA, Siddarth P, Krupa AK, Jones M, Martin-Harris L, Ercoli LM, Miller KJ, Werner A, von Kummer R, Sauer C, Small GW, Holthoff VA, Bookheimer SY (2013) APOE associated hemispheric asymmetry of entorhinal cortical thickness in aging and Alzheimer's disease. *Psychiatry Res* **214**, 212-220.
- [46] Yoo H, Lee H, Shin S, Park S-W, Choi J, Jung H, Cha J, Lee J-M, Lee J-Y (2014) Butyrylcholinesterase K and apolipoprotein E4 affect cortical thickness and neuropsychiatric symptoms in Alzheimer's disease. *Curr Alzheimer Res* **11**, 137-144.
- [47] Foley JM, Salat DH, Stricker NH, McGlinchey RE, Milberg WP, Grande LJ, Leritz EC (2016) Glucose dysregulation interacts with APOE- E4 to potentiate temporoparietal cortical thinning. *Am J Alzheimers Dis Other Demen* **31**, 76-86.
- [48] Mukherjee S, Kim S, Gibbons LE, Nho K, Risacher SL, Glymour MM, Habeck C, Lee GJ, Mormino E, Ertekin-Taner N, Montine TJ, Decarli C, Saykin AJ, Crane PK (2012) Genetic architecture of resilience of executive functioning. *Brain Imaging Behav* **6**, 621-633.
- [49] Laczó J, Andel R, Vyhnaek M, Matoska V, Kaplan V, Nedelska Z, Lerch O, Gazova I, Moffat SD, Hort J (2015) The effect of TOMM40 on spatial navigation in amnesic mild cognitive impairment. *Neurobiol Aging* **36**, 2024-2033.
- [50] Yang X, Liu P, Sun J, Wang G, Zeng F, Yuan K, Liu J, Dong M, von Deneen KM, Qin W, Tian J (2012) Impact of brain-derived neurotrophic factor Val66Met polymorphism on cortical thickness and voxel-based morphometry in healthy Chinese young adults. *PLoS One* **7**, e37777.

- [51] Zugman A, Pedrini M, Gadelha A, Kempton MJ, Noto CS, Mansur RB, Santoro ML, Gama CS, Bressan RA, McGuire P, Jackowski AP, Brietzke E (2015) Serum brain-derived neurotrophic factor and cortical thickness are differently related in patients with schizophrenia and controls. *Psychiatry Res* **234**, 84-89.
- [52] Guerreiro R, Wojtas A, Bras J, Carrasquillo M, Rogaeva E, Majounie E, Cruchaga C, Sassi C, Kauwe JSK, Younkin S, Hazrati L, Collinge J, Pocock J, Lashley T, Williams J, Lambert J-C, Amouyel P, Goate A, Rademakers R, Morgan K, Powell J, St George-Hyslop P, Singleton A, Hardy J, Alzheimer Genetic Analysis Group (2013) TREM2 variants in Alzheimer's disease. *N Engl J Med* **368**, 117-127.
- [53] Zhong L, Chen XF, Zhang ZL, Wang Z, Shi XZ, Xu K, Zhang YW, Xu H, Bu G (2015) DAP12 stabilizes the C-terminal fragment of the triggering receptor expressed on myeloid cells-2 (TREM2) and protects against LPS-induced pro-inflammatory response. *J Biol Chem* **290**, 15866-15877.
- [54] Průcha M, Zazula R, Müller M, Hyánek T, Dostál M, Sedláčková L (2011) TREM-1 expression on monocytes is not a parameter specific for infectious etiology of systemic inflammatory response syndrome. *Prague Med Rep* **112**, 205-215.
- [55] Rohn TT (2013) The triggering receptor expressed on myeloid cells 2: "TREM-ming" the inflammatory component associated with Alzheimer's disease. *Oxid Med Cell Longev* **2013**, 860959.
- [56] Jonsson T, Stefansson H, Steinberg S, Jonsdottir I, Jonsson PV, Snaedal J, Bjornsson S, Huttenlocher J, Levey AI, Lah JJ, Rujescu D, Hampel H, Giegling I, Andreassen OA, Engedal K, Ulstein I, Djurovic S, Ibrahim-Verbaas C, Hofman A, Ikram MA, van Duijn CM, Thorsteinsdottir U, Kong A, Stefansson K (2013) Variant of TREM2 associated with the risk of Alzheimer's disease. *N Engl J Med* **368**, 107-116.
- [57] Luis EO, Ortega-Cubero S, Lamet I, Razquin C, Cruchaga C, Benitez BA, Lorenzo E, Irigoyen J, Pastor MA, Pastor P (2014) Frontobasal gray matter loss is associated with the TREM2 p. R47H variant. *Neurobiol Aging* **35**, 2681-2690.
- [58] Lue LF, Schmitz CT, Serrano G, Sue LI, Beach TG, Walker DG (2015) TREM2 protein expression changes correlate with Alzheimer's disease neurodegenerative pathologies in post-mortem temporal cortices. *Brain Pathol* **25**, 469-480.
- [59] Tanaka S, Kawamata J, Shimohama S, Akaki H, Akiyuchi I, Kimura J, Ueda K (1998) Inferior temporal lobe atrophy and APOE genotypes in Alzheimer's disease. X-ray computed tomography, magnetic resonance imaging and Xe-133 SPECT studies. *Dement Geriatr Cogn Disord* **9**, 90-98.
- [60] Tai LM, Thomas R, Marottoli FM, Koster KP, Kanekiyo T, Morris AWJ, Bu G (2016) The role of APOE in cerebrovascular dysfunction. *Acta Neuropathol* **131**, 709-723.
- [61] Moreno-Ortega AJ, Ruiz-Nuño A, García AG, Cano-Abad MF (2010) Mitochondria sense with different kinetics the calcium entering into HeLa cells through calcium channels CALHM1 and mutated P86L-CALHM1. *Biochem Biophys Res Commun* **391**, 722-726.
- [62] Vingtdoux V, Chandakkar P, Zhao H, Blanc L, Ruiz S, Marambaud P (2015) CALHM1 ion channel elicits amyloid- clearance by insulin-degrading enzyme in cell lines and in vivo in the mouse brain. *J Cell Sci* **128**, 2330-2338.
- [63] Yao J, Irwin RW, Zhao L, Nilsen J, Hamilton RT, Brinton RD (2009) Mitochondrial bioenergetic deficit precedes Alzheimer's pathology in female mouse model of Alzheimer's disease. *Proc Natl Acad Sci U S A* **106**, 14670-14675.
- [64] Shen L, Kim S, Risacher SL, Nho K, Swaminathan S, West JD, Foroud T, Pankratz N, Moore JH, Sloan CD, Huentelman MJ, Craig DW, DeChairo BM, Potkin SG, Jack CR, Weiner MW, Saykin AJ (2010) Whole genome association study of brain-wide imaging phenotypes for identifying quantitative trait loci in MCI and AD: A study of the ADNI cohort. *Neuroimage* **53**, 1051-1063.
- [65] Carmona MA, Murai KK, Wang L, Roberts AJ, Pasquale EB (2009) Glial ephrin-A3 regulates hippocampal dendritic spine morphology and glutamate transport. *Proc Natl Acad Sci U S A* **106**, 12524-12529.
- [66] Braskie MN, Ringman JM, Thompson PM (2011) Neuroimaging measures as endophenotypes in Alzheimer's disease. *Int J Alzheimers Dis* **2011**, 1-15.
- [67] Pawlowski JW, Kellicker N, Bobst CE, Kaltashov I (2016) Assessing the iron delivery efficacy of transferrin in clinical samples by native electrospray ionization mass spectrometry. *Analyst* **141**, 853-861.
- [68] Printy BP, Verma N, Cowperthwaite MC, Markey MK (2014) Effects of genetic variation on the dynamics of neurodegeneration in Alzheimer's disease. *Conf Proc IEEE Eng Med Biol Soc* **2014**, 2464-2467.
- [69] Obirikorang C, Issahaku RG, Osakunor DNM, Osei-Yeboah J (2016) Anaemia and iron homeostasis in a cohort of HIV-infected patients: A cross-sectional study in Ghana. *AIDS Res Treat* **2016**, 1623094.
- [70] Silver JM, McAllister TW, Yudofsky SC (2005) *Textbook of traumatic brain injury*. American Psychiatric Pub.
- [71] Ceusters W, Smith B (2006) A realism-based approach to the evolution of biomedical ontologies. *AMIA Annu Symp Proc* **2006**, 121-125.
- [72] Amdouni E, Morvan Y, Gibaud B (2015) Towards an imaging biomarker ontology based on the open biological and biomedical ontologies foundry. In *Proceedings Semantic Web Applications and Tools for Life Sciences*, pp. 2-3.
- [73] Zhang L, Chang RC, Chu L, Mak HK (2012) Current neuroimaging techniques in Alzheimer's disease and applications in animal models. *Am J Nucl Med Mol Imaging* **2**, 386-404.
- [74] Malhotra A, Younesi E, Gurulingappa H, Hofmann-Apitius M (2013) "HypothesisFinder": A strategy for the detection of speculative statements in scientific text. *PLoS Comput Biol* **9**, e1003117.
- [75] Younesi E, Malhotra A, Gündel M, Scordis P, Kodamullil AT, Page M, Müller B, Springstube S, Wüllner U, Scheller D, Hofmann-Apitius M (2015) PDON: Parkinson's disease ontology for representation and modeling of the Parkinson's disease knowledge domain. *Theor Biol Med Model* **12**, 20.
- [76] Malhotra A, Gündel M, Rajput AM, Mevissen HT, Saiz A, Pastor X, Lozano-Rubi R, Martinez-Lapsicina EH, Zubizarreta I, Mueller B, Kotelnikova E, Toldo L, Hofmann-Apitius M, Villoslada P (2015) Knowledge retrieval from pubmed abstracts and electronic medical records with the multiple sclerosis ontology. *PLoS One* **10**, 1-12.
- [77] Dreses-Werringloer U, Lambert JC, Vingtdoux V, Zhao H, Vais H, Siebert A, Jain A, Koppel J, Rovelet-Lecrux A, Hannequin D, Pasquier F, Galimberti D, Scarpini E, Mann D, Lendon C, Campion D, Amouyel P, Davies P, Fosskett JK, Campagne F, Marambaud P (2008) A polymorphism in CALHM1 influences Ca²⁺ homeostasis, A β levels, and Alzheimer's disease risk. *Cell* **133**, 1149-1161.
- [78] Schupf N, Lee A, Park N, Dang L-H, Pang D, Yale A, Oh DK-T, Krinsky-McHale SJ, Jenkins EC, Luchsinger JA,

- Zigman WB, Silverman W, Tycko B, Kisselev S, Clark L, Lee JH (2015) Candidate genes for Alzheimer's disease are associated with individual differences in plasma levels of beta amyloid peptides in adults with Down syndrome. *Neurobiol Aging* **36**, 2907.e1-2907.e10.
- [79] Deming Y, Black K, Carrell D, Cai Y, Del-Aguila JL, Fernandez MV, Budde J, Ma S, Saef B, Howells B, Bertelsen S, Huang K, Sutphen CL, Tarawneh R, Fagan AM, Holtzman DM, Morris JC, Goate AM, Dougherty JD, Cruchaga C (2016) Chitinase-3-like 1 protein (CHI3L1) locus influences cerebrospinal fluid levels of YKL-40. *BMC Neurol* **16**, 217.
- [80] Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, Iwatsubo T, Jack CR Jr, Kaye J, Montine TJ, Park DC, Reiman EM, Rowe CC, Siemers E, Stern Y, Yaffe K, Carrillo MC, Thies B, Morrison-Bogorad M, Wagster MV, Phelps CH (2012) Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* **7**, 280-292.
- [81] Nelson PT, Braak H, Markesbery WR (2009) Neuropathology and cognitive impairment in Alzheimer disease: A complex but coherent relationship. *J Neuropathol Exp Neurol* **68**, 1-14.