SUPPLEMENTARY MATERIAL

**Appendix A.1: PRISMA checklist**

|  |  |  |  |
| --- | --- | --- | --- |
| Table A.1 : Prisma checklist and associated page | | |  |
| **Section/topic** | **#** | **Checklist item** | **Reported on page #** |
| **TITLE** | | |  |
| Title | 1 | Identify the report as a systematic review, meta-analysis, or both. | TITLE |
| **ABSTRACT** | | |  |
| Structured summary | 2 | Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number. | ABSTRACT |
| **INTRODUCTION** | | |  |
| Rationale | 3 | Describe the rationale for the review in the context of what is already known. | 3 |
| Objectives | 4 | Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS). | 3 |
| **METHODS** | | |  |
| Protocol and registration | 5 | Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number. | 3 |
| Eligibility criteria | 6 | Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale. | 4 |
| Information sources | 7 | Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched. | 4 |
| Search | 8 | Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated. | Appendix 2 |
| Study selection | 9 | State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis). | 5/6 |
| Data collection process | 10 | Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators. | NA |
| Data items | 11 | List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made. | NA |
| Risk of bias in individual studies | 12 | Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis. | 6/7 |
| Summary measures | 13 | State the principal summary measures (e.g., risk ratio, difference in means). | 6 |
| Synthesis of results | 14 | Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I2) for each meta-analysis. | N/A |

**Appendix A.2: details on search procedure**

1. Web of Science

In “Advanced research” :

TS=((Frail\* OR pre-frail OR sarcopenia) AND (("long term care" OR "home care" OR "home help" OR "home services" OR "community services" OR formal care OR home nurse OR care worker OR informal care OR nursing home OR institution) AND (use OR utilization OR visit OR consumption)))

Additional configuration:

* Restrict results to English (language) and document types (articles).
* Custom year range to 2000-2018.

1. PubMed/MedLine

In “Advanced”:

(((Frail\*[Title/Abstract] OR pre-frail[Title/Abstract] OR sarcopenia[Title/Abstract])) AND ("long term care"[Title/Abstract] OR "home care"[Title/Abstract] OR "home help"[Title/Abstract] OR "home services"[Title/Abstract] OR "community services"[Title/Abstract] OR “formal care”[Title/Abstract] OR “home nurse”[Title/Abstract] OR “care worker”[Title/Abstract] OR “informal care”[Title/Abstract] OR “nursing home”[Title/Abstract] OR “institution”[Title/Abstract])) AND (use[Title/Abstract] OR utilization[Title/Abstract] OR visit[Title/Abstract] OR consumption[Title/Abstract])

Additional configuration:

* Custom publication dates: 2000-2018

1. Embase

In “Advanced”:

(frail\*:ti,ab,kw OR 'pre frail':ti,ab,kw OR sarcopenia:ti,ab,kw) AND ('long term care':ti,ab,kw OR 'home care':ti,ab,kw OR 'home help':ti,ab,kw OR 'home services':ti,ab,kw OR 'community services':ti,ab,kw OR 'formal care':ti,ab,kw OR 'home nurse':ti,ab,kw OR 'care worker':ti,ab,kw OR 'informal care':ti,ab,kw OR 'nursing home':ti,ab,kw OR 'institution':ti,ab,kw) AND (use:ti,ab,kw OR utilization:ti,ab,kw OR visit:ti,ab,kw OR consumption:ti,ab,kw)

Additional configuration:

* Mapping: no options
* Custom publication dates: 2000-2018
* Sources: Embase, Pudmet-not-Medline
* Quick limits: only in English
* Publication types: articles

1. CINAHL

(Frail\* OR pre-frail OR sarcopenia) AND (("long term care" OR "home care" OR "home help" OR "home services" OR "community services" OR “formal care” OR “home nurse” OR “care worker” OR “informal care” OR “nursing home” OR “institution”) AND (use OR utilization OR visit OR consumption))

Additional configuration:

* Databases: Academic Search Premier / EconLit
* Publication dates: 2000-2018
* Document type: article
* Language : English
* Select “Academic journals (peer rewied) ”

**Appendix A.3: Quality assessment of each study according to the Berger et al (2014) grid**

To compare the quality of the evidence provided by the selected papers, we use the questionnaire proposed by the ISPOR Task Force to assess the relevance and credibility of observational studies in health care fields [33]. This task force provided a questionnaire to help decisions makers to evaluate observational health studies in a systematic manner.

The questionnaire is divided in two parts: the first relates to relevance while the second is focused on credibility. The relevance part of the questionnaire is centred on the relevance of the population (Is the population relevant?), the interventions and outcomes (Are any relevant intervention missing? Are the outcomes relevant?). It additionally considers the context of the study (Is the context applicable?). Due to the systematic review procedure, these criteria are systematically fulfilled. The second part focuses on credibility, referring to the extent to which the paper answers the research question it considers. Credibility considers the design of the paper (8 questions), the characteristics of the data used (4 questions), the way the analysis were conducted (3 questions), the reporting (7 questions) and the interpretation (4). The questions of each category are respectively reported in Table A.3.1 to A.3.5. When a question is answered positively (coded 1 in the Table), the paper checks quality requirements with respect to the dimension considered. Conversely, if the answer is negative (coded 0), it means that the paper meets some limitations with respect to the dimension considered. Tables A.3.1 to A.3.5 present the way each paper was assessed with respect to each item in the four categories.

An additional category in Berger et al.’s grid is related to conflict of interests: none of our studies declared conflicts of interest.

We provide here some precisions on the way some questions have been interpreted in the quality assessment. Question *“Was there evidence that a formal study protocol including an analysis plan was specified before executing the study?”* (Design, N°3) was coded 1 if there was evidence of a protocol (mention of the protocol, or ethics approval, or sample size computation). Question *“Was a study design used to minimize or account for confounding?”* was coded 1 if the methodology explicitly makes it possible to neutralize confounding (panel analysis, instrumental variable). Question *“Was there a thorough assessment of potential measured and unmeasured confounders?”* (Analysis, N°1) was coded 1 if the analysis controls for key confounders. Question *“Did the authors describe the statistical uncertainty of their findings?”(*Reporting, N°4) was coded 1 if confidence interval or standard errors were reported. Question *“Was the effect of unmeasured confounding discussed?*” was coded 1 if there is a discussion of potential bias in the study.

Three questions are regarded as having no discriminatory power, either because the point they mention is systematically not reported in the studies (“*Were sample size and statistical power to detect differences addressed*?” (Design, n°4); “Was the follow-up period of sufficient duration to detect differences addressed?” (Design, n°6) or because they only apply to longitudinal studies (“*Was the follow-up time similar among comparison groups or were the differences in follow-up accounted for in the analyses?”* (Data, N°4)). Consequently, they are not taken into account in the construction of aggregated score for the quality analysis.

For two questions, a negative answer is associated to what is called a “fatal flaw”, meaning that the limitation considered leads to strong validity issues. In the credibility part of the questionnaire, there are two potential fatal flows: the first corresponds to the situation when the exposure outcome is not valid. In the systematic review procedure, papers have been selected to have a relevant exposure outcome: then the first fatal flow is not an issue. The second criteria leading to a fatal flow relates to the assessment and control of confounding.

In the ‘Design” part (Table A.3.1), the main issue in this category refers to the inclusion of a formal protocol (question 3, 8 studies coded 1 out of 17 studies) and the use of a design to minimize or account for confounding (question 5, 2 studies over 17). With respect to data (Table A.3..2), all criteria are fulfilled by each study (except question 1, 16/17 studies). Regarding analysis (Table A.3.3), it is rare to have analyses of subgroups or interaction effect (question 2) and sensitivity analysis (question 3): these criteria are met by 1 study out of 17. The main issues with reporting lie in the reporting of adjusted estimates (question 4, 11/17), the statistical uncertainty (question 5, 13/17) and the reporting of both absolute and relative measure of treatment (frailty and sarcopenia). Finally, most studies fail to address the discussion of unmeasured confounders in the interpretation part (Table A.3.5, Question 4, 5/17) while other criteria are met by all the studies.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table A.3.1 : assessment of credibility with respect to the *study design* domain | | | | | | | | | | | | | | | | | | |  |
|  | Question | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |  |
| Rothman et al. (2008) | Rochat et al. (2010) | Gobbens et al. (2012) | Hoeck et al. (2012) | Hogan et al. (2012) | Schoufour et al. (2014) | Hirani et al. (2015) | Bentur et al. (2016) | Bock et al. (2016) | Butler et al. (2016) | Campitelli et al. (2016) | Kelaiditi et al. (2016) | Gonzalez-Colaço et al. (2017) | Ferrante et al. (2018) | Hajek et al. (2018) | Lambotte et al. (2018) | Verve et al. (2018) | Number of studies coded 1 |
|  | Reference number | [15] | [16] | [17] | [18] | [27] | [29] | [20] | [19] | [21] | [30] | [28] | [31] | [22] | [23] | [24] | [25] | [26] |  |
|  | **Design** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Subscale score (0 to 6)a** | **5** | **4** | **4** | **4** | **4** | **5** | **4** | **5** | **5** | **5** | **5** | **5** | **4** | **4** | **5** | **5** | **5** | **Average score : 4,6** |
| 1 | Were the study hypotheses or goals prespecified a priori? | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 17/17 |
| 2 | If one or more comparison groups were used, were they concurrent comparators or did they justify the use of historical comparison group(s)? | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 17/17 |
| 3 | Was there evidence that a formal study protocol including an analysis plan was specified before executing the study? | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 8/17 |
| 4 | Were sample size and statistical power to detect differences addressed? | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| 5 | Was a study design used to minimize or account for confounding? | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 2/17 |
| 6 | Was the follow-up period of sufficient duration to detect differences addressed? | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |  |
| 7 | Were the sources, criteria, and methods for selecting participants appropriate to address the study questions/hypotheses? | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 17/17 |
| 8 | Were the study groups selected so that comparison groups would be sufficiently similar to each other (e.g., either by rescriction or recruitment based on the same indications for treatment?) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 17/17 |
| Notes: questions from the grid assessing the credibility of observational studies (Berger et al. (2014)).  0 = quality item not fulfilled in the paper. 1 = quality item fulfilled in the paper. “NA” stands for “not available”.  a : 8 items are included in the Design section but two of them are systematically not available in our studies. | | | | | | | | | | | | | | | | | | |  |

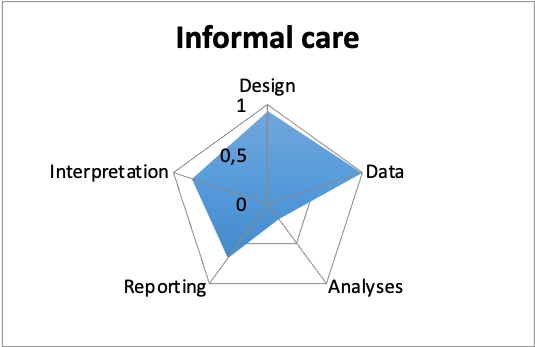
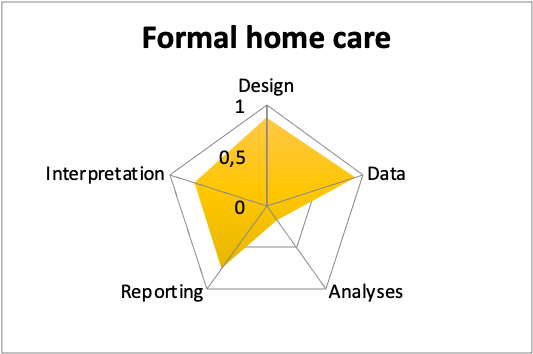
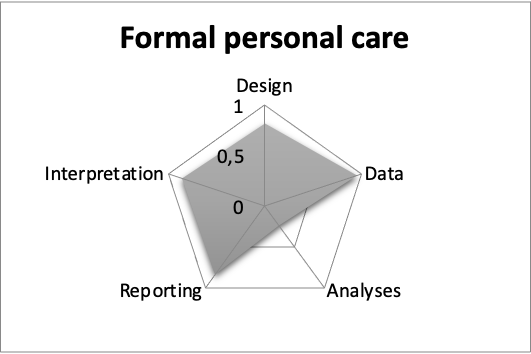
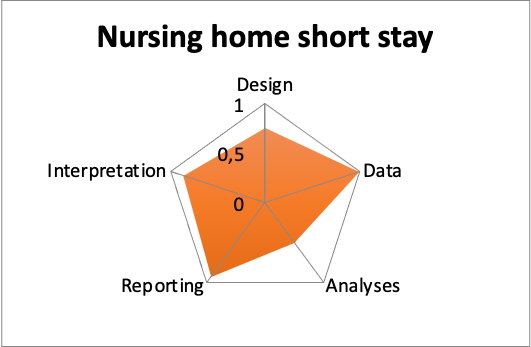
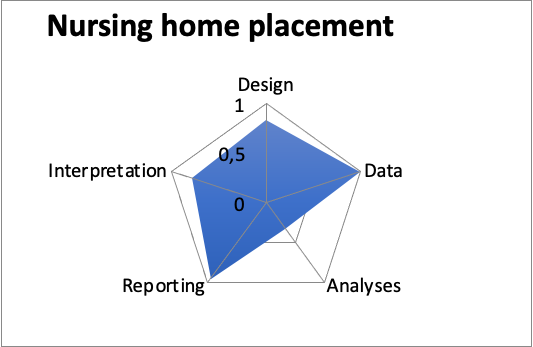
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table A.3..2 : assessment of credibility with respect to the *data* domain | | | | | | | | | | | | | | | | | | |  |
|  | Question | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |  |
| Rothman et al. (2008) | Rochat et al. (2010) | Gobbens et al. (2012) | Hoeck et al. (2012) | Hogan et al. (2012) | Schoufour et al. (2014) | Hirani et al. (2015) | Bentur et al. (2016) | Bock et al. (2016) | Butler et al. (2016) | Campitelli et al. (2016) | Kelaiditi et al. (2016) | Gonzalez-Colaço et al. (2017) | Ferrante et al. (2018) | Hajek et al. (2018) | Lambotte et al. (2018) | Verve et al. (2018) | Number of studies coded 1 |
|  | Reference number | [15] | [16] | [17] | [18] | [27] | [29] | [20] | [19] | [21] | [30] | [28] | [31] | [22] | [23] | [24] | [25] | [26] |  |
|  | **Data** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Subscale score (0 to 3/4)a** | **3** | **3** | **3** | **3** | **3** | **3** | **3** | **2** | **3** | **3** | **3** | **3** | **3** | **3** | **3** | **3** | **3** | **Average score : 3/3 if Q3 = NA**  **2,9/4 if Q4 = 1** |
| 1 | Were the data sources sufficient to support the study? | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 16 |
| 2 | Was exposure defined and measured in a valid way? | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 17 |
| 3 | Were the primary outcomes defined and measured in a valid way? | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 17 |
| 4 | Was the follow-up time similar among comparison groups or were the differences in follow-up accounted for in the analyses? | 1 | NA | 1 | NA | NA | NA | NA | 1 | 1 | NA | 1 | 1 | 1 | 1 | 1 | NA | NA | 9/9 |
| Notes: questions from the grid assessing the credibility of observational studies (Berger et al. (2014)).  0 = quality item not fulfilled in the paper. 1 = quality item fulfilled in the paper. “NA” stands for “not available”.  a : question 4 in relevant only for longitudinal studies. Then, the total subscore either have a maximum of 3 (cross-sectional studies) or 4 (longitudinal studies). | | | | | | | | | | | | | | | | | | |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table A.3.3 : assessment of credibility with respect to the *analyses* domain | | | | | | | | | | | | | | | | | | |  |
|  | Question | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |  |
| Rothman et al. (2008) | Rochat et al. (2010) | Gobbens et al. (2012) | Hoeck et al. (2012) | Hogan et al. (2012) | Schoufour et al. (2014) | Hirani et al. (2015) | Bentur et al. (2016) | Bock et al. (2016) | Butler et al. (2016) | Campitelli et al. (2016) | Kelaiditi et al. (2016) | Gonzalez-Colaço et al. (2017) | Ferrante et al. (2018) | Hajek et al. (2018) | Lambotte et al. (2018) | Verve et al. (2018) | Number of studies coded 1 |
|  | Reference number | [15] | [16] | [17] | [18] | [27] | [29] | [20] | [19] | [21] | [30] | [28] | [31] | [22] | [23] | [24] | [25] | [26] |  |
|  | **Analyses** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Subscale score (0 to 3)** | **1** | **2** | **1** | **1** | **1** | **0** | **1** | **0** | **2** | **1** | **1** | **1** | **1** | **1** | **1** | **0** | **0** | **Average score : 0,9** |
| 1 | Was there a thorough assessment of potential measured and unmeasured confounders? | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 13 |
| 2 | Were analyses of subgroups or interaction effects reported for comparison groups? | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 3 | Were sensitivity analyses performed to assess the effect of key assumptions or definitions on outcomes? | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| Notes: questions from the grid assessing the credibility of observational studies (Berger et al. (2014)).  0 = quality item not fulfilled in the paper. 1 = quality item fulfilled in the paper. | | | | | | | | | | | | | | | | | | |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table A.3.4 : assessment of credibility with respect to the *reporting* domain | | | | | | | | | | | | | | | | | | |  |
|  | Question | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |  |
| Rothman et al. (2008) | Rochat et al. (2010) | Gobbens et al. (2012) | Hoeck et al. (2012) | Hogan et al. (2012) | Schoufour et al. (2014) | Hirani et al. (2015) | Bentur et al. (2016) | Bock et al. (2016) | Butler et al. (2016) | Campitelli et al. (2016) | Kelaiditi et al. (2016) | Gonzalez-Colaço et al. (2017) | Ferrante et al. (2018) | Hajek et al. (2018) | Lambotte et al. (2018) | Verve et al. (2018) | Number of studies coded 1 |
|  | Reference number | [15] | [16] | [17] | [18] | [27] | [29] | [20] | [19] | [21] | [30] | [28] | [31] | [22] | [23] | [24] | [25] | [26] |  |
|  | **Reporting** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Subscale score (0 to 7)** | **7** | **7** | **6** | **7** | **7** | **5** | **7** | **6** | **7** | **4** | **7** | **6** | **6** | **6** | **6** | **4** | **4** | **Average score : 6** |
| 1 | Was the number of individuals screened or selected at each stage of defining the final sample reported? | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 16 |
| 2 | Were the descriptive statistics of the study participants adequately reported? | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 16 |
| 3 | Did the authors describe the key components of their statistical approaches? | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 17 |
| 4 | Were counfounder-adjusted estimates of treatment effects reported? | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 11 |
| 5 | Did the authors describe the statistical uncertainty of their findings? | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 13 |
| 6 | Was the extent of missing data reported? | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 16 |
| 7 | Were absolute and relative measures of treatment effect reported? | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 13 |
| Notes: questions from the grid assessing the credibility of observational studies (Berger et al. (2014)).  0 = quality item not fulfilled in the paper. 1 = quality item fulfilled in the paper. “NA” stands for “not available”. | | | | | | | | | | | | | | | | | | |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table A.3.5 : assessment of credibility with respect to *interpretation* domain | | | | | | | | | | | | | | | | | | |  |
|  | Question | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |  |
| Rothman et al. (2008) | Rochat et al. (2010) | Gobbens et al. (2012) | Hoeck et al. (2012) | Hogan et al. (2012) | Schoufour et al. (2014) | Hirani et al. (2015) | Bentur et al. (2016) | Bock et al. (2016) | Butler et al. (2016) | Campitelli et al. (2016) | Kelaiditi et al. (2016) | Gonzalez-Colaço et al. (2017) | Ferrante et al. (2018) | Hajek et al. (2018) | Lambotte et al. (2018) | Verve et al. (2018) | Number of studies coded 1 |
|  | Reference number | [15] | [16] | [17] | [18] | [27] | [29] | [20] | [19] | [21] | [30] | [28] | [31] | [22] | [23] | [24] | [25] | [26] |  |
|  | **Interpretation** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | **Subscale score (0 to 4)** | **3** | **3** | **4** | **4** | **3** | **4** | **4** | **3** | **4** | **3** | **3** | **3** | **3** | **3** | **3** | **3** | **3** | **Average score : 3,3** |
| 1 | Were the results consistent with prior known information or if not was an adequate explanation provided? | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 17 |
| 2 | Are the observed treatment effects considered clinically meaningful? | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 17 |
| 3 | Are the conclusions supported by the data and analysis presented? | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 17 |
| 4 | Was the effect of unmeasured confounding discussed? | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 |
| Notes: questions from the grid assessing the credibility of observational studies (Berger et al. (2014)).  0 = quality item not fulfilled in the paper. 1 = quality item fulfilled in the paper. “NA” stands for “not available”. | | | | | | | | | | | | | | | | | | |  |

Figure A.3.1. Spider chart for quality assessment by LTC outcome



Reading grid: The coloured area show for each type of LTC outcome, the average score (% of criteria fulfilled) on each quality dimension. For instance, Informal care studies have an average score (fulfilment criteria ratio) of 20% for the Analyses domain, and 100% for the Data domain.

**Appendix A.4. Details on each study: frailty/scarcopenia measures, outcomes, mains findings and statistical uncertainty**

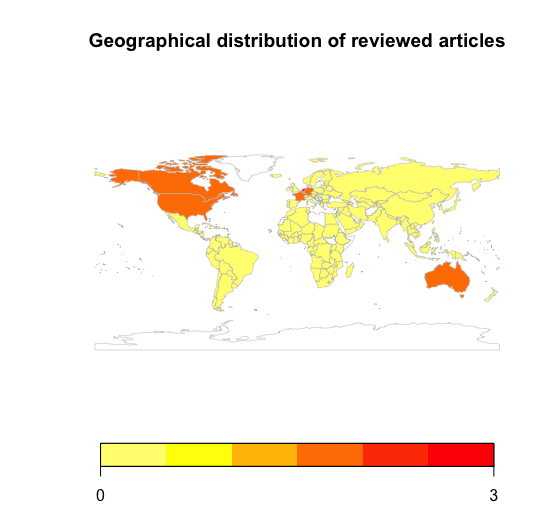
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **First Author, year, area, reference number** | **Frailty / sarcopenia measure** | | **LTC outcome** | **Crude estimate** | | **Adjusted estimate** |
|  | Definition / concept | Measurement | Outcome(s) analyzed | Value (95% CI or SE) | Estimate type | Value (95% CI or SE) |
| Rothman, 2008, USA, [15] | Slow gait speed | dummy (1/0) | Long-term nursing home (NH) stays | 5.9 (3.5-9.8) | HR | 3.9 (2.2-6.7) |
| Low physical activity | 3.5 (2.3-5.4) | 2.1 (1.3-3.3) |
| Weight loss | 2.0 (1.4-2.9) | 1.7 (1.2-2.4) |
| Exhustion | 2.0 (1.3-2.9) | 1.1 (0.8-1.7) |
| Weakness | 1.7 (1.1-2.7) | 1.0 (0.6-1.6) |
| Cognitive impairments | 3.7 (2.5-5.4) | 2.6 (1.7-4.0) |
| Depressive symptoms | 1.8 (1.3-2.6) | 1.4 (1.0-2.1) |
| Rochat., 2010, Australia | Fried criteria | Robust (0 criteria) [REF] Prefrail (1,2 criteria)  Frail (> 2 criteria) | Nurse visit (0/1) | prefrail = 2.08 (1.21-3.56) ;   frail = 10.08 (5.71-17.81) | OR | prefrail = 1.79 (1.01-3.17) ;   frail = 6.75 (3.42-13.30) |
| Help with meals or household duties (0/1) | prefrail = 3.44 (2.31-5.13) ;   frail = 11.54 (7.22-18.44) | OR | prefrail= 2.46 (1.59-3.82) ;   frail = 4.82 (2.76-8.43) |
| One night in NH (0/1) | prefrail = 1.95 (0.75-5.06) ;   frail = 8.0 (3.00-21.35) | OR | prefrail = 1.45 (0.54-3.93) ; frail = 3.42 (1.06-11.07) |
| Gobbens et al., 2012, Australia | Tilburg Frailty Indicator (TFI), part B | Score for physical frailty (0-8) | Personal care (0/1) | NA | - | Ph (2009, 2010) = (0.45\*\*, 0.63\*\*) |
| Nursing care (0/1) | NA | Ph (2009, 2010) = (0.13, -0.038) |
| Informal care (IC) (0/1) | NA | Ph (2009, 2010) = (0.14, 0.23) |
| NH / rehabilitation center stay (temporary) (0/1) | NA | Ph (2009, 2010) = (0.11, 1.26\*\*\*) |
| Hoeck, 2012, Belgium | Fried criteria | Robust (0 criteria) [REF] Prefrail (1,2 criteria)  Frail (> 2 criteria) | Home (0/1) | prefrail = 3.25 (2.41-4.37)   frail = 8.68 (5.69-13.22) | OR | prefrail = 2.09 (1.51-2.90)   frail = 4.29 (2.72-6.28) |
| Home help (0/1) | prefrail = 4.56 (3.04-6.82)   frail = 10.75 (6.34-18.22) | prefrail = 2.51 (1.61-3.90)   frail = 3.97 (2.21-7.16) |
| Meals-on-wheels (0/1) | prefrail = 2.89 (1.59-5.26)  frail = 6.91 (3.49-13.67) | prefrail = 1.86 (0.95-3.66)   frail = 3.23 (1.35-7.66) |
| Hogan, 2012, Canada | Armstrong Index | continuous scores coded into categorical :  Robust [REF] Prefrail Frail | Move to long-term care placemement (0/1) | prefrail = 2.25 (1.00-5.08)   frail = 4.21 (1.91-9.25) | RR | prefrail = 2.21 (0.98-4.99)   frail = 4.14 (1.87-9.14) |
| Full frailty index (83 items) | prefrail = 1.85 (1.26-2.72)  frail = 3.30 (2.30-4.75) | prefrail = 1.87 (1.27-2.75)   frail = 3.30 (2.29-4.76) |
| Cardiovascular Health Study (CHS) criteria | prefrail = 1.54 (0.95-2.51)  frail = 2.21 (1.41-3.46) | prefrail = 1.49 (0.91-2.43)  frail = 2.17 (1.38-3.41) |
| Changes in Health, End-stage disease and Signs and Symptoms; CHS: Cardiovascular Health Study (CHESS) scale | prefrail = 1.33 (0.98-1.81)   frail = 1.84 (1.38-2.47) | intermediate = 1.33 (0.95-1.82)   high = 1.87 (1.39-2.50) |
| Schoufour, 2014, Netherlands | Frailty index based on 51 items | continuous score (between 0 and 1) | increase in LTC intensity (expenses based on hours of care) after 3 years | NA | - | 1.09\*\*\* |
| Hirani, 2015, Australia | Appendicular lean mass (ALM)  Weakness Grait speed | [REF]: no sarcopenia Sarcopenia I = low ALM Sarcopenia II = low ALM + weakness Sarcopenia III = low ALM + weakness + poor gait speed | NH placement (0/1) | Sarcopenia I = 2.17 (1.41-3.36) Sarcopenia II =2.64 (1.44-4.86) Sarcopenia III = 2.57 (1.29-5.13) | HR | Sarcopenia I = 1.96 (1.14-3.35)   Sarcopenia II =2.53 (1.31-4.90)   Sarcopenia III = 2.27 (1.08-4.80) |
| Bentur, 2016, Israël | Vulnerable Elders Survey (13 items screening tool - VES-13). Analysis of frailty transitions | Robust [REF] Transition Frail More frail | Requires help on a regular basis (0/1) | Transition = 13.9 (3.6-49.6)  Frail = 17.6 (4.9-63.3)  More frail = 39.7 (11.5-137.1) | OR | none |
| Has a formal caregiver (0/1) | Transition = 10.5 (2.9-37.5)  Frail = 17.3 (4.7-63.6) More frail = 31.9 (9.2-110.5) | none |
| Home care (0/1) | Transition = 6.5 (1.4-30.6)   Frail = 17.3 (3.8-77.9)   More frail = 13.4 (3.1-58.2) | none |
| Bock, 2016, Germany | Fried criteria | Robust (0 criteria) [REF] Prefrail (1) Prefrail (2) Frail (3) Frail (4,5) | Nursing care costs (home and institution, not permanent) (0/1) | Prefrail (1) = 12.74 (10.35) Prefrail (2) = 8.31 (5.86)   Frail (3) = 76.96 (47.75)   Frail (4,5) = 334.93 (167.83) | Marginal effect | Prefrail (1) = 10.14 (9.06)   Prefrail (2) = 5.40 (4.81)   Frail (3) =48.97 (33.38)   Frail (4,5) = 181.0 (109.49) |
| Butler, 2016, Ireland | Fried criteria | Robust (0 criteria) [REF] Pre-frail (1-2) Frail (≥3) | Informal caregving costs (0/1) | NA | - | Frail 0.361 (p<0.0001) ; no CI REF: robust |
| Campitelli, 2016, Canda | Full FI, 83 items (Hogan et al) > 72 items in the data | Robust (FI < 0.2) [REF] Pre-frail (0.2 <= FI <= 0.3) Frail (FI > 0.3) | NH placement (0/1) | Pre-frail : 2.21 (2.16, 2.27) Frail : 3.92 (3.83, 4.01) | RR | Pre-frail : 2.20 (2.15, 2.26) Frail : 3.84 (3.75, 3.93) |
| Modified FI, 50 items (Armstrong et al) > 48 items in the data | Pre-frail : 2.13 (2.08, 2.19) Frail : 3.64 (3.55, 3.73) | Pre-frail : 2.11 (2.05, 2.16) Frail : 3.58 (3.50, 3.67) |
| CHESS scale | Robust (0 criteria) [REF] Pre-frail (1) Frail (≥ 2) | Pre-frail : 1.39 (1.35, 1.43) Frail : 1.96 (1.91, 2.02) | Pre-frail : 1.40 (1.36, 1.44) Frail : 1.99 (1.94, 2.04) |
| Kelaiditi, 2016, Europe | Frailty index : number of deficits over 30 identified, expressed in % | 1) Continuous variable  2) Categorical variable : Not frail (FI < 0.25) (REF) Frail (FI >= 0.25 ) | NH placement (0/1) | Continous : 1.018 (1.004, 1.032) Frail : 1.566 (1.112, 2.206) | HR | Continous :0.998 (0.983, 1.014)  Frail : 2.121 (1.352, 3.325) |
| Gonzalez-Colaço Harmand, 2017, France | Fried criteria | Robust (0 criteria) [REF] Pre-frail (1-2) Frail (≥ 3) | NH placement (0/1) | NA | - | Non-robust : 1.34 (0.97 - 1.84) (pvalue 0,07) |
| Rockwood frailty index> ratio of positive parameters from a list of preselected deficits. | REF: Robust (ratio< 0.2)  Non-robust (> 0.2) | Non-robust : 1.32 (0.87 - 2) (pvalue 0,20) |
| Tilburg Frailty Indicator (TFI), part B | 15 items considered;  REF : Robust : < 5 [REF] Frail (≥ 5) | Non-robust : 1.28 (0.92 - 1.80) (pvalue 0,15) |
| Ferrante, 2018, USA | Fried criteria | Robust (0 criteria) [REF] Pre-frail (1-2) Frail (≥ 3) | Incident NH admission post ICU | NA | - | Frail : 3.52 (1.23-10.08)  Pre-frail : 2.01 (0.77-5.24)  REF : Not frail |
| Hajek, 2018, Germany | Fried criteria | Robust (0 criteria) [REF] Pre-fail (1) Pre-fail (2) Frail (3) Frail (4-5) | Professional nursing care costs | NA | - | Pre-frail (1) : −0.04 (0.03) Pre-frail (2) : −0.03 (0.04) Frail (3): 0.04 (0.09) Frail (4 - 5) : 0.55\* (0.23) |
| Informal care costs | NA | Pre-frail (1) : 0.03 (0.04) Pre-frail (2) : 0.05 (0.06) Frail (3): 0.42\*\* (0.16) Frail (4 - 5) : 0.52 (0.35) |
| Lambotte, 2018, Belgium | Comprehensive Frailty Assessment Instrument - physical dimension | No-low frail  Middle frail   High frail | Combinations of formal and informal care (latent class analysis) | NA |  | NA |
| Verver, 2018, Netherlands | Tilburg Frailty Indicator (TFI), part B | Non frail (TFI < 6=5) Frail (TFI ≥ 5) | Number of formal caregivers | NA |  | NA |
| Presence of informal caregivers | NA |  | NA |
| Need for care or support | NA |  | NA |

**Notes:**

**Appendix A.5: Geographical distribution of studies**

|  |  |
| --- | --- |
| Table A.5.1 : studies’ country of origin | |
| Country | No. of publications per country |
| Netherlands | 3 |
| USA | 2 |
| Canada | 2 |
| Australia | 2 |
| Belgium | 2 |
| France | 2 |
| Germany | 2 |
| Ireland | 1 |
| Israel | 1 |

Figure A.5.1. Map of the geographical distribution of the 17 reviewed articles



Legend number of articles. Min = 0, max =3