Objectively measured physical capability levels and mortality: systematic review and meta-analysis

Rachel Cooper, MRC career development fellow Diana Kuh, professor of life course epidemiology and MRC unit director Rebecca Hardy, MRC programme leader, Mortality Review Group on behalf of the FALCon and HALCyon study teams

ABSTRACT
Objective To do a quantitative systematic review, including published and unpublished data, examining the associations between individual objective measures of physical capability (grip strength, walking speed, chair rising, and standing balance times) and mortality in community dwelling populations.

Design Systematic review and meta-analysis.
Data sources Relevant studies published by May 2009 identified through literature searches using Embase (from 1980) and Medline (from 1950) and manual searching of reference lists; unpublished results were obtained from study investigators.

Study selection Eligible observational studies were those done in community dwelling people of any age that examined the association of at least one of the specified measures of physical capability (grip strength, walking speed, chair rises, or standing balance) with mortality.

Data synthesis Effect estimates obtained were pooled by using random effects meta-analysis models with heterogeneity between studies investigated.

Results Although heterogeneity was detected, consistent evidence was found of associations between all four measures of physical capability and mortality; those who performed less well in these tests were found to be at higher risk of all cause mortality. For example, the summary hazard ratio for mortality comparing the weakest with the strongest quarter of grip strength (14 studies, 53 476 participants) was 1.67 (95% confidence interval 1.45 to 1.93) after adjustment for age, sex, and body size ($I^2=84.0\%$, 95% confidence interval 74% to 90%; $P$ from Q statistic $<0.001$). The summary hazard ratio for mortality comparing the slowest with the fastest quarter of walking speed (five studies, 14 692 participants) was 2.87 (2.22 to 3.72) ($I^2=25.2\%$, 0% to 70%; $P=0.25$) after similar adjustments. Whereas studies of the associations of walking speed, chair rising, and standing balance with mortality have only been done in older populations (average age over 70 years), the association of grip strength with mortality was also found in younger populations (five studies had an average age under 60 years).

Conclusions Objective measures of physical capability are predictors of all cause mortality in older community dwelling populations. Such measures may therefore provide useful tools for identifying older people at higher risk of death.

INTRODUCTION
Physical capability, a term used to describe a person’s ability to do the physical tasks of everyday living, can be assessed by self report or objectively by using tests such as grip strength, walking speed, chair rising, and standing balance. Growing evidence from single studies suggests that these objective measures of physical capability are also useful markers of current and future health. As a result, interest is increasing in these tests and their potential use as simple screening tools in the general population to identify people who may benefit from targeted intervention (such as strength training) or among patient groups to assess response to treatment.

No systematic review of the literature has been done to examine the associations of walking speed, standing balance, or chair rises with mortality. Existing reviews of grip strength have not done meta-analyses of results or investigated heterogeneity between studies.

Whether associations are restricted to older populations with comorbidities that adversely affect physical capability and increase risk of death or are also found when physical capability is assessed in other populations, including those that are younger, thus remains unclear.

We did a systematic review to test the hypothesis that lower levels of physical capability in community dwelling populations would be associated with higher subsequent risk of poor health and mortality. In this paper, we focus on all cause mortality as the end point. We aimed to do meta-analyses of results acquired from published studies and through contact with authors of studies to obtain estimates of the size of the associations of grip strength, walking speed, chair rise time, and standing balance performance with mortality. We also investigated whether these associations varied by sex, age at baseline, length of follow-up, or country of study.
METHODS

We did a systematic review of existing literature, following the meta-analysis of observational studies in epidemiology (MOOSE) guidelines and the PRISMA statement.67

Search strategy and selection criteria

Eligible observational studies were those done in people of any age living in the community that examined the association of at least one of the specified measures of physical capability (grip strength, walking speed, chair rises, or standing balance) with mortality (full review protocol available on request). We excluded studies of patient groups.

RC searched the electronic databases Medline (from 1950) and Embase (from 1980) to May 2009 by using free text search terms (see web appendix) with no restrictions by language. The initial search included search terms for other selected health outcomes, but those results are reported elsewhere.8

Figure 1 summarises the initial identification of studies. Combining the results of the electronic searches and removing duplicate records left abstracts of 2270 unique records to be screened independently by two authors (RC and RH). Examining the full text of 57 potentially eligible papers led to the identification of 25 papers with mortality as an outcome that were eligible for inclusion.8-93 RC and RH independently extracted the following data from these papers by using a standardised form: relevant published results; study population, selection, and baseline characteristics; study exclusion criteria; details of methods for ascertaining physical capability and mortality; and adjustments made for potential confounders. RC and RH made an assessment of each study’s quality on the basis of a modified version of the Newcastle-Ottawa quality assessment scale34; however, no studies were excluded on the basis of this assessment. The same two authors independently searched the reference lists of all eligible papers and identified a further 64 abstracts to screen, resulting in the identification of a further three papers that were eligible for inclusion.35-37 Disagreements about the eligibility of a study or differences between the two sets of information extracted were resolved through discussion.

Contact with study authors

We contacted the corresponding authors of 23 of the 28 eligible papers and asked them to complete standardised results tables (copies available on request) (fig 2). Of the five papers for which authors were not contacted, we could not find up to date contact details for two,15 22 the third had used the same study population as other authors being contacted but had shorter follow-up,35 and a further two were not identified until after the deadline for contacting authors.27 28

Citations identified and screened (n=2270)

Citations excluded on basis of content of title and abstract (n=2213)
Papers retrieved for detailed assessment (n=57)
Papers excluded (n=22):
  Letter to editor, no new results (n=1)
  Examined outcome that was not eligible (n=1)
  Cross sectional (n=7)
  Non-community dwelling populations (n=2)
  No clearly defined outcome (n=1)
  Review paper; no new results (n=1)
  Had relevant measures but did not examine association of interest (n=2)
  Used composite score of physical performance (n=7)
Papers for full data extraction (n=35) (25 examined mortality)

Additional citations identified from reference lists of 35 eligible papers (n=64)

Citations excluded on basis of content of title and abstract (n=43)
Papers retrieved for detailed assessment (n=21)
Papers excluded (n=13):
  Did not examine relevant physical capability measure (n=6)
  Used composite score of physical performance (n=4)
  Non-community dwelling population (n=1)
  Did not present relevant results (n=1)
  Had insufficient events for analysis (only 6 deaths) (n=1)
Papers for full data extraction (n=8) (3 examined mortality)
Papers for inclusion in reviews (n=43)
Papers for inclusion in mortality review (n=28)

Fig 1 | Flow diagram for identification of published studies. Includes identification of studies for additional review of other health outcomes, reported elsewhere8

Fig 2 | Flow diagram showing contact with authors and ascertainment of results for inclusion in review
We asked authors to analyse the associations between each individual measure of physical capability available and all cause mortality, using survival analyses where possible, on men and women separately and also combined. We requested three separate sets of models: unadjusted, age adjusted, and age and body size (height and weight or body mass index) adjusted, with the addition of sex in the second and third models when run on men and women combined. For all capability measures, we asked authors for a comparison of sex specific quarters with the best performing group used as the reference category. For grip strength, we also asked authors to provide estimates for a one unit increase (we did not request estimates for a one unit increase for the other three capability measures because of the variation in the units of measurement). After sending one reminder, we received responses to 17 of the 23 requests and 12 standard sets of results were provided; the other five responses informed us of the author’s inability to provide results (fig 2).

We also contacted the corresponding authors of a further nine papers identified during screening of abstracts as having potentially relevant data but who had not published results from tests of the association between physical capability measures and mortality.38-46 This resulted in five sets of additional results for inclusion (fig 2).

**Statistical methods**

We based analyses on the standard requested results if they had been provided, using published results when requested results were not available. If necessary, we converted effect estimates for a per unit change in grip strength to a per 1 kg increase in grip strength.

When more than three sets of comparable results were available, we did random effects meta-analyses.47 We ran these analyses first on the sex and age adjusted estimates and then with additional adjustment for body size. We then added estimates from models with alternative adjustments into the meta-analysis for those studies (n=4) for which standard adjustments were not available; this was the case only for analyses of per unit change in grip strength. In all these initial analyses, we used the combined sex adjusted effects for men and women when they were available and the sex specific estimates only when they were not. We used $I^2$ and Q statistics to investigate between study heterogeneity.48 49 We examined potential sources of heterogeneity, including sex, mean age of study participants, length of follow-up, and region of study (as pre-specified in the protocol), by stratifying meta-analyses by each of these factors and by doing meta-regression analyses when we had more than 10 sets of results.50 We used permutation tests to obtain P values corrected for multiple testing.51 We also re-ran the main meta-analyses with each study removed in turn to check that no one study was explaining any heterogeneity found. We examined funnel plots to assess publication bias.52 We also used the tests proposed by Egger and colleagues and Begg and Mazumdar for this purpose.53 54

**RESULTS**

**Study characteristics**

We considered most of the studies to be of good quality (web table A). This was partly because physical capability was usually measured by trained professionals either in a home or clinic setting and mortality was ascertained by methods considered to be reliable, such as record linkage which usually results in little loss to follow-up.

Most studies were done in older populations (web table A and tables 1 and 3). None of the studies that reported on associations of walking speed, chair rises, or standing balance with mortality was in a population with an average age under 61 years, and in most the average age was over 70 years. Studies of grip strength and mortality covered a wider range of ages, and five had an average age of under 60 years.

Although similar numbers of studies of grip strength and mortality had short (≤5 years), intermediate (6-20 years), and longer (>20 years) follow-up, follow-up in studies of walking speed has usually been five years or less (no studies reported follow-up of more than 10 years), and in studies of chair rises follow-up has usually been between six and 10 years (tables 1 and 3). Most studies have been done in the United States, but studies have also been done across Europe and in Japan, Hong Kong, and Australia.

Although we did not assess the protocols for measuring physical capability used in different studies as part of our review, we did record that a range of different instruments have been used to measure grip strength. Most of these instruments are different types of handheld dynamometer, and authors tend to use either the maximum value recorded or the average achieved over a fixed number of trials in their models. Walking speed has been measured in different studies over distances of 4, 6, and 10 metres and 8 and 16 feet. Of the studies of chair rises identified, all had asked participants to do five rises.

We had a total of 33 sets of results for consideration in this review, consisting of 16 sets of published results and 17 sets of results in a specified format (web table A). Grip strength was the most frequently examined measure in the published literature; 19 of the 28 papers identified reported on the association of this measure with mortality. Fourteen papers reported results for the association between walking speed or time and mortality, four reported on standing balance, and three reported on chair rises. Of the five additional studies included after contact with authors, four provided results for grip strength, two for walking speed, four for chair rises, and one for standing balance.

**Assessment of publication bias**

Examination of funnel plots suggested no clear evidence of publication bias. This was confirmed by the results from tests proposed by Egger and colleagues and Begg and Mazumdar,53 54 which produced P values of 0.28 or greater.
Table 1 | Results from stratified meta-analyses of age, sex, and body size adjusted hazard ratios of associations of grip strength with all cause mortality

| Stratification | No* | Summary hazard ratio (95% CI) | I² (%) | P value‡ | No* | Summary hazard ratio (95% CI) | I² (%) | P value‡ |
|---------------|-----|-----------------------------|--------|---------|-----|-----------------------------|--------|---------|
| Mean age at baseline (years): | | | | | | | | |
| ≤60 | 3 | 0.97 (0.97 to 0.98) | 13.7 | 0.31 | 4 | 1.43 (1.07 to 1.91) | 86.4 | <0.01 |
| 61-70 | 4 | 0.97 (0.94 to 1.01) | 90.7 | <0.01 | 2 | 1.81 (0.73 to 4.49) | 68.0 | 0.08 |
| >70 | 7 | 0.97 (0.96 to 0.98) | 90.5 | <0.01 | 8 | 1.80 (1.48 to 2.18) | 81.9 | <0.01 |
| Length of follow-up (years): | | | | | | | | |
| ≤5 | 5 | 0.96 (0.95 to 0.97) | 39.7 | 0.16 | 5 | 2.16 (1.70 to 2.75) | 31.9 | 0.21 |
| 6-10 | 4 | 0.98 (0.95 to 1.00) | 95.8 | <0.01 | 2 | 2.26 (1.88 to 2.72) | 0 | 0.70 |
| 11-20 | 2 | 0.97 (0.97 to 0.98) | 0 | 0.34 | 2 | 1.43 (1.33 to 1.54) | 0 | 0.91 |
| >20 | 3 | 0.98 (0.97 to 0.99) | 90.7 | <0.01 | 5 | 1.39 (1.15 to 1.70) | 81.9 | <0.01 |
| Region of study: | | | | | | | | |
| North America | 8 | 0.97 (0.96 to 0.97) | 60.4 | 0.01 | 7 | 1.64 (1.35 to 2.01) | 89.1 | <0.01 |
| Japan | 4 | 0.99 (0.97 to 1.00) | 87.1 | <0.01 | 1 | 1.98 (1.64 to 2.40) | – | – |
| Other | 2 | 0.98 (0.94 to 1.01) | 72.3 | 0.06 | 6 | 1.64 (1.31 to 2.06) | 59.1 | 0.03 |
| Sex: | | | | | | | | |
| Male | 10 | 0.97 (0.96 to 0.98) | 91.6 | <0.01 | 12 | 1.75 (1.41 to 2.16) | 83.5 | <0.01 |
| Female | 9 | 0.97 (0.96 to 0.98) | 71.3 | <0.01 | 10 | 1.46 (1.29 to 1.66) | 34.2 | 0.13 |

*Number of data points. †‡14/14 estimates included are from models with multiple adjustments rather than standard adjustments of age, sex, and body size. §From Cochran’s Q statistic.

Table 2 | Summary of results from studies of grip strength not included in meta-analyses

| Study name and reference/s* | Total No (No of deaths) | Category comparison/value of unit change | Effect estimate (95% CI) | Adjustments |
|-----------------------------|-------------------------|----------------------------------------|-------------------------|-------------|
| Prospective Japanese study, Fujita et al 1995†‡ | Men 2068 (113); women: 1988 (42) | Low v high performance (as judged against a standard) | Odds ratio 1.86 (1.13 to 3.07) | Age, sex |
| EVERGREEN project, Laukkanen et al 1995‡ | 463 (74) | Below v above mean | Relative risk: men 1.92 (1.16 to 3.16); women 0.84 (0.38 to 1.86) | – |
| Edinburgh Longitudinal Study of Ageing, Milne and Maule 1984‡ | 483 (135) | NA | No effect estimates presented—mean baseline grip strength lower in people who died during follow-up than in those who survived | NA |
| Precipitating Events Project, Rothman et al 2008‡ | 784 (286) | Weak (that is, lower than sex and body mass index specific threshold) v not weak | Hazard ratio: 1.8 (1.3 to 2.5) | Age, sex, race, education, chronic conditions |

NA = not available.

*See web table A for further details of studies.

Grip strength and mortality

A total of 23 sets of results from tests of the association between grip strength and mortality were available, and these were all from unique studies with the exception of one study (the Honolulu Heart Program). Different authors used this one study to contribute two sets of results, but as one author provided effect estimates for the per unit change in grip strength and the other provided results of comparisons between quarters,25,33 we could include both in meta-analyses without the same study population being included more than once in the same meta-analysis.

Meta-analysis of the association between a unit change in grip strength and mortality was based on 13 studies, one of which contributed two data points as men and women had been analysed separately (fig 3).30 Two other studies were not included: one because it had used the sum of grip strength in both hands,41 when all other studies had used the maximum measure achieved in one hand or an average per hand, and the other because grip strength had been measured in N/m² rather than kg.20 (We instead included these two studies in the meta-analyses of comparisons of quarters described below.)

Most of the 13 studies (with a total of 44 636 participants) included in this first meta-analysis found that higher grip strength was associated with lower subsequent mortality; the overall summary hazard ratio of...
Table 3 | Results from stratified meta-analyses of age, sex, and body size adjusted hazard ratios of associations of walking speed and chair rise time with all cause mortality: lowest versus highest quarter comparison

| Stratification | Walking speed | Chair rise time |
|---------------|--------------|----------------|
|               | No*          | Summary hazard ratio (95% CI) | I² (%) | P value† | No*          | Summary hazard ratio (95% CI) | I² (%) | P value† |
| None          | 5            | 2.87 (2.22 to 3.72) | 25.2 | 0.25     | 5            | 1.96 (1.56 to 2.46) | 81.9 | 0.01     |
| Mean age at baseline (years): | | | | | | | | |
| 60–69         | 0            | –              | –     | –       | 0            | –              | –     | –       |
| 70–79         | 1            | 6.25 (2.79 to 14.02) | –    | –       | 0            | –              | 81.9 | 0.01     |
| >79           | 4            | 2.68 (2.14 to 3.35) | 0    | 0.71     | 5            | 1.96 (1.56 to 2.46) | 81.9 | 0.01     |
| Length of follow-up (years): | | | | | | | | |
| ≤5            | 4            | 3.11 (2.37 to 4.09) | 13.3 | 0.33     | 1            | 1.90 (1.26 to 2.87) | –    | –       |
| 6–10          | 1            | 2.16 (1.38 to 3.38) | –    | –       | 3            | 2.14 (1.86 to 2.48) | 8.7  | 0.34     |
| ≥11–20        | 3            | 1.50 (1.38 to 1.63) | –    | –       | –            | –              | –    | –       |
| ≥20           | 0            | –              | –     | –       | –            | –              | –    | –       |
| Region of study: | | | | | | | | |
| North America | 2            | 3.95 (1.89 to 8.28) | 61.9 | 0.11     | 3            | 1.85 (1.42 to 2.41) | 88.0 | 0.01     |
| Japan         | 0            | –              | –     | –       | –            | –              | –    | –       |
| Other         | 3            | 2.61 (2.02 to 3.37) | 0    | 0.54     | 2            | 2.26 (1.58 to 3.23) | 27.7 | 0.24     |
| Sex‡          |              |                |      |         |              |                |      |         |
| Male          | 4            | 2.65 (2.01 to 3.48) | 0    | 0.58     | 3            | 2.17 (1.82 to 2.58) | 17.3 | 0.30     |
| Female        | 4            | 3.19 (2.20 to 4.63) | 6.7  | 0.36     | 2            | 2.17 (1.36 to 2.30) | 67.0 | 0.05     

*Number of data points.
†From Cochran’s Q statistic.
‡Total numbers differ for sex stratified meta-analyses for following reasons: walking speed quarter comparison estimates include two studies of men only and two studies of women only (unstratified summary hazard ratio for comparison 2.83 (2.28 to 3.51), n=8, I²=90%, P<0.05); chair rise time quarter comparison estimates include two studies of men only and two studies of women only (unstratified summary hazard ratio for comparison 2.00 (1.62 to 2.46), n=6, I²=78.1%, P<0.01).

mortality associated with a 1 kg increase in grip strength estimated from a random effects model was 0.97 (95% confidence interval 0.96 to 0.98) with adjustment for age, sex (where appropriate), and body size or with multiple adjustments for four studies (see footnote to fig 3). Although all effect estimates were in the same direction, we found evidence of heterogeneity between studies (I²=89.5%, 95% confidence interval 84% to 93%; P from Q statistic<0.001). However, when we stratified the meta-analyses by the pre-specified characteristics and did meta-regression analyses we found no clear evidence that any of these factors explained the heterogeneity (table 1). Removal of each study from the meta-analyses did not greatly affect estimates of the level of heterogeneity [results not shown]. Findings were similar when we analysed results from models adjusted for age and sex only.

Effect estimates from 14 studies (53 476 participants) that had compared quarters of grip strength were available for inclusion in meta-analyses (fig 4). Most of these studies found that people in the weakest sex specific quarter of grip strength had significantly higher rates of mortality than did those in the strongest quarter; the overall summary hazard ratio of mortality comparing the lowest with the highest quarter was 1.67 (1.45 to 1.93) with adjustment for age, sex, and body size. Summary hazard ratios from the meta-analyses of effect estimates comparing the second weakest and second strongest quarter with the strongest quarter showed a graded effect (fig 5).
Although most of the estimates from comparisons of quarters were in the same direction, we found evidence of heterogeneity between studies (I²=84.0%, 74% to 90%; P from Q statistic<0.001). When we stratified meta-analyses (table 1), the association between grip strength and mortality seemed to be weaker in studies with an average age at baseline under 60 years compared with studies with an older average age at baseline and also in studies with follow-up of more than 20 years compared with studies with shorter follow-up. We found further evidence of a weaker association between grip strength and mortality in studies with longer follow-up in meta-regression analyses comparing the weakest and strongest quarters of grip strength (I²=0.054 from permutation test controlling for multiple testing). The inclusion of follow-up time reduced the estimate of between study variance from 0.06 to 0.02, and the variation in effect size by follow-up time was independent of age (I²=0.065 for length of follow-up from permutation test). We found no evidence from meta-regression analyses in support of differences in summary estimates between categories of the other pre-specified characteristics, and the removal of each study from meta-analyses did not greatly affect estimates of the level of heterogeneity.

Of the four studies of grip strength that could not be included in at least one of the meta-analyses, most results were consistent with the findings from meta-analyses (table 2). The one exception was among women in the study by Fujita et al.15

Walking speed and mortality

A total of 16 sets of results from analyses of the association between walking speed or time and mortality were available; analyses were done in 11 different study populations. The Hispanic Established Population for the Epidemiological Study of the Elderly (H-EPESE)10,35,36 and another unnamed US study17,18,24 had both been used three times, and the Health, Aging and Body Composition (ABC) study had been used twice,12,14, at different follow-up times (web table A).

We had seven sets of results in the form of comparisons of quarters, but three were based on the same study population (H-EPESE) and we included only one set of these (those with the requested standard adjustments) in the main meta-analysis. These five studies (14 692 participants) all found that people in the slowest quarter of walking speed had significantly higher mortality rates than did those in the fastest quarter (fig 6); the overall summary hazard ratio of mortality comparing the slowest with the fastest quarter was 2.87 (2.22 to 3.72) with adjustment for age, sex, and body size. Heterogeneity was moderate (I²=25.2%, 0% to 70%; P from Q statistic<0.05), and the studies varied little in the pre-specified characteristics that could have explained this (table 3). Summary hazard ratios from the meta-analyses of effect estimates comparing the second slowest and second fastest quarters with the fastest quarter were also consistent and showed a graded effect (fig 5). Findings were similar when we re-ran meta-analyses including estimates from the other analyses of H-EPESE in place of those provided by Al Snih et al.10 and Ostir et al36 in place of those provided by Al Snih et al.10.
Table 4 | Summary of results from studies of walking speed and chair rises not included in meta-analyses

| Study name and reference/s* | Total No (No of deaths) | Category comparison/ value of unit change | Effect estimate (95% CI) | Adjustments |
|-----------------------------|-------------------------|-----------------------------------------|-------------------------|-------------|
| **Walking speed**           |                         |                                         |                         |             |
| Health ABC study, Cesari et al 2005 and 2009*14 | Total 1016 (163)‡; 3024 (653)‡ | Speed -1.00 m/s v 1.00 m/s | Relative risk 1.64 (1.14 to 2.37); 1.49 (1.23 to 1.80)‡ | Age, sex, race, study site, smoking, BMI, MMSE score, physical activity, comorbidities (and also alcohol consumption and education for 6.9 year follow-up) |
| iISIRENTE, Cesari et al 2008*13 | 335 (71) | Per 1 SD increase | Hazard ratio 0.73 (0.54 to 0.99) | Age, sex, BMI, cognitive performance, No of clinical conditions, albumin, total cholesterol |
| **Chair rises**             |                         |                                         |                         |             |
| Health ABC study, Cesari et al 2009*14 | Total 3024 (653) | Time for 5 rises ≥ 17 seconds v < 17 seconds | Hazard ratio 1.40 (1.17 to 1.68) | Age, sex, race, study site, smoking, BMI, MMSE score, physical activity, comorbidities, alcohol consumption, education |
| iISIRENTE, Cesari et al 2008*13 | 335 (71) | Per 1 SD increase in chair rise score—from 0 (unable) to 4 (≤ 11 seconds) | Hazard ratio 0.51 (0.36 to 0.72) | Age, sex, BMI, cognitive performance, No of clinical conditions, albumin, total cholesterol |

BM=body mass index; MMSE=mini-mental state examination.
*See web table A for further details of studies.
‡4.9 year follow-up (on restricted sample).
§6.9 year follow-up (on full sample).
¶Results taken from Perera et al 2005,24 but results from two other papers on this study are consistent.17,18

meta-analyses (table 4). All evidence supported an association between slower walking speed and higher mortality.

Chair rise time and mortality
All seven sets of results from analyses of the association between chair rise time and mortality were from unique studies. Five were in the form of comparisons of quarters and so could be included in meta-analyses. These five studies (28 036 participants) all found that people in the slowest quarter of chair rise time had significantly higher mortality rates than did those in the fastest quarter (fig 7); the overall summary hazard ratio of mortality comparing the slowest with the fastest quarter was 1.96 (1.56 to 2.45) with adjustment for age, sex, and body size. Although all effect estimates were in the same direction, we found evidence of a high degree of heterogeneity between studies (I²=81.9%, 58% to 92%; P from Q statistic<0.001). When we stratified the meta-analyses by the pre-specified characteristics (table 3), the effect seemed to be stronger in men than in women, but we had too few data points to investigate this further. The variation in the other characteristics was also insufficient to explain the heterogeneity, and the removal of each study from meta-analyses did not affect estimates of the level of heterogeneity (results not shown). Summary hazard ratios from the meta-analyses of effect estimates comparing the second slowest and second fastest quarters with the fastest quarter were also consistent and showed a graded effect (fig 5). Three of the studies also reported effect estimates from comparisons of people unable to do chair rises with those in the fastest quarter; the summary hazard ratio for mortality from a meta-analysis of these three results suggested that those unable to do chair rises had the highest rates of mortality (4.09, 2.24 to 4.42). The findings from the two studies not included in the meta-analysis (table 4) were consistent with those from the meta-analysis and also reported that slower chair rise time was associated with higher mortality.

Standing balance and mortality
All five sets of results from analyses testing the association between standing balance and mortality were
from unique studies. However, as standing balance had not been measured and categorised in comparable ways across studies, we could not do meta-analyses of these results. Although all five studies found some evidence that poorer performance in standing balance tests was associated with higher mortality rates, these associations were not all statistically significant at conventional levels (table 5).

DISCUSSION

We have found evidence of associations between all four measures of physical capability investigated (grip strength, walking speed, chair rises, and standing balance) and all cause mortality. People in community dwelling populations who perform less well in these tests were consistently found to be at higher risk of death. The estimates from meta-analyses for grip strength, walking speed, and chair rises show a dose-response relation. With the exception of grip strength, studies have been done exclusively in older populations, and most have relatively short follow-up, so whether similar associations would be found at younger ages or after participants have been followed-up for longer than 10 years is unclear. For grip strength, we found evidence of an association even in populations with an average age at baseline of less than 60 years, although the association weakened with increasing length of follow-up.

Explanation of findings

Several possible explanations exist, which are not necessarily exclusive, for finding associations between objective measures of physical capability and mortality in community dwelling populations. Firstly, these findings could be explained by confounding (for example, by socioeconomic position or levels of physical activity), as effect estimates have only been adjusted for age, sex, and body size. These factors were considered to be the most likely confounders yet did not explain the associations of the physical capability measures with mortality. Apart from these, no standard set of multiple adjustments across papers existed and we thought that requesting further adjustments would have affected the response from study authors and led to inconsistencies in adjustments across studies.

Secondly, these measures of physical capability could be markers of disease and general health status. Some of the community dwelling populations included in this review consisted of people with diseases or comorbidities that were not considered severe enough to warrant exclusion from the study but that may have affected both their physical performance and mortality risk. This could definitely apply to walking speed, chair rises, and standing balance, for which some of the community dwelling populations who perform less well in these tests were consistently found to be at higher risk of death. The estimates from meta-analyses for grip strength, walking speed, and chair rises show a dose-response relation. With the exception of grip strength, studies have been done exclusively in older populations, and most have relatively short follow-up, so whether similar associations would be found at younger ages or after participants have been followed-up for longer than 10 years is unclear. For grip strength, we found evidence of an association even in populations with an average age at baseline of less than 60 years, although the association weakened with increasing length of follow-up.

Explanation of findings

Several possible explanations exist, which are not necessarily exclusive, for finding associations between objective measures of physical capability and mortality in community dwelling populations. Firstly, these findings could be explained by confounding (for example, by socioeconomic position or levels of physical activity), as effect estimates have only been adjusted for age, sex, and body size. These factors were considered to be the most likely confounders yet did not explain the associations of the physical capability measures with mortality. Apart from these, no standard set of multiple adjustments across papers existed and we thought that requesting further adjustments would have affected the response from study authors and led to inconsistencies in adjustments across studies.

Secondly, these measures of physical capability could be markers of disease and general health status. Some of the community dwelling populations included in this review consisted of people with diseases or comorbidities that were not considered severe enough to warrant exclusion from the study but that may have affected both their physical performance and mortality risk. This could definitely apply to walking speed, chair rises, and standing balance, for which some of the community dwelling populations who perform less well in these tests were consistently found to be at higher risk of death. The estimates from meta-analyses for grip strength, walking speed, and chair rises show a dose-response relation. With the exception of grip strength, studies have been done exclusively in older populations, and most have relatively short follow-up, so whether similar associations would be found at younger ages or after participants have been followed-up for longer than 10 years is unclear. For grip strength, we found evidence of an association even in populations with an average age at baseline of less than 60 years, although the association weakened with increasing length of follow-up.

Table 5 | Summary of results from studies of standing balance

| Study name and reference/ | Total No (No of deaths) | Category comparison/value of unit change | Effect estimate (95% CI) | Adjustments |
|--------------------------|-------------------------|-----------------------------------------|--------------------------|-------------|
| Health ABC study, Cesari et al 2009 | 3024 (653) | Balance ≥53 seconds v 8 53 seconds (note: max score 90 seconds) | Hazard ratio 1.35 (1.12 to 1.62) | Age, sex, BMI, MMSE score, smoking, alcohol consumption, education |
| IJSRENTIE, Cesari et al 2008 | 335 (71) | Per 1 SD increase in standing balance score —from 0 (unable) to 4 (hold tandem stand for 10 seconds) | Hazard ratio 0.77 (0.60 to 1.00) | Age, sex, BMI, cognitive performance, No of clinical conditions, albumin, total cholesterol |
| EPESE, Guralnik et al 1994 | 5264 (1741) | 1) Unable to hold side by side stand v able to hold tandem stand for 10 seconds; 2) Able to hold side by side stand for 10 seconds but unable to hold semi-tandem stand for 10 seconds v able to hold tandem stand for 10 seconds | Hazard ratio 3.54 (3.04 to 4.13); 2) 1.78 (1.51 to 2.09) | Age, sex, height, weight |
| EPIDOS study, Rolland et al 2006 | 7092 (722) | Lowest (0-20 seconds) v highest third (27-30 seconds) | Hazard ratio 1.57 (1.32 to 1.87) | Age, sex, body mass index |
| Study of Fukuoka Prefecture residents, Takata et al 2007 | 551 (72) | Per 1 unit change in balance time | Hazard ratio 0.99 (0.97 to 1.01) | Sex, smoking, BMI, systolic blood pressure, marital status, total cholesterol, glucose, complications from prevalent disease |

BMI=body mass index; MMSE=mini-mental state examination.

*See web table A for further details of studies.

Fig 7 | Hazard ratios of mortality comparing slowest with fastest quarter of chair rise time with adjustment for age, sex (where appropriate), and body size. B=both sexes; F=women only; M=men only; MrOS=Osteoporotic Fractures in Men Study; SOF=Study of Osteoporotic Fractures
the nature of their design (for example, recruitment of men from the active workforce) excluded people with health problems.25

A related possibility is that underlying ageing processes led to poorer performance and a higher probability of chronic disease and death. Walking speed, chair rising, and standing balance require strength, balance, and motor control; walking speed and chair rising also require muscle power and speed and adequate cardiorespiratory function; standing balance requires mental concentration. These functions decline with age, may co-occur, and contribute to the risk of frailty. The progressive dysregulation of homoeostatic equilibrium across multiple systems may be the biological basis of frailty; common pathways proposed include endocrine dysfunction, inflammation, oxidative stress, and disequilibrium between the sympathetic and parasympathetic systems.35

Functional status in later life reflects the peak achieved during growth and development, as well as the rate of decline. Thus, the relation between these measures of physical capability, even when measured at younger ages, and mortality could also reflect initial differences in development that affect both. Evidence that early growth, cognitive and motor development, and childhood social environment are associated with adult physical capability, chronic disease, and mortality support this possibility.35-62

Heterogeneity

Many of the estimates of $I^2$ calculated in these meta-analyses would be judged to be high.46 We have presented summary, overall estimates from meta-analyses despite this, as most effect estimates from individual studies were in the same direction and doing meta-analyses has been argued to still be appropriate.63-64 Furthermore, the value of $I^2$ depends on both the within study and the between study variance,48 65 and as many of the studies included in this review have precise estimates (figs 3 and 4), either because the study population was large or a high proportion of participants died, the values of $I^2$ will have been affected by this. However, these summary estimates, which are an average of estimates across populations with different characteristics, should be cited with caution, and the reasons for finding these levels of heterogeneity should be explored.

Insufficient variation often existed in the characteristics we proposed a priori as possible sources of heterogeneity to allow us to examine their role in explaining heterogeneity fully—for example, four of the five studies of walking speed and all five studies of chair rise time included in meta-analyses were in populations aged over 70. However, meta-regression analyses of comparisons of quarters of grip strength suggested that associations with mortality were weaker in those studies with longer follow-up, even after adjustment for age. Comorbidities at the time of assessment of capability, which would increase short term mortality, may be more likely to explain the association between grip strength and mortality in studies with shorter follow-up. With increasing length of follow-up, the proportional hazards assumption within a study may be violated.

In addition to the characteristics investigated, many others vary between studies and could result in heterogeneity. These include differences in exclusion criteria, the instruments used, the main causes of death, levels of underlying comorbidity, and ethnic diversity.

Implications

This review has highlighted the paucity of studies that have measured physical capability in younger populations with subsequent follow-up for mortality. This situation is expected to change; these measures are being introduced in studies of younger populations as overall markers of functioning at the multi-system level, rather than as markers of severity and stage of specific chronic diseases.66 However, investigation of associations with mortality in studies with measurement of capability at younger ages will obviously need lengthy follow-up. Research is also needed to examine the associations between changes in capability with age and mortality, as a steep decline in physical capability may be a better predictor of mortality than is the absolute level at a single point in time. In addition, associations between these measures and cause specific mortality and other health outcomes may help to elucidate the pathways underlying the associations with all cause mortality, although few studies identified reported on these. Elucidating the underlying biological pathways that link poorer capability to mortality will inform the development of effective interventions.

We chose to examine the relation between each individual measure of physical capability, by using a standardised exposure measurement (comparisons of quarters), and mortality. The rationale behind this is that a variety of composite scores exist that are derived by using these measurements in combination, but whether results with such scores are driven by one measure or whether they each make a similar contribution is unclear. Although our findings suggest that all four measures of physical capability assessed are associated with all cause mortality, the relative paucity of data for walking speed, chair rises, and standing balance makes us cautious about drawing conclusions on their relative strengths. As these measures of physical capability are highly correlated with each other, more studies are needed that consider the value of each additional test once the findings for one test are known.14 For clinical practice, investigating whether a derived composite score representing overall lower or upper body function, such as the short physical performance battery score65 or one of the frailty indices, may be a stronger predictor of mortality than any of the individual measures are by themselves would be of interest.

The associations found between measures of physical capability and mortality seem to operate across the whole range of ability, with no apparent threshold effect. Therefore, if these measures were to be used as screening tools, clinicians and researchers would need to identify thresholds with caution and recognise that...
differences in the most appropriate place to set these may exist, depending on the characteristics of the population to be screened. Ultimately, randomised controlled trials will be needed to determine whether interventions aimed at improving physical capability are effective at improving capability and as a consequence are effective at reducing morbidity and mortality.

Strengths and limitations

The main strengths of this systematic review are its inclusion of several measures of physical capability and its inclusion of as many relevant studies as possible by making contact with study authors. By following a strict protocol, testing a priori hypotheses, and including unpublished results, we have minimised bias.

The study also has some limitations, although none of these affects our conclusions. Results from each study are based on the assumption of proportional hazards, which may not hold in all studies. The meta-regression analyses were likely to be underpowered, as were the formal tests of publication bias, especially for walking speed and chair rises for which we had less than the recommended number of data points. That the funnel plots and formal statistical tests produced no clear evidence of publication bias should thus be interpreted with caution. However, our success in obtaining unpublished results should limit publication bias.

Conclusions

This review shows the value of objective measures of physical capability as predictors of subsequent mortality in older community dwelling populations. Grip strength measured at younger ages also predicted mortality, but whether walking speed, chair rise time, and standing balance performance are associated with mortality in younger populations remains to be seen.

We thank the following people who helped to provide data for this review: Kaarin Anstey (Aging Research Unit and Centre for Mental Health Research, Australian National University, Canberra, Australia), Cora L Craig (Canadian Fitness Survey), Peter Katzmarzyk (Canadian Fitness Survey, Pennington Biomedical Research Center), Ronald Klein, Michael D Knutson, Kristine E Lee (Department of Ophthalmology and Visual Sciences, University of Wisconsin, Madison, WI, USA), Kevin Morgan (School of Sport, Exercise and Health Sciences, Loughborough University, UK), Michiko Yamada, and Shoichiro Fujita. We also thank all other members of the FALCon study team (Eleni Bakra, Michaela Benzeval, Fiona Matthews, and Graciela Munz Turres) for their helpful comments and Hugh McGuire (National Collaborating Centre for Women’s and Children’s Health), who provided advice on systematic reviews and performing literature searches.

The HALCyon study team also includes Yoav Ben-Shlomo, Cyrus Cooper, Leonie Craig, Ian Day, Richard Martin, Kate Tilling, Tamuno Alfred, Mike Gardner, John Gallacher, Ian Deary, John Starr, Paul Sheils, Thomas von Zglinicki, Humphrey Southall, Paulaaucott, Jane Elliott, Andrew Steptoe, Chris Power, Geraldine McNeill, Alison Lennox, Marcus Richards, Gita Mishra, Zeinab Mulla, Emily Lemelin, and James Goodwin.

The following are members of the Mortality Review Group, all of whom contributed equally to this work: Rabiah Ahmad (Universiti Teknologi Malaysia), Awan Azieh Sayer (MRC Epidemiology Resource Centre, University of Southampton), Sohaim Al Snih (Sealy Center on Aging), Peter A Bath (University of Sheffield), Peggy M Cawthon (California Pacific Medical Center Research Institute), J David Curb (Keck Medical Center and University of Hawaii), Kristine E Ensrud (National Institute on Aging), Catherine R Gale (MRC Epidemiology Resource Centre, University of Southampton), Jack M Guralnik (National Institute on Aging), Suzanne Ho (Chinese University of Hong Kong), Fumiyoshi Kasagi (Radiation Effects Research Foundation, Japan), Barbara E K Klein (University of Wisconsin-Madison), Valérie Lauwers-Cancès (Service d’Épidémiologie, CHU Toulouse), Debbie A Lawlor (MRC Centre for Causal Analyses in Translational Epidemiology, University of Bristol), E Jeffrey Metter (National Institute on Aging), (Kuakini V Patel) (National Institute on Aging). Tanya Rantanen (University of Jyväskylä), Yves Rolland (University of Toulouse III), Hideo Sasaki (Hiroshima University School of Medicine), Holly Syddall (MRC Epidemiology Resource Centre, University of Southampton), Annweike van den Beld (Erasmus Medical Center), Bradley Willcox (Kuakini Health System), Andrew K Wills (MRC Unit for Lifelong Health and Ageing), and Jean Woo (Chinese University of Hong Kong).
Growing evidence indicates that simple objective measures of physical capability may be useful markers of future as well as current health.

Interest is increasing in these tests and their potential use as simple screening tools.

The inverse association between grip strength and all cause mortality was also seen in younger populations.

These measures may provide useful tools for identifying older people at higher risk of death.

**WHAT IS ALREADY KNOWN ON THIS TOPIC**

1. Katzmarzyk PT, Craig CL. Musculoskeletal fitness and risk of mortality.
2. Ahmad R, Bath PA. Identification of risk factors for 15-year mortality among elderly people.
3. Bergman H, Ferrucci L, Guralnik J, Hogan DB, Hummel S, Newman AB, et al. Prognostic value of usual gait speed in well-functioning older people.
4. Sasaki H, Kasagi F, Yamada M, Fujita S. Grip strength predicts cause-specific mortality in middle-aged and elderly persons.
5. Rosano C, Newman AB, Katz R, Hirsch CH, Kuller LH. Association between lower digit symbol substitution test score and slower gait and greater risk of mortality and of developing incident disability in well-functioning older adults.
6. Hardy SE, Studenski S, Chandler JM, Guralnik JM. Magnitude and patterns of decline in health and function in 1 year affect subsequent 5-year survival.
7. Rantanen T, Harris T, Leveille SG, Visser M, Foley D, Masaki K, et al. Muscle strength and body mass index as long-term predictors of mortality in initially healthy men.
8. Lukanen P, Heikkinen E, Kauppinen M. Muscle strength and mobility as predictors of survival in 75-84-year-old people.
9. Metter EJ, Talbot LA, Schrager M, Conwit R. Skeletal muscle strength as a predictor of all-cause mortality in healthy men.
10. Gordon LJ, Hsieh F, Simonsick EM, Guralnik JM, Harris TB. Grip strength, walking speed, chair rise, and standing balance performance and all cause mortality in older community dwelling populations.
11. The inverse association between grip strength and all cause mortality was also seen in younger populations.

**WHAT THIS STUDY ADDS**

Despite heterogeneity between studies, consistent evidence shows associations between grip strength, walking speed, chair rise, and standing balance performance and all cause mortality in older community dwelling populations.

**Ethical approval:** All included studies had received the relevant ethical approval. No additional approval was required for this review.

Data sharing: No additional data available.

**Conclusion:** Future research should investigate the potential uses of objective measures as simple screening tools for identifying older people at higher risk of mortality and disability.
46 Van den Beld AW, Visser TJ, Feelders RA, Grobbee DE, Lamberts SW. Thyroid hormone concentrations, disease, physical function, and mortality in elderly men. J Clin Endocrinol Metab 2005;90:6403-9.
47 DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177-88.
48 Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539-58.
49 Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557-60.
50 Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. Stat Med 1999;18:2693-708.
51 Higgins JP, Thompson SG. Controlling the risk of spurious findings from meta-regression. Stat Med 2004;23:1663-82.
52 Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. J Clin Epidemiol 2000;53:1119-29.
53 Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629-34.
54 Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994;50:1088-101.
55 Walston J, Hadley EC, Femucci L, Guralnik JM, Newman AB, Studenski SA, et al. Research agenda for frailty in older adults: toward a better understanding of physiology and etiology: summary from the American Geriatrics Society/National Institute on Aging Research Conference on Frailty in Older Adults. J Am Geriatr Soc 2006;54:991-1001.
56 Kuh D, Bassey EJ, Buttersworth S, Hardy R, Wadsworth MEJ. Grip strength, postural control, and functional leg power in a representative cohort of British men and women: associations with physical activity, health status, and socioeconomic conditions. J Gerontol Med Sci 2005;60A:224-31.
57 Kuh D, Hardy R, Buttersworth S, Okeil L, Richards M, Wadsworth MEJ, et al. Developmental origins of midlife physical performance: evidence from a British birth cohort. Am J Epidemiol 2006;164:110-21.
58 Kuh D, Hardy R, Buttersworth S, Okeil L, Wadsworth M, Cooper C, et al. Developmental origins of midlife grip strength: findings from a birth cohort study. J Gerontol Med Sci 2006;61:702-6.
59 Guralnik JM, Buttersworth S, Wadsworth MEJ, Kuh D. Childhood socioeconomic status predicts physical functioning a half century later. J Gerontol Med Sci 2006;61:694-701.
60 Sayer AA, Syddall HE, Gilbody HL, Dennison EM, Cooper C. Does sarcopenia originate in early life? Findings from the Hertfordshire cohort study. J Gerontol Med Sci 2004;59:M930-4.
61 Galobardes B, Smith GD, Lynch JW. Systematic review of the influence of childhood socioeconomic circumstances on risk for cardiovascular disease in adulthood. Ann Epidemiol 2006;16:91-104.
62 Turrell G, Lynch JW, Lelli C, Raghu Nathan T, Kaplan GA. Socioeconomic disadvantage in childhood and across the life course and all-cause mortality and physical function in adulthood: evidence from the Alameda County Study. J Epidemiol Community Health 2007;61:723-30.
63 Higgins JP. Commentary: heterogeneity in meta-analysis should be expected and appropriately quantified. Int J Epidemiol 2008;37:1158-60.
64 Ioannidis JP, Patsopoulos NA, Rothstein HR. Reasons or excuses for avoiding meta-analysis in forest plots. BMJ 2008;336:1413-5.
65 Rucker G, Schwarzer G, Carpenter JR, Schumacher M. Undue reliance on I2 in assessing heterogeneity may mislead. BMC Med Res Methodol 2008;8:79.
66 Guralnik JM, Branch LG, Cummings SR, Curb JD. Physical performance measures in aging research. J Gerontol Med Sci 1999;44:M141-6.

Accepted: 16 June 2010