Associations between device-measured physical activity and performance-based physical function outcomes in adults: a systematic review and meta-analysis

Joshua Culverhouse, Melvyn Hillsdon, Brad Metcalf, Michael Nunns, Rebecca Lear, Gemma Brailey, Richard Pulsford

ABSTRACT
This systematic review and meta-analysis aimed to examine the association between device-measured physical activity (PA) and performance-based measures of physical function (PF). Databases searched included CINAHL, Embase, MEDLINE/PubMed, SPORTDiscus, and Web of Science (last search conducted on November 11, 2022). Observational studies (cross-sectional or prospective) reporting associations between wearable device-measured PA and PF outcomes in non-clinical adults were eligible. Forty-two studies with a pooled sample of 27,276 participants were eligible, with 34 studies reporting a standardised regression coefficient ($\beta$) between at least one of four PA measures and one of six PF outcomes. All measures of PA were positively associated with all measures of PF, except for step count with grip strength. Largest associations were seen with lower-body PF tests; gait speed ($\beta$=0.11–0.26), walk tests ($\beta$=0.18–0.41), chair-rise test ($\beta$=0.10–0.26), balance ($\beta$=0.07–0.24) and Timed Up-and-Go ($\beta$=0.10–0.24) all p<0.01. Small or no association was seen with body mass index, body fat percentage, and lower-limb flexibility. The findings provide provisional support for the use of device-measured PA to remotely monitor people for risk of low PF. Prospective designs are needed to determine the direction of the relationship. Future studies should also expand beyond simple aggregate measures of PA to use different accelerometer values as there is evidence that the temporal distribution of activity is related to health and functional outcomes.

INTRODUCTION
The increase in healthy life expectancy has not kept pace with the increase in absolute life expectancy, resulting in a greater proportion of years lived in poor health. The UK’s latest figures suggest between 16 and 19 years of life will be lived in poor health for males and females, respectively. The societal and economic burden of these additional years lived in poor health is hard to quantify, but health and social care costs are rising, with the increased demand in later life in part due to loss of independence and disability.

Physical function (PF) is a broad concept that relates to the capacity of an individual to perform the physical tasks of everyday life required for independence, which reflects motor function and control, and components of physical fitness. Disablement models support the causal pathway from limitations in PF to disability, and loss of independence once these limitations interfere with activities of daily living. Relatively simple performance-based measures of PF such as grip strength, gait speed, chair rise tests, walk tests and balance can be strong predictors of adverse future health outcomes in older adults and late mid-life. Weak grip strength and slow gait speed are also characteristics of Fried’s frailty phenotype. Chair rise tests and grip strength have been recommended as screening and diagnostic tools for sarcopenia. However, PF assessments largely take place in clinical settings and only tend to occur when a person is attending a medical setting due to an adverse health event.

Declining PF is a common factor of ageing and, despite impairments typically being considered in older age, they can occur much earlier in mid-life (45–64 years). Depending on the point of intervention, declines in PF can potentially be prevented, retarded or reversed. However, identifying opportunities to intervene in mid-life relies on the ability to detect impairments in function prior to the point that reduced function results in presentation in medical settings. Remote health monitoring, through wearable devices, is one possible solution to early detection of presymptomatic and preclinical changes in
PF. Wearable devices for monitoring health outcomes are already being employed by both individuals, to track their own health through activity levels, and by clinicians as a method of early detection.\(^{18}\)

Wearable devices, such as accelerometers, have become increasingly popular for measuring physical activity (PA) in health research.\(^{19}\) There is strong evidence that structured PA, defined as any bodily movement produced by skeletal muscles that results in energy expenditure,\(^{20}\) and exercise interventions can improve or delay the loss of PF in older adults.\(^{21,22}\) Therefore, it is reasonable to consider that PA measures may be a potential proxy for PF. Prior to this, it is necessary to know what measures of PA are most strongly associated with, or even predictive of, PF. However, there is a paucity of evidence on the association between PA and PF in mid-life, when function is likely to be good but declining.

Systematic review level evidence of the associations between free-living PA and PF is limited, with reviews often focussing on interventions in people with reduced function.\(^{21,22}\) A meta-analysis has shown light intensity (LPA) and moderate-to-vigorous intensity PA (MVPA) to be associated with grip strength and chair rise tests;\(^{23}\) however, the focus of the meta-analysis was on the association between PA and strength rather than PF. In addition, included studies were limited to older adults, preventing insight into important associations of PA and PF in midlife. It also included a mix of studies of healthy populations as well as studies of specific clinical populations (eg, chronic obstructive pulmonary disease, diabetes, osteoarthritis), where the association between PA and PF might be confounded by these long-standing health conditions. No analysis of the differences in associations between studies of healthy and clinical populations was performed. To the authors’ knowledge, there are no systematic reviews of the association between PA and PF indicators such as gait speed, walk tests, balance or the Timed Up-and-Go test (TUG); and no reviews that examine the associations of PA and PF in both mid-life and older adulthood. This systematic review and meta-analysis examines associations between wearable, device-measured PA and a range of performance-based PF outcomes in non-clinical adults. The findings will inform the potential of remote monitoring of early declines in PF, which could inform the development of future screening programmes and interventions.

**METHODS**

The review was conducted according to the Conducting Systematic Reviews and Meta-Analyses of Observational Studies of Etiology (COSMOS-E)\(^{24}\) guidance and the Cochrane handbook\(^{25}\); and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.\(^{26}\) The protocol was registered in the International Prospective Register of Systematic Reviews—PROSPERO (CRD42021282861).

**Search strategy**

Systematic literature searches were conducted in PubMed (including Ovid MEDLINE, HMIC and Embase), EBSCO-host (including CINAHL and SPORTDiscus) and Web of Science for studies published between database inception and 15th June 2021; a top-up search was performed on 11 November 2022. The search strategy included keywords related to PA, device-based measures of PA, PF outcomes and observational study designs (online supplemental file 1). In addition, supplementary searches were performed through bibliography screening of included papers to identify any other potentially relevant publications.

**Study selection**

Inclusion was determined by two independent reviewers (JC+GB or RL). Disagreements were resolved by discussion with the third author (GB or RL), if required. Study selection was completed in two phases: title and abstract screening was performed to exclude clearly irrelevant studies, after which full texts were screened. If two or more studies reported similar associations for the same cohort, we included the study with highest quality score or largest sample size, respectively.

**Eligibility criteria**

**Population**

Participants were adults (≥18 years of age) recruited from non-clinical, community-dwelling populations. Studies of adults recruited specifically due to the presence of, or expected progression to, a disease or other clinical condition were excluded. These inclusion criteria allow for generalisation to the general population, including those in mid-life; these assertions cannot be made from studies of clinical populations of solely older adults.

**Exposure**

Studies reporting continuous wear data from remote wearable, device-based measures of PA were included. Depending on device, this included studies that advised participants to wear the device for 24 hours continuously, or to only remove the device during sleep and water-based activity. Reported PA metrics were total step count, total volume of PA (TPA), LPA and MVPA, classified using published cut-points or proprietary algorithms. Studies which collapsed continuous PA data were contacted to try to obtain the continuous association. We excluded studies that exclusively reported estimates of sedentary behaviour.

**Outcome**

Studies reporting performance-based PF instruments, adopted by clinicians and researchers, were included. These include; grip strength, gait speed, chair rise tests, walk tests, balance tests or composite assessments of these measures.\(^{14,27,28}\)

**Study design**

The review included observational studies (both cross-sectional and prospective designs), which reported...
associations between the exposures and outcomes. Experimental studies and randomised controlled trials were excluded.

No restrictions were placed on country or date. Only full texts, in English, were included.

Data extraction
Two authors (JC+RL) independently extracted the following data from included studies: (1) author, study year and country of origin; (2) cohort and study design; (3) sample size and sex distribution; (4) age of study participants; (5) device used for PA measurement and metrics reported; (6) test used for assessing PF and metrics reported; (7) statistical analyses undertaken including and covariates included and (8); key results for the association between PA and PF. Discrepancies in extracted data were resolved by discussion with a third author (GB), if required.

Assessment of study and evidence quality
Two authors (JC+RL) independently assessed the quality of included studies using an adapted version of the National Institutes of Health Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies (online supplemental file 2). Scores were given ranging from 0 to 12, with higher scores indicating higher quality. Discrepancies in quality assessment were resolved by discussion with a third author (GB), if required. The continuous quality rating scores were used in sensitivity analyses.

Statistical analysis
The required association statistic was the standardised regression coefficient (β) and SE, see detailed explanation of β coefficient below. Using the β coefficient allowed for synthesis across different metrics and units for both PA and PF variables. If only a partial correlation coefficient was obtainable, this was used as an approximation the β coefficient, with sensitivity analysis performed to ensure these coefficients would not bias the pooled effect.29

Some PF outcomes have slight different measurement protocols, and these are grouped together in this review as follows; the chair-rise test outcome includes the 30 s and the five-repetition variants; gait speed includes any protocol measuring normal/usual or maximal gait speed over a distance ≤10 m; grip strength includes any protocol using a hand dynamometer to obtain maximal grip strength; walk tests included the 6 min walk test and 400 m walk test, or any variant covering a similar time or distance in different units; the TUG test includes both the 8-foot and 3 m variations; and balance includes any continuous measure of tandem, semi-tandem or single-leg stance, with eyes closed or open. Where composite scores of the above measures were reported for an overall PF score, we sought to obtain the associations for the individual components.

The adjusted β coefficients were extracted from included papers, or obtained from converting the unstandardised regression coefficient (b) where possible using the following equations: \[ \beta = \frac{\text{SE}(b)}{\text{SE}(\beta)} \]

where SDx is the SD of the PA exposure and SDy is the SD of the PF outcome.30 If the SDx or SDy was reported in two subgroups and needed to be combined the following equation was used to obtain the full sample SD:

\[ \text{SD}_{\text{full sample}} = \sqrt{\frac{(n_1-1)\text{SD}_1^2+(n_2-1)\text{SD}_2^2+\frac{n_1n_2}{n_1+n_2}(M_1-M_2)^2}{n_1+n_2-1}} \]

where \( n_1 \) and \( n_2 \) are the sample sizes of the two subgroups, \( \text{SD}_1 \) and \( \text{SD}_2 \) are the subgroup SDs, and \( M_1 \) and \( M_2 \) are the subgroup means.30 If SE was not reported, it was calculated from the 95% CIs using the following equation:

\[ \text{SE} = \frac{\text{upper limit} - \text{lower limit}}{3.92} \]

where the upper limit and lower limit refer to the 95% CI of the effect size.31 In cases where the partial correlation was used, the following equation was used to calculate the SE of the partial correlation:

\[ \text{SE} = \frac{1-r^2}{\sqrt{n-2}} \]

where \( r \) is the partial correlation coefficient and \( n \) is the sample size.31 If a study reported associations separately for two subgroups (eg, males and females) these were combined using the following equations to provide a composite effect size:

\[ \beta_p = \frac{(W_1\beta_1 + W_2\beta_2)}{(W_1 + W_2)} \]

\[ \text{SE}(\beta_p) = \sqrt{\frac{1}{W_1+W_2}} \]

where \( \beta_1 \) and \( \beta_2 \) are the β coefficients for the two subgroups, and SE(\( \beta_1 \)) and SE(\( \beta_2 \)) are the respective SEs. The weightings for the two subgroups are \( W_1 = 1/\text{SE}(\beta_1) \) \(^2\) and \( W_2 = 1/\text{SE}(\beta_2) \) \(^2\).31

Where required, we contacted authors to request the β coefficient adjusted for age+sex, or additional unpublished data to allow us to estimate the β coefficient from the effect size published in the paper. If authors had measured additional PA or PF outcomes but not reported these associations, these were also requested. β coefficients were inverted for PF outcomes where a lower score indicated better function, so that all positive effects in this review indicate better/higher PF.

Meta-analyses were performed to obtain a pooled estimate of individual β coefficients for associations between the reported PA measures and PF outcomes, visualised as forest plots. Ideally, included effect sizes would be adjusted for the same covariates,31 32 however, due to varying adjustment models across papers, the included estimates were extracted from the following order of models: (1) age+sex and (2) age, sex+additional factors. We used random-effects models to account for both between and within study variance, with inverse variance as the weighting method. Statistical heterogeneity was estimated using the I² analysis. An I² (the variation across studies due to heterogeneity rather than chance)
of <40% was considered low heterogeneity and an $I^2$ of >75% was considered high heterogeneity. Heterogeneity, along with the number of studies within each meta-analysis should be considered when interpreting the pooled effects. Where possible (≥10 studies in the meta-analysis) meta-regressions were run to examine the individual effects of sex (percentage female), age, quality assessment and study sample size (n) on the associations.

Sensitivity analyses
Leave-one-out sensitivity was performed on each meta-analysis to explore the influence of individual studies on the overall pooled effect. In addition, for meta-analyses with ≥10 studies, a visual and statistical evaluation of publication bias was performed using funnel plots and Egger’s regression tests (p<0.05 indicated publication bias). For the purpose of quantifying the magnitude of the pooled effect size, the following values were used: 0.10–0.19=small, 0.20–0.29=medium and ≥0.30=large. So as not to entirely exclude them from the review, studies for which a β coefficient was not obtained were included in a vote count summary and the directions of associations compared with those studies included in the meta-analysis via $\chi^2$ test. All analyses were performed in Stata V.17 (StataCorp. 2021. Stata Statistical Software: Release 17. College Station, StataCorp).

RESULTS
Search and study selection results
The original and top-up database searches identified 2741 articles after duplicates were removed, of which 2533 were excluded based on title and abstract screening. Two hundred and seven full-text articles were reviewed, 43 of which fulfilled the inclusion criteria. Two studies, by the same author, used data from the same pool of participants, the study with the larger sample size and greater number of reported associations was chosen for inclusion. Resulting in a total of 42 included publications (figure 1).

![Figure 1](http://bmjpublichealth.bmj.com/)

**Figure 1** PRISMA flow diagram showing the screening process and the search results. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.
Study characteristics

The 42 included studies represented N=27 276 participants (range: n=64–4702), with an average mean sample age of 70.3 years (range: 46–90 years) (table 1). Study samples were on average 63.6% female. Three studies were prospective36–38 and the other 39 were cross-sectional44–75. Most studies used accelerometers to measure PA (k=39), with one study using a pedometer,68 and two using the Actiheart combined accelerometer and heart rate sensor.4370 Device locations across studies were as follows; hip/waist (k=27), thigh (k=3), wrist (k=3), other (k=9). Studies reported the following PA dimensions; MVPA (k=31), LPA (k=17), TPA (k=15) and total step count (k=14). A range of accelerometer cut-points were used for classifying LPA and MVPA across studies, the most common non-proprietar classifi-
cations were Troiano76(k=6) and Freedson77(k=5) (online supplemental file 3, table 2).

Studies also reported the following PF outcomes; gait speed (k=27), handgrip strength (k=24), chair rise tests (k=17), TUG (k=15), balance (k=12), endurance walk tests (k=10) and composite PF tests (k=6) (online supplemental file 3). There was an insufficient number of studies employing composite measures of PF for these to be pooled; only one of the four studies that did report composite measures was excluded from meta-analyses, where the associations of individual measures within the composite score were not reported or obtainable.47

Of the 42 studies identified for inclusion in this review, a standardised regression coefficient (β), adjusted for at least age+sex, was obtained for 34 studies and thus were included in pooled analyses. Authors of 14 of these studies provided either additional data to allow the estimation of the β coefficient, or effect sizes for additional associations that were not reported in the original paper. The variations in PA exposures and PF outcomes reported across included studies prevented the computation of a single overall effect size. Instead, multiple pooled analyses (n=24) were performed for each combination of PA and PF measure, as described above. Overall, the 34 studies included in meta-analyses represent 22 774 participants (range: 64–4702), with a mean sample age of 69.3 (range: 46–83.5) and comprising 63.4% females. Two studies reported prospective associ-
The limited number of studies reporting some of the associations meant that only 6 of the meta-analyses contained ≥10 studies, and therefore, meta-regressions and Egger’s test were only performed on these 6. Due to an unbalanced number of studies across the device locations (27 studies adopted waist/hip), we refrained from conducting subgroup analysis on this factor. All extracted data are provided in online supplemental file 3.

Methodological quality

For all 42 included studies, the mean quality assessment rating was 8.1±1.2 (range: 3–13). The 34 studies included in meta-analyses, the mean rating was 8.2±1.2 (range: 6–13). Study design (only 4 studies were prospective), sample size justification and participation rate of eligible persons were the most problematic domains of study quality (online supplemental file 4).

RESULTS OF META-ANALYSES

Gait speed

There were positive associations for each of the PA measures with gait speed (figure 2). The magnitudes of association varied between PA measures, with medium strength associations seen in MVPA (β=0.26, p<0.001) and step count (β=0.26, p<0.001), and small associations seen with TPA (β=0.17, p<0.001) and LPA (β=0.11, p<0.001). Statistical heterogeneity was high step count, and moderate for TPA, LPA and MVPA. Meta-regressions for age, sex, sample size and quality assessment score for TPA and MVPA were non-significant (online supplemental file 6). Egger’s test for TPA and MVPA was non-significant (online supplemental file 7).

Chair rise tests

All PA measures were positively associated with chair rise tests (figure 3). The magnitudes of association varied between PA measures; step count was the largest but with wide CIs (β=0.26 (0.09 to 0.41), p=0.003), followed by MVPA (β=0.18, p<0.001), TPA (β=0.14, p<0.001) and LPA (β=0.10, p<0.001). Heterogeneity was high MVPA and step count, moderate for TPA and low for LPA. Meta-regressions for MVPA were non-significant (online supplemental files 67). Egger’s test for MVPA was non-significant.

Balance

There were a limited number of studies reporting associations with balance. All measures of PA were positively associated with balance (figure 4). The largest associations were seen with step count (β=0.24, p<0.003), followed by MVPA (β=0.15, p<0.001) and TPA (β=0.12, p<0.001); the smallest association was with LPA (β=0.07, p<0.036). Heterogeneity was moderate for MVPA and low TPA, LPA and step count.

Walk tests

Similar to balance, there were a limited number of studies reporting associations with walk tests. All measures of PA were positively associated with walk tests (figure 5). The magnitudes were largest with step count (β=0.41, p<0.001) and MVPA (β=0.35, p<0.001); followed by LPA (β=0.19, p<0.001) and TPA (β=0.18, p<0.001). Heterogeneity was high for TPA and step count, moderate for MVPA and low for LPA.

Timed Up-and-Go

All measures of PA were associated with the Timed Up-and-Go test (figure 6). The magnitudes were largest with MVPA (β=0.24, p<0.001) and step count (β=0.24, p<0.001); followed by TPA (β=0.19, p<0.001) and LPA.
Table 1  Characteristics of articles assessing the association between device-measured physical activity metrics with performance-based physical function outcomes in adults

<table>
<thead>
<tr>
<th>Author (year), country</th>
<th>Sample (n)</th>
<th>Age (F%)</th>
<th>PA measures</th>
<th>PF measures</th>
<th>Adjustments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>LPA</td>
<td>MVPA</td>
<td>Steps</td>
</tr>
<tr>
<td>Adachi (2018), JP</td>
<td>308</td>
<td>79.9 (3.6)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Aggio (2016),* GB</td>
<td>1286</td>
<td>78.2 (4.5)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Aoyagi (2009), JP</td>
<td>170</td>
<td>72.6 (4.6)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Cooper (2015), GB</td>
<td>1727</td>
<td>63.3 (1.1)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Cooper (2020), GB</td>
<td>4702</td>
<td>46 (0)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Davis (2014), GB</td>
<td>217</td>
<td>78.1 (5.8)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Duck (2019), USA</td>
<td>99</td>
<td>74 (6.5)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Gobbo (2020), BR</td>
<td>68</td>
<td>69.4 (6.5)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Hall (2017), USA</td>
<td>775</td>
<td>62.1 (N/R)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Hsueh (2020), TW</td>
<td>127</td>
<td>70.8 (5.3)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Izawa (2017), JP</td>
<td>290</td>
<td>74.5 (N/R)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Jantunen (2017), FI</td>
<td>695</td>
<td>70.7 (2.7)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Johanssøn (2021), NO</td>
<td>3653</td>
<td>68.5 (5.9)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Kim (2015), JP</td>
<td>207</td>
<td>83.5 (2.6)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Kruger (2016), SA</td>
<td>247</td>
<td>57.0 (10.2)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Lai (2020), TW</td>
<td>118</td>
<td>70.0 (5.0)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Lerma (2018), USA</td>
<td>91</td>
<td>70.7 (10.2)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Lohne-Seiler (2016), NO</td>
<td>161</td>
<td>72.8 (5.1)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Manas (2019)*, ES</td>
<td>771</td>
<td>76.8 (4.9)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Meier (2020), USA</td>
<td>304</td>
<td>72.8 (5.8)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Mendham (2021), SA</td>
<td>111</td>
<td>67 (64, 71)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Mizumoto (2015)*, JP</td>
<td>201</td>
<td>79.7 (3.8)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Nagai (2018)*, JP</td>
<td>886</td>
<td>73.6 (7.0)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Oguma (2017)*, JP</td>
<td>155</td>
<td>90.2 (1.4)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Osuka (2015), JP</td>
<td>802</td>
<td>72.5 (5.9)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Pina (2021), SA+GB</td>
<td>288</td>
<td>68.5 (N/R)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Reid (2016)*, AU</td>
<td>602</td>
<td>58.1 (10.0)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Ribeiro (2020), BR</td>
<td>230</td>
<td>66 (63, 71)</td>
<td>70.4</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Author (year), country Cohort</th>
<th>Design</th>
<th>Sample (n)</th>
<th>Age</th>
<th>Sex (F%)</th>
<th>PA measures</th>
<th>PF measures</th>
<th>Adjustments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LPA MVPA Steps TPA</td>
<td>Bal. Chair Gait HGS TUG Walk Comp. Age Sex Add.</td>
<td></td>
</tr>
<tr>
<td>Rojer (2018), NL(^{78})</td>
<td>N/A</td>
<td>236</td>
<td>66.9 (N/R)</td>
<td>64.8</td>
<td>✓ ✓</td>
<td>✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>Sanchez-Sanchez (2019), ES(^{57})</td>
<td>TSHA</td>
<td>497</td>
<td>78.1 (5.7)</td>
<td>54.3</td>
<td>✓ ✓</td>
<td>✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>Santos (2012), PT(^{58})</td>
<td>N/A</td>
<td>312</td>
<td>74.3 (6.6)</td>
<td>62.5</td>
<td>✓ ✓</td>
<td>✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>Savikangas (2020), FI(^{59})</td>
<td>Password CS</td>
<td>293</td>
<td>74.4 (3.8)</td>
<td>58.4</td>
<td>✓ ✓</td>
<td>✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>Schrack (2019), USA(^{60})</td>
<td>BLSA</td>
<td>680</td>
<td>67.9 (13.2)</td>
<td>49.9</td>
<td>✓ ✓</td>
<td>✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>Spartan (2019), USA(^{62})</td>
<td>FOS</td>
<td>1352</td>
<td>68.6 (7.5)</td>
<td>54</td>
<td>✓ ✓</td>
<td>✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>Thiebaud (2020),* JP(^{63})</td>
<td>N/A</td>
<td>86</td>
<td>67 (7)</td>
<td>100</td>
<td>✓ ✓</td>
<td>✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>van der Velde (2017), NL(^{84})</td>
<td>Maastricht CS</td>
<td>1962</td>
<td>59.7 (8.2)</td>
<td>48.6</td>
<td>✓ ✓</td>
<td>✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>Ward-Ritacco (2014), USA(^{65})</td>
<td>N/A</td>
<td>64</td>
<td>58.6 (3.6)</td>
<td>100</td>
<td>✓ ✓</td>
<td>✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>Ward-Ritacco (2020), USA(^{66})</td>
<td>N/A</td>
<td>80</td>
<td>52.6 (6.1)</td>
<td>100</td>
<td>✓ ✓</td>
<td>✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>Westbury (2018), GB(^{67})</td>
<td>HSS</td>
<td>131</td>
<td>78.8 (2.4)</td>
<td>75.6</td>
<td>✓ ✓</td>
<td>✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>Yamada (2011),* JP(^{58})</td>
<td>N/A</td>
<td>515</td>
<td>77.0 (7.2)</td>
<td>67.5</td>
<td>✓ ✓</td>
<td>✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>Yasunaga (2017), JP(^{69})</td>
<td>N/A</td>
<td>287</td>
<td>74.4 (5.2)</td>
<td>37.3</td>
<td>✓ ✓</td>
<td>✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>Yerrakalva (2022), UK(^{98})</td>
<td>EPIC-Norfolk PR</td>
<td>1488</td>
<td>69.9 (6.0)</td>
<td>54.4</td>
<td>✓ ✓</td>
<td>✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓</td>
<td>✓ ✓</td>
</tr>
</tbody>
</table>

Age in years is presented as mean (SD) or median (IQR). Sex distribution is presented as the percentage of females within the study sample. *Asterisk denotes not included in meta-analyses.
*Add, additional; AU, Australia; Bal, balance test; BLSA, Baltimore Longitudinal Study of Aging; BR, Brazil; BRHS, British Regional Heart Study; Chair, chair rise test; Comp, composite measure; CS, cross-sectional; ES, Spain; FI, Finland; FOS, Framingham Offspring Study; Gait, gait speed; GB, Great Britain; HBCS, Helsinki Birth Cohort Study; HGS, handgrip strength; HSS, Hertfordshire Sarcopenia Study; JP, Japan; LPA, light intensity physical activity; MURDOCK, The Measurement to Understand the Reclassification of Disease Of Cabarrus/Kannapolis; MVPA, moderate-to-vigorous physical activity; N/A, not applicable; NL, Netherlands; NO, Norway; N/R, not reported; NSHD, National Survey of Health and Development; OPAL, Older People and Active Living; PA, physical activity; PF, physical function; PPAOI, Population-Based and Inspiring Potential Activity for Old-Old Inhabitants; PR, prospective; PT, Portugal; PURE, Prospective Urban and Rural Epidemiological study; SA, South Africa; SPPB, short physical performance battery; Steps, average or total step count; TOOTH, The Tokyo Oldest-Old Survey of Total Health; TPA, total physical activity; TSHA, Toledo Study of Healthy Aging; TUG, Timed Up-and-Go; TW, Taiwan.
Handgrip strength showed small, positive associations with TPA ($\beta=0.07$, $p<0.001$), LPA ($\beta=0.05$, $p=0.002$) and MVPA ($\beta=0.07$, $p<0.001$), but had no association with step count ($\beta=0.02$, $p<0.406$) (figure 7). Heterogeneity was moderate for TPA, LPA and MVPA, and low for step count. Egger’s test for TPA, LPA and MVPA was non-significant (online supplemental file 7). As detailed in the methods, effect sizes from studies reporting subgroups were pooled, except in two instances for grip strength, where the effects were in the opposite direction in each subgroup (figure 7C,D).

**Sensitivity analyses**

The results of the ‘leave-one-out’ sensitivity analyses suggests that, in general, our estimates of associations were robust to sensitivity analyses. The $\beta$ coefficients did not change more than; $-0.04$ to $+0.03$ for balance, $-0.04$ to $+0.08$ for chair rise tests, $-0.02$ to $+0.04$ for gait speed, $-0.03$ to $+0.03$ for grip, $-0.03$ to $+0.05$ for TUG and $-0.07$ to $+0.12$ for walk tests. Importantly, $\beta$ coefficients from the ‘leave-one-out’ analyses were always within the 95% CIs of the original estimates derived from ‘all studies’ (online supplemental file 5). Even for the three associations that became non-significant, the magnitude of the change in the $\beta$ coefficient was very small (eg, $\beta$ coefficients of 0.12, 0.07 and 0.41 fell no more than 0.04). The sample study sizes for these associations were 3, 3 and 4, respectively, and we are impacted when the studies with large sample sizes were removed; therefore, we suggest caution when interpreting the associations with smaller numbers of studies.

All meta-regressions were non-significant. Bubble plots suggested that some meta-regressions might have studies with high leverage. According to Borenstein et al, there are no current methods in which meta-regression deals with ‘high leverage’. Leverage was calculated for each study within each meta-regression, and the formula reported in Borenstein et al was used to identify studies with ‘high’ leverage. In the absence of an optimal process to deal with high leverage, analysis was re-run excluding any studies with large sample sizes.

**Table 2** Vote counting across all reported associations of included studies

<table>
<thead>
<tr>
<th>Total PA and gait speed (k = 11)</th>
<th>Light PA and gait speed (k = 9)</th>
<th>MVPA and gait speed (k = 16)</th>
<th>Step count and gait speed (k = 7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author &amp; Year</td>
<td>n</td>
<td>Std. Beta [95% CI]</td>
<td>Std. Beta [95% CI]</td>
</tr>
<tr>
<td>Alabili 2018</td>
<td>313</td>
<td>0.13 (0.08, 0.22)</td>
<td>0.04 (0.00, 0.18)</td>
</tr>
<tr>
<td>Anayi 2009</td>
<td>170</td>
<td>0.16 (0.02, 0.32)</td>
<td>0.10 (0.00, 0.21)</td>
</tr>
<tr>
<td>Davis 2014</td>
<td>217</td>
<td>0.23 (0.10, 0.36)</td>
<td>0.11 (0.06, 0.17)</td>
</tr>
<tr>
<td>Gabbs 2020</td>
<td>151</td>
<td>0.13 (0.08, 0.25)</td>
<td>0.11 (0.06, 0.16)</td>
</tr>
<tr>
<td>Hall 2017</td>
<td>460</td>
<td>0.09 (0.00, 0.18)</td>
<td>0.16 (0.00, 0.27)</td>
</tr>
<tr>
<td>Shaub 2020</td>
<td>324</td>
<td>0.32 (0.14, 0.48)</td>
<td>0.18 (0.07, 0.28)</td>
</tr>
<tr>
<td>Story 2017</td>
<td>296</td>
<td>0.09 (0.03, 0.16)</td>
<td>0.11 (0.06, 0.17)</td>
</tr>
<tr>
<td>Kim 2015</td>
<td>267</td>
<td>0.19 (0.12, 0.28)</td>
<td>0.21 (0.10, 0.32)</td>
</tr>
<tr>
<td>Kang 2016</td>
<td>247</td>
<td>0.23 (0.10, 0.36)</td>
<td>0.11 (0.06, 0.16)</td>
</tr>
<tr>
<td>Littman 2018</td>
<td>91</td>
<td>0.14 (0.08, 0.20)</td>
<td>0.20 (0.11, 0.29)</td>
</tr>
<tr>
<td>Nunt 2020</td>
<td>206</td>
<td>0.35 (0.16, 0.53)</td>
<td>0.22 (0.10, 0.34)</td>
</tr>
<tr>
<td>Mondhan 2012</td>
<td>111</td>
<td>0.21 (0.10, 0.32)</td>
<td>0.19 (0.08, 0.30)</td>
</tr>
<tr>
<td>Pinta 2021</td>
<td>288</td>
<td>0.20 (0.09, 0.31)</td>
<td>0.21 (0.10, 0.32)</td>
</tr>
<tr>
<td>Reiter 2018</td>
<td>236</td>
<td>0.09 (0.03, 0.16)</td>
<td>0.11 (0.06, 0.16)</td>
</tr>
<tr>
<td>Sandhu-Sinha 2019</td>
<td>497</td>
<td>0.10 (0.03, 0.17)</td>
<td>0.10 (0.00, 0.20)</td>
</tr>
<tr>
<td>Sankaran 2020</td>
<td>253</td>
<td>0.20 (0.10, 0.30)</td>
<td>0.22 (0.11, 0.30)</td>
</tr>
<tr>
<td>Schick 2019</td>
<td>150</td>
<td>0.18 (0.07, 0.29)</td>
<td>0.19 (0.08, 0.30)</td>
</tr>
<tr>
<td>Spottis 2019</td>
<td>132</td>
<td>0.29 (0.13, 0.44)</td>
<td>0.21 (0.10, 0.32)</td>
</tr>
<tr>
<td>Story 2018</td>
<td>131</td>
<td>0.20 (0.10, 0.30)</td>
<td>0.11 (0.06, 0.16)</td>
</tr>
<tr>
<td>Youyang 2017</td>
<td>287</td>
<td>0.21 (0.10, 0.30)</td>
<td>0.10 (0.00, 0.15)</td>
</tr>
<tr>
<td>Yendrali 2012</td>
<td>1419</td>
<td>0.20 (0.10, 0.30)</td>
<td>0.17 (0.10, 0.24)</td>
</tr>
</tbody>
</table>

Figure 2 Forest plots showing the associations between physical activity measures and gait speed. k, number of studies per meta-analysis; MVPA, moderate-to-vigorous physical activity; N, sample size; PA, physical activity.
studies with high leverage. All meta-regressions remained non-significant.

**Vote count summary**

Of the 42 studies that met the inclusion criteria, β coefficients were not obtainable for eight studies; and therefore, these were not included in the meta-analysis. To avoid completely omitting these studies from the review and to acknowledge any potential bias, a vote count summary is provided with all studies and subgroup vote count comparing those studies included in the meta-analysis and those excluded (table 2).

**DISCUSSION**

The aim of this systematic review was to examine associations between wearable, device-measured PA and a...
range of performance-based PF outcomes in community-dwelling adults. Forty-two studies met the inclusion criteria and 34 studies provided suitable data for meta-analyses, across 24 different associations between PA and PF. All measures of PA were positively associated with all measures of PF, except for step count with grip strength. In general, the more physically active people were the better their PF. Associations were generally higher with lower-body PF tests, particularly gait speed, chair rises and walk tests. Within each measure of PF, the associations with either MVPA or step count were generally larger than compared with LPA or TPA. The associations of PA with chair-rise tests and grip strength were similar to those reported in a previous meta-analysis.25 Direct comparisons between this review and that of Ramsey et al23 are not possible due to this review excluding studies that recruited studies based on the presence of a specific clinical condition. Our decision was taken to increase the external validity of the results but also because the expected association between PA and PF would be condition specific and there were too few studies for each specific condition to carry out analysis separately, comparing studies in healthy populations to each clinical condition. Our inclusion of all adults (not just older

<table>
<thead>
<tr>
<th>Author &amp; Year</th>
<th>Total PA and TUG (k = 3)</th>
<th>Light PA and TUG (k = 4)</th>
<th>MVPA and TUG (k = 9)</th>
<th>Step count and TUG (k = 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Std. Beta [95% CI]</td>
<td>Std. Beta [95% CI]</td>
<td>Std. Beta [95% CI]</td>
<td>Std. Beta [95% CI]</td>
</tr>
<tr>
<td>Closer 2015</td>
<td>0.19 [0.10, 0.32]</td>
<td>0.09 [0.00, 0.18]</td>
<td>0.15 [0.10, 0.20]</td>
<td>0.15 [0.00, 0.30]</td>
</tr>
<tr>
<td>Gubbo 2020</td>
<td>0.15 [0.06, 0.33]</td>
<td>0.10 [0.00, 0.20]</td>
<td>0.05 [0.00, 0.10]</td>
<td>0.24 [0.07, 0.40]</td>
</tr>
<tr>
<td>Rea 2018</td>
<td>0.12 [0.03, 0.32]</td>
<td>0.08 [0.00, 0.18]</td>
<td>0.03 [0.00, 0.06]</td>
<td>0.35 [0.10, 0.60]</td>
</tr>
<tr>
<td>Van der Vliet 1992</td>
<td>0.13 [0.04, 0.30]</td>
<td>0.08 [0.00, 0.15]</td>
<td>0.03 [0.00, 0.06]</td>
<td>0.32 [0.08, 0.55]</td>
</tr>
<tr>
<td>Ward-Riggs 2014</td>
<td>0.10 [0.00, 0.20]</td>
<td>0.05 [0.00, 0.10]</td>
<td>0.02 [0.00, 0.04]</td>
<td>0.14 [0.05, 0.35]</td>
</tr>
<tr>
<td>Ward-Riggs 2020</td>
<td>0.09 [0.00, 0.21]</td>
<td>0.05 [0.00, 0.10]</td>
<td>0.02 [0.00, 0.04]</td>
<td>0.14 [0.05, 0.35]</td>
</tr>
</tbody>
</table>

Pooled Estimate

| p-value       | 0.081          | 0.081          | 0.081          | 0.081          |
| Homogeneity (I^2) | 50%           | 50%           | 50%           | 50%           |

**Figure 5** Forest plots showing the associations between physical activity measures and walk tests. k, number of studies per meta-analysis; MVPA, moderate-to-vigorous physical activity; N, sample size; PA, physical activity.

**Figure 6** Forest plots showing the associations between physical activity measures and the timed up-and-go test. k, number of studies per meta-analysis; MVPA, moderate-to-vigorous physical activity; N, sample size; PA, physical activity.
adults) adds to the previous review in this area. The number of studies within many of the meta-analyses did not allow for meta-regression; though in the six which did, there was no apparent effect of sample age on the observed associations.

The differences observed in the magnitude of associations between PA and specific measures of PF may be explained, at least in part, by the specificity of exercise. For example, grip strength, a general measure of muscular strength, would be expected to improve as a result of resistance type exercises rather than ambulatory activity. Therefore, measuring PA with devices that largely capture ambulatory behaviour, not resistance exercise, would likely underestimate the association between PA and grip strength, especially in participants undertaking a higher level of resistance exercise. Similarly, measures of PF more related to ambulation (eg, gait speed and walk tests) would be expected to produce larger associations with device-based measures of PA that mainly represent ambulatory activity. Although device-based measures of PA overcome recall and social desirability biases associated with self-report measures, they do not adequately capture strength or resistance-based activities. 80 81

The reliance on single thresholds of acceleration to define activity intensity categories, for all study participants, can lead to the misclassification of time spent in different intensities of activity. The approach assumes that a given value of acceleration represents the same intensity of PA for all individuals regardless of their fitness. 82 For example, if two people (one low fit and one high fit) were walking at the same speed on a treadmill the accelerometer would record approximately the same level of acceleration assuming both people had similar stride lengths. However, the less fit person would be exercising at a higher relative intensity (% of maximum) than the fitter person. Consequently, in less fit participants the single threshold method would lead to an underestimate of time spent in MVPA—misclassified as LPA, and for fitter participants an overestimate of time in MVPA. Further, the most common thresholds used by included studies were derived in calibration studies of young adults (<30 year old) which is unlikely to generalise to older populations with lower fitness levels. 76 77  Our findings show that more time at higher acceleration values is associated with better function, but it is difficult to know what level of relative intensity these thresholds represent in the populations being studied, even though in general higher accelerations are correlated with higher V02 levels. 83 84

The reporting of PA volume alone ignores other dimensions of activity and the temporal distribution, including event-based outcomes of free-living behaviour. 85 This is despite evidence that two people with the same volume of activity, accumulated in different patterns will vary in their risk of mortality, 84 and that patterns (eg, number and duration of bouts of activity) may also be associated with PF. 86 Developments in data processing allow for additional PA metrics to be derived from accelerometers, metrics that better reflect the frequency, duration, intensity and volume of PA, as well as how the PA was...

**Figure 7** Forest plots showing the associations between physical activity measures and handgrip strength. k, number of studies per meta-analysis; MVPA, moderate-to-vigorous physical activity; N, sample size; PA, physical activity.
accumulated within and between days. It is also possible to estimate specific movements, for example, sit-to-stand posture transitions, which have not been widely reported in this literature, but which might be more relevant to certain measures of PF (eg, chair rise tests and TUG). The ability to detect more specific types of activity, such as postural transitions, particularly the ‘quality’ of these activities (eg, duration, velocity and power), holds promise for better understanding of links between specific device-measures of PA and PF. This in turns raises the potential for remote monitoring of PF in free-living settings rather than being reliant on clinic-based measures. It is already documented that clinic and laboratory measures of PF do not capture the same broad dynamic of free-living PF.

Only two of the studies included in this meta-analysis reported prospective associations, meaning the direction of causation cannot be determined. It is logical that the relationship is somewhat bidirectional, given the likely cyclical relationship between impaired function, disability and reduced PA. Prospective associations between PA in mid-life and preserved PF at follow-up have been demonstrated, although with self-report measures of the exposure and outcome. Further examination of these prospective associations should be performed with device-measured PA, to avoid the biases associated with self-report. The association between PA and PF, or even prevalence of impairment, in mid-life is poorly understood, despite the potential for early screening and intervention. The WHO specifically refers to reduced gait speed and muscle strength as early markers for declines in intrinsic capacity, and emphasises the need for early detection to prevent these declines in capacity. Prospective studies with measures of both PA and function collected in mid-life are required to better understand whether device-based measures of PA in mid-life are associated with the risk of low function later in life.

Strengths and limitations

To the authors’ knowledge, this is the first meta-analysis of the associations between device measured free-living PA and PF in observational studies of adults from mid-life to older adulthood. Specifically, this is the first review to examine pooled associations of PA with gait speed, walk tests, balance and TUG. We build on previous analyses of associations with grip strength and chair rise tests by focussing on non-clinical populations where associations are less likely to be confounded by the presence of health conditions. The multiple dimensions of PA and broad range of performance-based PF outcomes provide a comprehensive review of the relative magnitudes of PA associations between PF measures, and the associations of different PA dimensions within those measures. The inclusion of studies employing device-based measures removes the impact of error and bias associated with self-report measures from pooled effects. However, we note that the number of studies within certain analyses was low, contributing to considerable heterogeneity, and an inability to explore potential effect modifiers using meta-regression. As such we interpret the reported pooled effects of these meta-analyses with a degree of caution. Adopting the standardised regression coefficient as the effect size for the pooled analysis allowed for the inclusion of studies employing different statistical inference methods, measurement methods and descriptive statistics. However, only evidence of an association should be interpreted from a significant meta-analysis, as the strength of associations is not comparable across standardised regression output. The minimum adjustment model for inclusion was age+sex may have meant some important confounding factors were overlooked; however, it allowed inclusion of a greater number of studies than if the criteria had been stricter. We could not include eight studies within meta-analyses, however, the proportion of these studies reporting positive associations between PA and PF was similar to those included in meta-analysis.

CONCLUSION

In community-dwelling adults, higher levels of PA regardless of intensity were associated with higher levels of a broad range of PF measures. These findings support the potential of device-based measures of movement being used to remotely monitor people for risk of low PF without the need to attend a clinic or laboratory. The cross-sectional nature of all but one study and the focus on older age populations prevents generalisability of these associations to younger populations. Future research should also investigate a broader range of potentially important PA measures, especially those that capture how PA is accumulated within and between days.

Acknowledgements We would like to thank all the authors that provided additional data to allow us to include their studies in this review. We would also like to acknowledge the cohort studies, which are mentioned in table 1, and all the participants.

Contributors JC, RP and MH conceived the presented review. JC and MN developed the protocol. JC, RL and GB performed the screening, data extraction and risk of bias. JC and BM developed and performed the statistical analysis. JC wrote the manuscript with contributions from RP, MH, BM, MN, RL and GB. RP supervised the project.

Funding The lead author (JC) is funded by the ESRC through the South West Doctoral Training Partnership (grant number: ES/P000630/1). J. Culverhouse @joshculverhouse

Twitter Joshua Culverhouse @joshculverhouse

Conflicts of interest None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines,
REFERENCE

1 Marmot M. Health equity in England: the Marmot review 10 years on. BMJ 2020;368:m693. 10.1136/bmj.m693 Available: http://dx.doi.org/10.1136/bmj.m693


