Author's response to reviews

Title: A Population-based study of dementia in the oldest old: the Monzino 80-plus Study Design, methodological challenges, and population characteristics

Authors:

Ugo Lucca (lucca@marionegri.it)
Mariateresa Garri (mariateresa.garri@marionegri.it)
Angela Recchia (angela.recchia@marionegri.it)
Giancarlo Logroscino (giancarlo.logroscino@neurol.uniba.it)
Pietro Tiraboschi (ptiraboschi@yahoo.com)
Massimo Franceschi (massimo.franceschi@multimedica.it)
Chiara Bertinotti (chiarabertinotti@libero.it)
Anna Biotti (anna.biotti@libero.it)
Elena Gargantini (elena.gargantini@tin.it)
Marilena Maragna (malinenam.m@libero.it)
Alessandro Nobili (alessandro.nobili@marionegri.it)
Luca Pasina (luca.pasina@marionegri.it)
Carlotta Franchi (carlotta.franchi@marionegri.it)
Emma Riva (emma.riva@marionegri.it)
Mauro Tettamanti (mauro.tettamanti@marionegri.it)

Version: 2 Date: 31 March 2011

Author's response to reviews: see over
To the attention of the
Editor-in-Chief
BMC Neurology

Re: MS: 9468233484732506
A Population-based study of dementia in the oldest old: the Monzino 80-plus
Study design, methodological challenges, and population characteristics

Milano, March 31, 2011

Dear EDITOR,

We have revised our manuscript following the reviewers’ suggestions and hope that in the present form it is now acceptable for publication in BMC Neurology.

All changes made when revising the manuscript have been highlighted in yellow. Below you will find the point-by-point response to the reviewers’ concerns.

We do not have an electronic format of the study questionnaire, the original form of which was created several years ago. Please, let us know an address where we can mail the paper version.

Thank you for your kind attention and consideration.

Sincerely,

Ugo Lucca

Laboratory of Geriatric Neuropsychiatry
Istituto di Ricerche Farmacologiche "Mario Negri"
Via Giuseppe La Masa, 19
20156 Milano - Italy
phone: +39-02-39014427
fax: +39-02-39001916
e-mail: lucca@marionegri.it
Point-by-point reply

Reviewer's report
Title: A Population-based study of dementia in the oldest old: the Monzino 80-plus Study Design, methodological challenges, and population characteristics
Version: 1 Date: 18 January 2011
Reviewer: Michal Schnaider Beeri
Reviewer's report:
This manuscript describes the design and methods of the Monzino 80-plus Study. The oldest old are the fastest growing segment of the Western world population and at the highest risk for developing dementia. Thus learning about their dementia prevalence, incidence, course, and risk factors is a public health imperative. This study documents important biases particularly encountered in the study of the oldest old.

Compulsory revisions:
Although the title refers to this study of dementia, no results on the dementia ratings are presented. This may reflect the present unavailability of the final diagnosis, but even results of initial clinical diagnosis, which presumably will not be dramatically different from the final diagnosis, would enhance the manuscript since this is the main characteristic of interest of this study.

We agree with the reviewer that the word “dementia” in the title could create false expectations, but to investigate dementia in the oldest old was/is the general reason we started the Monzino 80-plus study. Actually the subtitle (as well as the cover letter to the editor) clearly specified the objective of the present report and the kind of results here presented: “Study design, methods, and baseline population characteristics”.

With regard to the estimate of the prevalence of dementia, the manuscript reported that: “... Since interviews of centenarians and those of men 95-99 years old have just finished, data on these individuals are not yet available for analyses ...” (page 11 of the original manuscript [OM]) ”, and that “Diagnoses of dementia made at a certain visit but not confirmed in the following visits were reclassified as “non dementia” at that visit. (page 10 OM)”, while the first follow-up of the group of men 95-99 years old is still ongoing. Thus, we are not yet in the condition to report prevalence data. Besides, reporting the prevalence of dementia here would completely change the aim and the focus of the report and, consequently, quite a different paper should be written. However, following the reviewer’s suggestion we have now added a sentence to the manuscript in the Results section anticipating that:

“Based on the initial diagnoses, the preliminary estimate of over-eighties affected by dementia is rather high accounting for about one third of the present study population (32%, 35% in women and 24% in men), with prevalence increasing with advancing age (16% at 80-84 years, 34% at 85-89, 43% at 90-94, and 57% over 94).”

For interpretation of the MMSE values, it would be very helpful to discuss the differential population norms for the MMSE, based on education, since this sample consisted almost entirely of subjects with little formal education.

As noted by the reviewer, the population investigated in the Monzino 80-plus Study has a low level of formal education, reflecting the general poor schooling of the population
of over-eighties all over Italy. However, we are not sure we have understood the reviewer’s suggestion “to discuss the differential population norms for the MMSE, based on education”. (1) With respect to the Italian population of cognitively normal oldest old? In this case, since the (poor) education of the Monzino population mirrors that of the Italian population of that age class, the comparison would be between the norms for MMSE of the present study normative population (i.e. including demented persons) versus the norms for MMSE in a cognitively normal Italian population of oldest old by age class. (2) With respect to non-Italian normative populations (of course, not to “non-Italian population of normals”) with different educational levels? In this case, without considering item variations and different administration procedures between diverse MMSE versions, very few studies have reported data (and in different ways) useful for comparisons. In any case, we have tried to discuss both these possible comparisons in the “Discussion” section of the revised manuscript.

“Cognitive performance in the oldest old
In the present study population performance on the MMSE steadily decreased with increasing age (Table 5). Using the same standardized version and administration and scoring procedures of the MMSE in a population of 1,680 cognitively normal overeighties from four Italian communities (mean education: 6.2 years), we found that the mean MMSE score declined slowly with increasing age: 26.9 at age 80-84, 25.7 at age 85-89, 25.2 at age 90-94, and 24.4 at age 95 and over (unpublished data). Difference in mean MMSE score between this cognitively normal population and the present study general population (Table 5) progressively increased with increasing age (1.4 at age 80-84, 3.7 at age 85-89, 4.6 at age 90-94, and 7.6 at 95 and over) strongly suggesting that in the general population the proportion of oldest old with cognitive impairment rises with age. Mean MMSE score in the present study population was consistently lower also compared to that of other populations with a higher educational level. Without considering possible item variations and different administration and scoring procedures of MMSE in the diverse studies, the mean MMSE score of the present study population was 22.4 (n = 1,268; 0-8 years of education: 93.1%) versus 24.6 in the 80 years or older individuals of five US community populations (n = 951; 0-8 years of education: 56.9%) (Crum et al. 1993) and 23.9 in the 80 years or older individuals of the Kungsholmen Project (n = 1,106; 0-7 years of education: 26.4%) (Guo et al. 1997). In a sample of 435 adults aged 90 or older enrolled in the 90+ Study, a well-educated, upper middle class cohort, the mean MMSE score was 22.2 (0-12 years of education: 26.7%) (Kahle-Wroblewski et al. 2007) versus 19.9 in the 90 and older individuals of the present study (n = 441, 0-12 years of education: 95.7%). The two Göteborg cohorts had a higher proportion of elderly with a low educational level (0-6 years: 75% in the 85-year-old cohort and 67% in the 95-year-old cohort) and their mean MMSE scores (~23.5 in the 85-year-old cohort and ~18.1 in the 95-year–old cohort) were similar to those of the individuals of the same age in the present study (85-year-old: 23.4; 95-year–old: 18.5) (Skoog 1993 and Borjesson-Hanson 2004). These results suggest that the comparison of prevalence and incidence rates of dementia between populations with quite different educational levels could represent an interesting model to investigate the actual and still controversial role of formal education in cognitive decline and in the development of dementia and Alzheimer’s disease.”

Comparisons of the characteristics of the Monzino study to the few other oldest old
As we wrote in the Introduction of the manuscript ("Even though disease and disability are highly prevalent among the oldest old and represent the background against which cognitive decline occurs, information on the health status of this age segment from large cohorts of the elderly population is lacking"), descriptive data of population-based studies of the oldest old are generally poor and mainly limited to age, sex, marital status, living situation, and mean MMSE score (though not always). Even the educational level is often omitted. Following the reviewers’ suggestion we have contextualized our research within the existing literature and described the other studies in the oldest old.

“Most of the numerous population-based studies on cognitive impairment and dementia have included a broad age span of elderly persons (commonly 65 years or older) with the consequence that the number of oldest old investigated has usually been too small to accurately estimate the prevalence and incidence of dementia in this segment of the elderly population. Even when the number of oldest old included was large enough (O’Connor 1989, Fratiglioni 1991, Bachman 1992, Ebly 1994, Ott 1995, Brayne 1998, Breitner 1999, Von Strauss 1999, Riedel-Heller 2001, Yamada 2001, Lopez-Pousa 2004, Gatz 2005), extreme ages were often poorly represented. To overcome this limitation and more reliably investigate the relationship between age and dementia in the advanced ages, a few studies have focused on the oldest old. All available inhabitants aged 85 years or older (n = 891; 95+ years = 34) were included in the “Leiden 85 plus study” (Heeren 1991). A random sample of 358 citizens aged 85 years and over (90+ years = 91) was assessed in the Munich study (Fichter 1995). Prevalence and severity of dementia was investigated in two birth cohorts of 494 85-year-old and 338 95-year-old persons born in Göteborg (Skoog 1993 and Borjesson-Hanson 2004). The Vantaa 85+ Study included all available persons of at least 85 years of age (n = 553; 90+ years: 105 out of 399 deceased individuals) but only the overall point prevalence of dementia was reported (Polvikoski 2001). Survivors aged 90 and older (n = 911) from the Leisure World Cohort Study, a retirement community in California, were enrolled in the 90+ Study (Corrada 2008). Information on these populations of oldest old was mostly limited to some demographic characteristics while no or very few data on general health status, lifestyle habits, functional disability, and behavioral disturbances have yet been reported.”

Minor essential revisions:
The abstract does not present clearly the questions of the study. Through the conclusions section one can derive that the “important methodological challenges” that are being addressed particularly in an oldest old cohort study are the biases introduced by non-response or by death prior to interview and the crucial contribution of informants.

Following the reviewer’s observation, in the abstract of the revised manuscript we have presented the questions of the study more analytically and better specified them also at the end of the introduction.

Abstract: “The purpose of the study was to describe the design, methods, and baseline
population characteristics of the survey and investigate possible differences in demographic, cognitive, functional, and behavioral characteristics between oldest old with and without any performance on cognitive tests and between oldest old alive and deceased prior to the interview”.

Abstract/conclusions: “… [Use of informant-based instruments in surveys in the oldest old is crucial in assessing everyday functioning and change,] especially in participants with no cognitive test performance available.”

Background (end): “[Here, we describe the study design, methods, and baseline population characteristics of the study and discuss important methodological issues in performing epidemiological studies in the very old], such as possible differences in demographic, cognitive, functional, and behavioral characteristics between oldest old with and without any performance on cognitive tests and between oldest old alive at first visit and those deceased after prevalence day but prior to interview.”

The results of the neuropsychological tests would be very informative since they are affected by extreme age, both with regard to test completion and with regard to test score.

We quite agree with the reviewer, but, precisely because of the interest per se of this topic (a thorough neuropsychological testing in this age group is quite a challenging issue), we have decided to address it separately and specifically. Within the context of the present manuscript, the further addition (and discussion) of other data, moreover concerning a different subject and a sub-population, to the quite large body of results already reported would turn out to be overwhelming and confusing. In this paper we have decided to focus on the main characteristics of the whole study population.

In the methods section, degree of confidence in the preliminary diagnosis of dementia had four categories: dementia, probably dementia, probably no dementia, and no cognitive impairment. It is not clear whether mild cognitive impairment (MCI)/cognitive impairment with no dementia or any of the “transitional” dementia phase definitions are examined and determined. This is crucial if in the future the authors intend to assess predictors of conversion to dementia and most importantly, predictors for intact cognition (which would, by definition, rule out MCIs).

The terms adopted in the present study for rating the degree of confidence with the diagnosis of dementia have nothing to do with any diagnostic criteria for the “transitional” dementia phase between normality and dementia. They aim to reflect both the evaluators’ subjective gradings of the individual’s cognitive functioning but also the completeness and trustworthiness of the information gathered. We definitely agree with the reviewer on the importance of a clear definition of this “transitional” phase “to assess predictors of conversion to dementia and predictors for intact cognition”. We think that the information collected will permit classifying the non demented individuals according to several definitions of this “transitional” phase (such as questionable dementia or CDR 0.5, MCI, CIND, and the kind).

To make the text clearer in the revised manuscript we have specified that:
“Degree of confidence reflects both the evaluators’ subjective grading of the individual’s cognitive functioning and the completeness and trustworthiness of the information gathered for each subject.”
Since the manuscript provides a detailed description of the methods, it is not clear why the specific laboratory tests were not mentioned.

Following the reviewer’s remark we have added the list of laboratory tests performed: “... (complete blood count, glucose, creatinine, total protein, protein electrophoresis, AST, ALT, total cholesterol, HDL cholesterol, triglycerides, sodium, potassium, calcium, vitamin B₁₂, folate, TSH, FT3, FT4, TPHA, and urine test)”.

How did the authors address the differential age recruitment in the different municipalities?

The Monzino 80-plus Study was able to gradually expand the size of the population investigated over time in parallel with the renewals of the funding of the project by the Monzino Foundation. As explained in the manuscript, we started to interview “all registered individuals of 80 years or older residing in Castellanza, Gorla Maggiore, Gorla Minore, Marnate, Olgiate Olona, and Solbiate Olona and of 85 years or older residing in Fagnano Olona” (so as not to inflate the largest class of individuals aged 80-84). Then, in order to increase the number of individuals investigated in the least represented age class (90+), the survey was extended to all registered elderly of 90 years or older residing in the neighbouring and relatively populous (some 50,000 inhabitants) municipality of Gallarate (see Figure 1). Finally, seeking to have a sufficiently large population of elderly also in the extreme ages, the survey tried to reach all the elderly 100 or older and a random sample of men aged 95-99 (very few men survive age 94) who were residents in all the other municipalities of the province of Varese. Having investigated all the over-eighties of the initial survey area, to contrast the natural tapering of the number of subjects alive with increasing age in order to include a sufficiently large number of individuals also in the extreme ages, we obviously had to expand the area of the survey. The study remained confined in any case to the relatively small area of Varese province, historically and culturally very homogenous (moreover considering the age class investigated). Having deliberately expanded the proportion of individuals 90 years or older, of course prevalence and incidence rates of dementia for the whole population will have to be standardized on the general population of the Varese province and Italy.

It would be useful to have some English-language editing, especially in the discussion section.

The revised manuscript has been edited by a second English mother-tongue professional editor.

**Level of interest:** An article of importance in its field  
**Quality of written English:** Needs some language corrections before being published  
**Statistical review:** No, the manuscript does not need to be seen by a statistician.  
**Declaration of competing interests:** I declare that I have no competing interests.
Reviewer's report

Title: A Population-based study of dementia in the oldest old: the Monzino 80-plus Study Design, methodological challenges, and population characteristics

Version: 1 Date: 1 February 2011
Reviewer: Ananya Roy

Reviewer's report:

General comments
This is an important area of investigation and this article has the potential to advance our understanding of cognition among a relatively understudied general population. The study has been well designed with indepth and varied assessment of cognitive and behavioral assessment. However, this article is descriptive and does not address a clearly outlined question.

We think that the title clearly specified the purposes of the article: “... Study design, methodological challenges, and population characteristics”. The two related methodological issues addressed in the manuscript are (see legends of Tables 6 and 7):
1. Are the baseline characteristics of the sub-population of oldest old who did not perform, for whatever reason, any cognitive test different from those of the sub-population of oldest old who instead did?
2. Are the baseline characteristics of the sub-population of oldest old who died IN BETWEEN THE PREVALENCE DAY AND THE DAY OF THE INTERVIEW different from those of the sub-population of oldest old alive at the day of the interview?

The common, underlying question to both of these issues is: if one does not take into consideration the sub-populations of those who do not perform any objective cognitive test and of those who die after the prevalence day but prior to the interview (thus persons who are part of the initial population under study), will the prevalence of cognitive and functional disability in the general elderly population be estimated accurately?

We have now made the methodological questions addressed in the study explicit:

Abstract: “The purpose of the study was to describe the design, methods, and baseline population characteristics of the survey and investigate possible differences in demographic, cognitive, functional, and behavioral characteristics between oldest old with and without any performance on cognitive tests and between oldest old alive and deceased prior to the interview.”

Background (end): “[Here, we describe the study design, methods, and baseline population characteristics of the study and discuss important methodological issues in performing epidemiological studies in the very old], such as possible differences in demographic, cognitive, functional, and behavioral characteristics between oldest old with and without any performance on cognitive tests and between oldest old alive at first visit and those deceased after prevalence day but prior to interview.”

The major limitation of this study is the handeling of the data. Also the statistical analysis needs further refinement to confirm the correlations presented in the paper and later discuss them.

The answers to these general remarks are reported below where the issues are more
analytically raised.

- Major Compulsory Revisions

Methods

1) Study design, para 6:” since dementia has been associated with shorter survival…… at least six months after the last interview”

The reasoning behind estimation of mental status of the participants who died before the interview is not clear. In fact it seems that this may actually bias any assessment of cognitive function among the living elderly, by skewing the distribution.

The original sentence: “Since dementia has been associated with a shorter survival and the probability of dying in the oldest ages is very high, information regarding all aspects of the resident elderly who have died after the prevalence day but before interview was collected from a proxy using the full questionnaire (except, of course, for the cognitive battery) and, for individuals in nursing homes, also from the institution’s records.”

Actually, our aim (see above the complete sentence) was not at all to estimate the ‘mental status of the participants who died before the interview” and to mix up actual scores on MMSE with estimated ones (what for?). Our aim was to compare demographic, cognitive, functional, and behavioral aspects of “the elderly alive at first visit versus those deceased between prevalence day and first visit” (see Table 7). All the information on cognitive competence, activities of daily living, and behavioral disturbances “was collected from a proxy using the full questionnaire (except, of course, for the cognitive battery)”, that is, interviewing in the same way the proxies of both the elderly alive and of those deceased at the time of the interview. The fact that, based on the result of SBI-SI (which is an informant-based instrument to assess cognitive competence in ALL the individuals investigated) we could also report a reasonable and reliable (see later) estimate of the global cognitive performance of the elderly just before their deaths (and after prevalence day!), was/is merely in view of orienting the reader at first sight with respect to the mean level of cognitive functioning of the two groups through a worldwide well known score easy to interpret. The line reporting the MMSE/c-MMSE in Table 7 (the least important) could be skipped without jeopardizing anything in the article. Besides, it should be considered that “cognitive scores from retrospective interviews [taken 0 to 7 years after death] are strongly associated with late-life MMSE scores taken close to death.” (Marioni et al. Int Psychogeriatr 2011; 23: 274-279).

Also given that the overall design of the study is prospective such retrospective assessments would muddy the study design as this is retrospective assessment of mental status after the death has happened.

Please reconsider this.

Though the Monzino 80-plus is a prospective study, data reported in the manuscript are cross-sectional. All the information collected at first visit (as well as at following visits) is - as any anamnestic collection of information in any study as well as in any clinical
setting - retrospective in nature. In this respect, there is no difference between the information collected from the informants of elderly subjects alive and that from the informants of elderly already deceased prior to first visit. The fact that we also reported, among all the other features, an estimate of the global cognitive “performance” of the deceased individuals was only to roughly orientate the reader.

We do not understand why this “would muddy the study design”. Elderly deceased after prevalence day and prior to first visit being excluded from the prosecution of the study, of course can and will not contribute to any prospective results. Moreover, “this is [NOT] retrospective assessment [?] of mental status after the death has happened”, but a reliable estimate of the level of cognitive functioning just before the recent death (Marioni et al. 2011), that is to the time close to death to which all the other scales administered to the informants were referred (i.e. at the cross-sectional point in time under study). Besides, “when used retrospectively, the IQCODE [one of the several scales administered to the informants in the present study] is able to identify cognitive impairment in subjects with neuropathologically defined AD [not “dementia”!].. The accuracy of this instrument is impressive as the test was administered to the informants some years after death (an average delay of 22 months)” (Thomas et al. Int J Geriatr Psychiat 1994; 9: 233-236). And this is true also for other instruments (Ellis et al. Arch Neurol 1998; 55: 360-365). It should also be considered that all of the informant-based scales (not only in the Monzino Study) investigate the behavior of any participant with reference not to the day of the (informant’s) interview, but to a time period (one month for SBI, ten years for IQCODE, etc.) before the interview.

Statement added to the Discussion:
“Retrospective interviews (on average two years after death) were shown to be sensitive instruments to detect the antemortem presence of dementia and were strongly associated with late-life MMSE scores taken close to death [Thomas 1994, Ellis 1998, Marioni 2011].”

2) Study design, para 6: “ modifications which arose….. were not taken into consideration”

This is not clear. If it means that cognitive changes associated with entering the last stages of life, ailing health etc, were not taken into consideration, then this has the pit fall of leading to immense recall bias and misclassification of cognitive and behavioral function, as it would be hard to determine which aspects of behavior and mental ability were due to aging and those due to illness preceeding death. This ...

The original sentence: “All aspects investigated refer to the time preceding death; modifications which arose as a consequence of a possible critical phase preceding death were not taken into consideration. In order to minimize attrition bias, the same was done at follow-ups for participants who had died between two following visits but at least six months after the last interview.”

We wrote “critical phase” not “last stages of life” or “ailing health”. By definition, a critical phase is an unfavourable clinical condition of uncertain prognosis quoad vitam. Had we not also taken the post stroke phase into consideration, we could have simply written “agonal phase”. Thus, a dramatic clinical condition when the patient is commonly unconscious or in a confusional state (delirium) and, as according to any
criteria, it is not possible to make a new diagnosis of dementia. It is making a (mis)diagnosis of dementia under this condition that would lead to an “immense bias”. If a person died because of, let’s say, a heart attack, a stroke, an accident, or, in general, with no critical condition preceding death, “the time preceding death” considered to investigate “all aspects” would obviously begin from the last day of life (as for the living individuals).

… This could either bias any association between cognitive function and risk of death away from the null or over correct and bias it strongly towards the null, which ever direction, it would be impossible to determine the true relationship between cognitive ability and risk of death. Instead limiting the assessment to those who are alive, would give a robust estimate of risk, with less chance of missclassification.

We did not investigate any of these associations in our manuscript and if we do, we will surely not take into consideration individuals already dead to investigate the “association between cognitive function and risk of death”.

3) Methods, Measurements, Cognitive performance and competence,

It needs to be clear where proxy measures were taken and where the subjects themselves were assessed/ or answered questions. These would need o be treated differently, as the patterns of recall bias would be different among the two groups. Also information gathered from GP’s or from institution documents can not be lumped with that gathered from family members, as assessment of family members is purely dependent on their experience with only their loved one, where frequency reported by institutions or GP’s is tempered by experience of dealing with large numbers of patients. This can introduce systematic bias, as those with higher levels of disability would be institutionalized.

As in any door-to-door population-based study, subjects were interviewed where they were living the day of the interview: at home and in an institution (if still capable of answering). As reported in the original manuscript (p. 12 and Table 2), informant-based instruments were administered mainly to family members (in about 77% of the cases) or to relatives (in about 20% of the cases) at home if the subject was at home, in an institution or at the informant’s home at their convenience (mostly at home) if the subject was in an institution. In this respect, we do not see what makes this study different from other population-based studies on dementia.

These would need o be treated differently, as the patterns of recall bias would be different among the two groups. … This can introduce systematic bias, as those with higher levels of disability would be institutionalized.

We are not sure we have understood to which two groups (“where PROXY measures were taken” vs “where the SUBJECTS themselves were assessed/ or answered questions”?) and thus to what kind of systematic bias the reviewer is making reference. If the two groups are that of subjects at home and that of subjects in an institution (why, in fact, would “the pattern of recall bias ... be different” if the informant of a subject in an institution answered at home or in institution? In any case, the large majority of informants answered at home.), we agree with the reviewer that “those with higher levels of disability would be institutionalized”, but do not see why this would lead to a
systematic bias. The more the subjects are cognitively deteriorated the less reliable the information gathered from them would be (wherever they live), however, this is precisely the reason why, in studying dementia, information should be collected also by collateral sources. Moreover and on the other hand, the more cognitively deteriorated a subject is the easier the diagnosis of dementia is (and frequently they already have it). Thus, we do not see where any bias could be. Going back to the aim of the present study, it is exactly because of this common judgement (subjects in institutions are more likely deteriorated than those at home) that inspecting the percentages of “living in institution” in Tables 6 and 7 one realizes that not taking into consideration subjects without any cognitive assessment and subjects who died prior to interview would tend to underestimate the prevalence of cognitive and functional disability in the elderly general population.

Also information gathered from GP’s or from institution documents can not be lumped with that gathered from family members, as assessment of family members is purely dependent on their experience with only their loved one, where frequency reported by institutions or GP’s is tempered by experience of dealing with large numbers of patients.

Apart from the evidence showing that a high level of accuracy in the diagnosis of even very mild dementia (histologically confirmed in AD cases) can be achieved with information provided by an informant (“usually the spouse or an adult child”) (Morris et al. Neurology 1991; 41: 469-478; Cacchione PZ et al. J Am Geriatr Soc 2003; 51: 819-823) while “only around one third of people with some form of dementia are clinically diagnosed as such [by GPs]” (BMJ 2010; 341: 865-867), the study evaluators did/do not lump information from the different sources, but, as any clinician would do in clinical practice, carefully reviewed/review all the information collected.

Statement added to the Methods (Measurements):
“When used retrospectively (on average 22 months after death), the IQCODE was a valid instrument to identify cognitive impairment in subjects with neuropathologically defined Alzheimer’s disease [Thomas LD 1994].”

Statement added to the Discussion:
“Collateral sources, usually the spouse or an adult child, were shown to be accurate in reporting the cognitive abilities of subjects even in a very mild stage of dementia [Morris 1991, Cacchione 2003].”

4) Methods, statistical analysis

The data collected in this study has the depth and breadth to illuminate a large number of questions. Statistical analysis is usually guided by questions and hypotheses. The fact that there is no clearly defined question/s causes confusion in understanding the results and discussion.

The clarification added in the revised manuscript on the aims of the study should help understand the results and discussion.

From the results it seems that emphasis is to provide:
a) a description of the population in terms of demographic information, medical history and lifestyle, and socioeconomic status.
b) Describe the neurocognitive and neurobehavioral profile of the population
c) explore the relationships between neurocognitive and neurobehavioral status and age
d) Compare the profile of deceased and alive participants as well as those with MMSE vs. those with the c-MMSE

If:
point a) refers to Tables 2, 3, and 4;
point b) refers to Table 5;
point c) (“explore the relationships between neurocognitive and neurobehavioral status [?] and age [?]”) should probably refer to the correlation coefficients of assessed and estimated cognitive performance scores with cognitive competence ratings reported on page 13 under “Cognitive, Functional, Mood, and Behavioral Assessments” (last three lines);
point d) refers to Tables 6 and 7.

It would be best to clearly outline these and then describe the statistical methods used to address each.

Following the reviewer’s suggestion, in the “Statistical Analysis” section we have now first outlined the issue at stake and then described the statistical method used to analyse it.

“To describe the baseline status of the study sample we tabulated the percentages (for nominal variables) and means and standard deviations (for numerical variables) of socioeconomic characteristics, lifestyle habits, activities and clinical features of the population. Cognitive, functional, mood and behavioral profiles were described using both means and standard deviations and medians and upper and lower quartiles so as to give a comprehensive view of the distribution of the variables, particularly for the ones that were not normally distributed. All characteristics were tabulated in the overall sample and by quinquennia. Trend across (ordered) age groups was studied using a test developed by Cuzick [40]. Characteristics derived from interview of the proxy informants were compared between subjects alive at the time of the study visit and subjects deceased after prevalence day but prior to the study visit by means of chi-squared (for nominal variables) or Student’s t-test (for numerical variables).

The correlation between cognitive performance and daily life cognitive competence was calculated by means of Pearson’s r linear coefficient.

Elderly subjects with and without MMSE were compared on a number of characteristics retrieved from the interviews with proxy informants by means of chi-squared (for nominal variables) or Student’s t-test (for numerical variables). Comparisons were repeated to adjust for age, sex and education using ANOVA.

Student’s t-test and Pearson’s linear correlation coefficient results were cross-checked with the non-parametric Mann-Whitney test and Spearman’s rank correlation coefficient, respectively, finding very similar results (and no change from significance to non-significance or vice versa).

Analyses were done using JMP v9.0.1 (Sas Institute Inc., Cary, NC) and Stata/IC v11.1 (StataCorp, College Station, TX).”
The statistical analysis presented here is appropriate for simple descriptive purposes (a and b). It would be important to explore the distribution of the cognitive and behavioral assessments (normal/ skewed). It seems from the comparison of the mean and the medians presented that they are skewed.

Of course we explored the distribution of the cognitive, functional and behavioral scores, in fact we reported both means and medians as well as related measures of dispersion (SDs and 25th and 75th percentiles) so as to give the reader a good idea of the shapes of the distributions and let the reader compare the cognitive, functional and behavioral characteristics of the present population with those of others, as most often only means and SDs are reported. The reviewer is right in pointing out the presence of asymmetry in the distribution of the measures and we have acknowledged this in the revised statistical analysis section.

This means that later utilization of the Pearson correlation coefficient is not appropriate, spearman rank correlation coefficient should be used.

We think that using the Pearson correlation coefficient gives a more "naturalistic" view of the concordance between the two measures (directly measuring the "linear strength") than the use of a non-parametric test that measures something more "artificial" (measuring the "linear strength" between the ranks). We agree with the reviewer that in principle this measure should not be used for non-normal distributions, but we used it notwithstanding this problem after inspecting the bivariate scatterplot and after looking at the estimates of its non-parametric counterpart Spearman’s rank ordered correlation coefficient, which are very high and very similar to those of the parametric coefficient: MMSE/IQCODE: -0.76, MMSE/SBI-SI: -0.81; c-MMSE/IQCODE: -0.83, c-MMSE/SBI-SI: -0.88. By the way, even though almost no measure in the neuropsychological domain is normally distributed, regression analysis/analysis of variance are methods always used (as far as a numerical variable is concerned), as presently they are the only methods available to correct for confounding and as such they are also indicated by the reviewer (see next point).

The later analyses (c &d) (correlation coefficients and bivariate ANOVA or Chi square tests) are only correlational and susceptible to bias due to confounding, by socio economic status, social activity, gender, education, clinical conditions etc. It would be important to carry out regression analysis, with appropriate selection of confounders and present these in the results.

Here there must have been some misinterpretation.
Point c, the correlation coefficients: what would be the purpose, at least within the context of the present report, of adjusting “with appropriate selection of confounders” the correlations between scores calculated in the SAME population? Confounders of what?
Point d, Tables 6 and 7: if the purpose of these two tables was to compare the “Baseline characteristics” “of living individuals with MMSE or with c-MMSE score” (Table 6 title) and “of elderly alive at first visit versus those deceased between prevalence day and first visit” (Table 7 title) in order to investigate if the profiles of these two pairs of subpopulations were different, why should those figures be adjusted “with appropriate selection of confounders” (of what ?)? We were not interested in knowing if “socio
economic status, social activity, gender, education, or clinical conditions” can partially
explain why individuals without MMSE are, for example, more likely institutionalized
or have a score on SBI-SI almost twice that of individuals with. What matters here is
that, if, for whatever reason, the percentage of individuals institutionalized is much
higher or the score on SBI-SI, IQCODE, IADL, bADL is consistently worse in one
group than in another, then the probability of observing a higher frequency of cognitive
and functional disability in the first group is awfully high. As a result, if one does not
take this evidence into account the prevalence of dementia in the general population
will be underestimated. Actually, we would have expected the opposite question: why,
in Table 6, did you correct p-values for age, sex and education? In fact, it was not
necessary. The finding of an increasing percentage of individuals without MMSE with
age (upper part of Table 6) aroused our suspicions that a higher proportion of
cognitively impaired elderly in this group was to be expected (considering the positive
association of cognitive decline with age). To see how much the other differences could
be associated with age, we adjusted the p-values for demographic variables. But even if
the adjusted p-values were no longer significant, those differences would still be there
and would continue to suggest that a much higher proportion of cognitively impaired
individuals is likely to be found in the group without MMSE.

5) Discussion

a) The discussion of the results need to be contextualized with existing literature.
For example, How do the scores of cognitive ability in this population compare to those
in other studies/populations?

A similar question was also raised by the other reviewer and we have now tried to
contextualize our research within the existing literature and compare ours with other
study results in the Discussion section.

b) Limitations of the study need to be elaborated.

Though not under the caption “Limitations”, the main limitation of the study was
already reported on page 14: “the common indisposition of the elderly affected by
dementia to carry out a brain imaging study to assist in the differential diagnosis of
dementia type”.

c)“ Failing to consider the deceased or untested results in bias”
The arguments outlined in this has problems.

It is to be expected that people with and without MMSE scores would have different
levels of disability, as that is why the subjects were unable to carry out the MMSE to
begin with. This limits the population that can be studied using the MMSE, but does not
bias any association, just limits the generalizability of the study.

Actually, subjects who cannot perform MMSE (a very easy global cognitive test)
because of physical disability are very few, while the opposite is common: subjects
cognitively deteriorated become disabled because of cognitive impairment. But
whatever the causal direction between cognitive impairment and disability is, we do not
see why “the arguments ... has problems”. As already explained, for whatever reason
“Subjects alive and without MMSE are ... older, more often institutionalized, and with a much worse cognitive, functional, and behavioral profile than those with”, the failure to include these subjects would tend to underestimate the prevalence of cognitive and functional disability in the general elderly population. “Hence the importance of also measuring observable behaviors in the everyday environment to gather more complete and reliable information in a field survey in the very old.” In the present study we did not investigate any association.

This is not the case with the observed behaviors and the informant survey information and thus is an advantage of those tools of assessment as has been pointed out by the authors. These arguments should also be supported with by appropriate references.

To our knowledge, the present is the first study to directly investigate these important methodological issues, especially the different characteristics of oldest old with and without cognitive test performance. Relations between dementia, cognitive impairment, and mortality have been explored, in different ways, in many other studies, several (eight, now nine) of which were already reported in the original manuscript (p. 15).

“Subjects alive and without MMSE.....behavioral profile than those with.”

Without correcting for medical conditions any association between institutionalization, age and mental status would be confounded. As people with severe medical disability have a higher probability of dementia and they have a higher probability of being institutionalized. Thus without correcting for medical conditions no association between institutionalization and mental status can be parsed out.

“As people with severe medical disability have a higher probability of dementia”: is any “severe [how severe?] medical disability” associated with “a higher probability of dementia”? This is something that should be demonstrated. Even the association with specific pathologies is for the most part still controversial.

In any case, we did not study “any association (between institutionalization, age and mental status” or “between institutionalization and mental status)” and adjusting for whatever “confounder” (of what?) will not reverse the situation: those without MMSE will remain older, more institutionalized, less cognitively competent, more functionally disabled, more behaviorally disturbed.

- Minor Essential Revisions

1) Methods

The study design of this particular article is cross-sectional with collection of covariate information and cognitive assessment being carried out at one time point. It is not clear whether the study is to be carried forward in time with further follow up visits making it prospective. Use of the term “prospective” causes confusion.

Yes, results here presented are cross-sectional, while the Monzino 80-Plus Study is a prospective study as correctly described by the reviewer in a previous point (“the over all design of the study is prospective”) and as suggested by sentences like: “... the same was done at follow-ups for participants ...” [page 5], “The same schedule was repeated at each follow-up” [page 6], “... agreement between interviewers was evaluated at one of the study follow-ups” [page 9], “Diagnoses of dementia made at a certain visit but
not confirmed in the following visits ...” [page 10], “a great effort was made at follow-up ...” [page 15]. In the manuscript the term “prospective” has been used on pages 2, 3 and 15 as a general descriptor of the Study in close relation with the phrase “The Monzino 80-plus Study is...” and never in association with present study results (all seven Tables reported the term “baseline”).

2) Results
Given that the title of the article is “A population based study of dementia in the oldest old” it would be important to report the prevalence to diagnosed dementia in this population. Also, the methods describe the procedure for diagnosis, so the reader does look for this in the article. Also then this would need to be discussed later.

Please see the answer to an analogous question asked by the other reviewer.

- Discretionary Revisions

Methods, Measurements, Cognitive performance and competence,
Para 3:
“MMSE showed a very high correlation coefficient with both SBI-SI....[0.87 X SBI-SI]

How much of the variation in MMSE is explained by the substitute measures?

The percentage of variability of MMSE explained by BIMC was 81% and that by SBI-SI 72%.

A strong correlation may not be sufficient to estimate another measure as different aspects are measured by the different tests. The BIMC does not measure any function dependent on visual or manual ability and thus would differ from the whole MMSE. It would be interesting to see how accurately the formula provided here tracks the assessed MMSE in this sub-population that may have all the measures (the BIMC, SBI-SI and MMSE), providing an empirical assessment of this method.

BIMC (score range 0-33) and MMSE (score range 0-30) are two global cognitive function tests sharing 6 identical questions, that is some 20% of the score. We investigated the correlation between MMSE and BIMC or SBI-SI because we hypothesized a common underlying trait, global cognition. We agree with the reviewer that “an empirical assessment of this method” is “to see how accurately the formula provided here tracks the assessed MMSE in this sub-population that may have all the measures (the BIMC, SBI-SI and MMSE)” and thus followed her suggestion. In the 1,268 individuals with both baseline MMSE and SBI-SI score available, the mean MMSE score was 22.44 (SD 7.37) and the mean estimated c-MMSE score was 22.38 (SD 7.22). In the 161 individuals with both MMSE and BIMC scores available, the mean MMSE score was 17.81 (SD 7.10) and the mean estimated c-MMSE score was 17.99 (SD 6.07). We have now reported these results in the revised version of the manuscript.

“The accuracy of the formulas provided by regression analyses to estimate a c-MMSE
score in individuals who did not perform the test was empirically tested in the sub-population of oldest old for whom the actual scores of MMSE and BIMC or of MMSE and SBI-SI were available. In the 1,268 individuals with both baseline MMSE and SBI-SI scores available, the mean MMSE score was 22.44 (SD 7.37) and the mean estimated c-MMSE score was 22.38 (SD 7.22). In the 161 individuals with both MMSE and BIMC scores available, the mean MMSE score was 17.81 (SD 7.10) and the mean estimated c-MMSE score was 17.99 (SD 6.07).”

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Needs some language corrections before being published

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:** I declare that I have no competing interests