First update on the EADC-ADNI Harmonization of Protocols for Hippocampal Tracing
ICAD, Honolulu, Wednesday, July 14, 2010

Participants:

Giovanni Frisoni - IRCCS S.Giovanni di Dio – Fatebenefratelli, BS, Italy
Marina Boccardi - IRCCS S.Giovanni di Dio – Fatebenefratelli, BS, Italy
Simon Duchesne - Laval University, Québec City, Canada

Josephine Barnes - University College London, Institute of Neurology, Queen Square, London, UK.
Richard Camicioli - University of Alberta Edmonton AB Canada
Bruno Dubois - Neurological Institute of the University Salpetriere Hospital, Paris
Michael Ewers - University of Dublin, Ireland
Perminder Sachdev - University of New South Wales, Sydney, Australia.
Heather Snyder - Alzheimer Association Chicago IL
Andy Simmons – NEUROMED, London, UK
Hilkka Soininen – Kuopio University Hospital, Kuopio, Finland
Pieter Jelle Visser – University of Maastricht, Netherlands
Stephan Teipel – Rostock University, Germany
Lars-Olof Wahlund – Karolinska Institute, Sweden

Remote participants:

Martina Bocchetta - IRCCS S.Giovanni di Dio – Fatebenefratelli, BS, Italy
Timothy Brown - Johns Hopkins University
Haroon Burhanullah - Johns Hopkins University
Leyla deToledo-Morrell – Rush UMC, Chicago, IL
Harald Hampel - University of Frankfurt
John Ratnanather - Johns Hopkins University
Nicolas Robitaille - Laval University, Québec City, Canada

GB Frisoni describes project outline: rationale (slides 4-7), objective (sld 8), working group (sld 9-11), publications plan (sld 12) .

M Boccardi describes briefly the methods already described in Toronto: protocols selection (sld 15-17), protocols certification procedure sld 18), extraction of key landmarks and differences among protocols (sld 19-25), operationalization of the differences among the certified protocols into tracing units(sld 26); modelling of tracing units (sld 26-27), modelling of the normal and AD hippocampi as the sum of all tracing units ("maximum hippo") (sld 28-29), or as they would be obtained using the 12 originally examined protocols (sld 30-41). New results are provided, regarding the re-test reliability (intra-rater) of all tracing units, in a sample of 20 ADNI subjects, selected in order to have all degrees of hippocampal atrophy as evaluated at the MTA scale (Scheltens et al., 1992) (four subjects for each of the 5 severity degrees) (sld 45-46). The segmentation unit "alveus/fimbria" had the lowest reliability (icc=0.863). All of the other segmentation units were associated to particularly high re-test reliability (the lower one had icc=0.964) (sld 48). This may have reflected an overtraining of the tracer (scatterplots of first and second values in slides 49-55), and these results may be of little help for deciding which segmentation units should be included or not in the harmonized protocol. Therefore, new intra-rater values will be computed with a naïf tracer (after her learning session). This tracer will also provide

the tracings for the computation of inter-rater reliability. Moreover, values should not be considered per se, but in a wider context: for example, alveus/fimbria have the lowest re-test reliability, but this problem would be reflected more in the attempt to exclude this segmentation unit rather than in including it (comparison with histological images corresponding to MRI slices (sld 57) and icc values for hippocampal body including alveus and fimbria in a single tracing (sld 58) are provided to support discussion about this point).
Also, preliminary data on the impact of segmentation units on total hippocampal volume and on the difference between patients and controls are provided (slide 47). The tracing unit "subiculum" did not reflect the same amount of atrophy as the other tracing units, and one criterion (oblique line) resulted associated to "hypertrophy" in patients compared to controls. This is a consequence of the fact that the arbitrary lines "vertical" and "oblique" (i.e., same inclination of parahippocampal white matter) form a larger angle in atrophic brains (sld 59). This should also be considered when deciding which criterion should be included in the harmonized protocol, to adopt a criterion that would not overestimate atrophy in AD.
The groups used for these last computations will be expanded (the 20 subjects selected to compute re-test reliability were 8 controls, 9 MCI and 3 AD, i.e., not enough to address the aim of measuring impact of segmentation units on total - normal - hippo volume and on the difference due to AD).

Project outline and part of the results are available at **www.hippocampal-protocol.net**

M Boccardi and GB Frisoni acknowledge the contribution of Alzheimer Association, Lilly International, Wyeth International for supporting the project; Rossana Ganzola and Nicolas Robitaille for the work carried out; all the main authors of the selected protocols for hippo tracing for their active participation in the certification phase.

Questions/comments from the audience:

Q=question
A=answer
C=comment

Automatic algorithms

Q: are you also considering automatic algorithms?
A: no
C: after a harmonized protocol is available, it may be possible to get back to the creators of existing automatic algorithms, proposing to make a version of their algorithms that uses the same landmarks/criteria of the harmonized protocol

Icc values

Q: are you sure that these particularly high icc were properly obtained?
A: this tracer was particularly expert, she had the highest reliability since her first learning phase, from then has always traced hippocampi, and finally with this harmonization work she has further improved her knowledge. It may be that, due to overtraining, this was not the proper person to provide the data that we need to evaluate the differential impact of tracing units on reliability. The whole work will be re-done by a naïf tracer.

C: it may be considered not only volume reliability, but also topographic reliability (volume may be the same, but traced points may not).
C: many ways exist to compute icc values, suggests papers by Chupin (Neuroimage, 2007) and Crum (IEEE, 2006), you may also try those and re-consider these results.

"holes" of liquor

Q: which tracing criterion did you use when liquoral "holes" were found within the hippocampus?
A: we used Jack's criterion of excluding liquoral "holes" only when connected to the liquor external to the hippocampus, and we included those regions as hippocampal tissue when they were not connected to external liquor. Anyway, this issue will need to be considered by the Delphi panel, who will have the responsibility to decide which criterion should be used in these cases.

Normalization

Q: did you normalize your volume values?
A: no, all these are raw values
C: it needs to be decided which kind of normalization (TIV, head/brain..) should be carried out, as well as the whole pre-processing procedure

Image quality

Q: Were images selected based on their quality?
A: Images were all from the ADNI dataset, i.e. ADNI standard parameters. Anyway, only the severity of hippocampal atrophy at the MTA scale was considered in the selection.
Q: has there been any correction for field inhomogeneities aimed to improve image quality or any other pre-processing?
A: no, only alignment along the AC-PC line was carried out.
C: similar image quality may be considered in the selection of subjects, and preprocessing and scaling should also be considered

Pharmacologic treatment

Q: were the ADNI subjects that you used for these tracings pharmacologically treated?
A: probably the (3) AD subjects were, but this variable was not taken into consideration, as only the severity of hippocampal atrophy at the MTA scale was considered in the selection
C: it may be useful to consider treated and non treated subjects
C: based on the new knowledge and criteria, it may be opportune to consider early and late MCI

3T

Q: wouldn't it be useful to replicate this work also on 3T images?
A: this is planned for the validation of the harmonized protocol

Alveus/fimbria

C: you say that the segmentation unit “alveus/fimbria” had lowest reliability, but I think that the exclusion of alveus and fimbria leads to lower reliability.
A: indeed, we provide quantitative info, but it does not speak *per se*, rather, we need to understand the meaning of these numbers. And in fact we provided some evidence that this lowest reliability would actually be reflected in the exclusion, rather than in the inclusion, of this segmentation units [see slides 57-58]. This is tricky, in fact, the lower border of alveus/fimbria is also the upper border of the segmentation unit: “minimum hippocampal body”, which means that the lower reliability originated by this border (which is the real problem) should have been reflected also in a lower reliability of the “minimum hippo body”, which, instead, is associated to high icc. Probably, that is reflected in lower reliability only in the “alveus/fimbria” because this is a much smaller unit, while the “minimum hippo body” (which is the largest segmentation unit – almost 8 times larger than the “alveus/fimbria”) may be more invariant even if traced with this more variable border. The Delphi panel will consider the meaning of the provided value, and will not decide using these values *per s*e.

Validation with pathology

Q: you say that you will validate the harmonized protocol with pathology, but how can you be sure that you get a proper correspondence? And, in case, would you privilege truth or reliability?
A: we have availability of autopsy data and we may benefit of examining it in the validation phase. But, indeed, a further question for the Delphi panel may be whether the objective of this work should be to define a protocol extracting the “real” hippocampus, or a protocol “reliably” extracting it, etc.

Monitor

Q: which kind of monitor was used for the tracing?
A: we used a normal LCD monitor. It is large (24”), but this is not a particular advantage, as magnification of the image also corresponds to reduction of its definition.

One or more protocols?

Q: could this work be extensible to other conditions than AD? Should there be a harmonizad protocol for each disorder?
A (GBF): we aim to a harmonized protocol, but of course in other fields (psychiatric, ..) different protocols, more specific for that field, may be used
A (MB): we are gathering info based on AD, but hopefully the Delphi panel may give precedence to the “hippocampus” rather than to the “AD hippo”. If a harmonized protocol for the hippocampus is achieved, this would guarantee not only sensitivity of measurements in AD, but would also help specificity of measures. This would therefore be an advantage for the AD, and at the same time the extension of its use to other fields may be appropriate and preferable. Again, the Delphi panel may decide on which is the objective of this work, by considering also this aspect.
C: you are providing a bench work that may be used in the other fields to take measures specific to their studies.