**Supplemental Digital Content 5: Qualitative assessment within-group studies**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Study** | **Bias due to confounding**  **Y/N** | **Bias level** | **Description of bias** | **Bias in selection of participants into the study**  **Y/N** | **Bias level** | **Description of bias** | **Bias in classification of interventions**  **Y/N** | **Bias level** | **Description of bias** |
| Jácome et al (2014)22 | N | NA | No baseline confounding or time-varying confounding | N | NA | No selection of participants based on participant characteristics after the start of the intervention | N | NA | Only one intervention group |
| Beauchamp et al (2010)19 | N | NA | No baseline confounding or time-varying confounding | N | NA | No selection of participants based on participant characteristics after the start of the intervention | N | NA | Only one intervention group |
| Marques et al (2015-1) 23 | N | NA | No baseline confounding or time-varying confounding | N | NA | No selection of participants based on participant characteristics after the start of the intervention | N | NA | Only one intervention group |
| Marques et al (2015-2)26 | N | NA | No baseline confounding or time-varying confounding | N | NA | No selection of participants based on participant characteristics after the start of the intervention | N | NA | Only one intervention group |
| Rinaldo et al (2017) CT 30 | N | NA | No baseline confounding or time-varying confounding | N | NA | No selection of participants based on participant characteristics after the start of the intervention | N | NA | Clearly defined groups, block randomisation performed without contact or knowledge of participants, baseline assessment before start of the interventions |
| Rinaldo et al (2017) EDU 30 | N | NA | No baseline confounding or time-varying confounding | N | NA | No selection of participants based on participant characteristics after the start of the intervention | N | NA | Clearly defined groups, block randomisation performed without contact or knowledge of participants, baseline assessment before start of the interventions |
| Harrison et al (2015) 31 | N | NA | No baseline confounding or time-varying confounding | N | NA | No selection of participants based on participant characteristics after the start of the intervention | N | NA | Only one intervention group |
| Harrison et al (2019) 24 | N | NA | No baseline confounding or time-varying confounding | N | NA | No selection of participants based on participant characteristics after the start of the intervention | N | NA | Only one intervention group |
| Liu et al (2019) 29 | N | NA | No baseline confounding or time-varying confounding | N | NA | No selection of participants based on participant characteristics after the start of the intervention | N | NA | Only one intervention group |
| Mesquita et al (2016) 28 | N | NA | No baseline confounding or time-varying confounding | N | NA | No selection of participants based on participant characteristics after the start of the intervention | N | NA | Only one intervention group |

**Supplemental Digital Content 5 (continued)**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Study** | **Bias due to missing data Y/N** | **Bias level** | **Description of bias** | **Bias in measurement of outcomes**  **Y/N** | **Bias level** | **Description of bias** | **Bias in selection of the reported result Y/N** | **Bias level** | **Description of bias** |
| Jácome et al (2014)22 | Y | Low | 13.3% drop-out | Y | Serious | Evaluators were the same healthcare professionals that delivered the PR-program, this might have influenced the outcome assessment | N | NA | TUG is a valid measure for functional balance |
| Beauchamp et al (2010)19 | Y | Low | 12% drop-out | Y | Low | Evaluators at post rehab-assessment were unaware of baseline test scores | N | NA | BBS and TUG are valid to assess balance |
| Marques et al (2015-1) 23 | Y | Serious | 35.3% drop-out | Y | No information | No information on whether or not evaluators were aware of baseline scores at post-test | N | NA | TUG is a valid measure for functional balance |
| Marques et al (2015-2)26 | Y | Low | 10% drop-out | Y | No information | No information on whether or not evaluators were aware of baseline scores at post-test | N | NA | TUG is a valid measure for functional balance |
| Rinaldo et al (2017) CT 30 | Y | Low | 14.3% drop-out | Y | No information | No information on whether or not evaluators were aware of baseline scores at post-test | N | NA | Timed one leg stance is a valid measure for balance |
| Rinaldo et al (2017) EDU 30 | Y | Low | 14.3% drop-out | Y | No information | No information on whether or not evaluators were aware of baseline scores at post-test | N | NA | Timed one leg stance is a valid measure for balance |
| Harrison et al (2015) 31 | Y | Serious | 32.1% drop-out | Y | Serious | Evaluators performing post-tests were aware of baseline scores | N | NA | BBS and BEStest are valid measures for balance |
| Harrison et al (2019) 24 | Y | Serious | High drop-out rates:  Post PR: 20%  3 mo. FU: 57.5%  6 mo. FU: 72,5% 12 mo. FU: 85% | Y | No information | No information on whether or not evaluators were aware of baseline scores at post-test | N | NA | BBS and BEStest are valid measures for balance |
| Liu et al (2019) 29 | N | Low | No drop-out | Y | Low | Evaluators at post rehab-assessment were aware of baseline test scores, however this will not have influenced the main outcomes of the study (gait-parameters) since no instructions were given with regards to the outcomes during the tests | N | NA | TUG is a valid measure for functional balance |
| Mesquita et al (2016) 28 | Y | Moderate | 24% drop-out | Y | Moderate | Evaluators at post rehab-assessment were aware of baseline test scores | N | NA | TUG is a valid measure for functional balance |

**Supplemental Digital Content 5 (continued)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Study** | **Overall bias Y/N** | **Bias level** | **Description of bias** |
| Jácome et al (2014)22 | Y | Serious | Evaluators were the same healthcare professionals that delivered the PR-program, this might have influenced the outcome assessment |
| Beauchamp et al (2010)19 | Y | Low | 12% drop-out to be expected in PR |
| Marques et al (2015-1) 23 | Y | Serious | High percentage of drop-out and no information on knowledge of baseline-results in post-test evaluators |
| Marques et al (2015-2)26 | Y | No information | 10% drop-out  No information on whether or not evaluators were aware of baseline scores at post-test |
| Rinaldo et al (2017) CT 30 | Y | No information | No information on whether or not evaluators were aware of baseline scores at post-test |
| Rinaldo et al (2017) EDU 30 | Y | No information | No information on whether or not evaluators were aware of baseline scores at post-test |
| Harrison et al (2015) 31 | Y | Serious | 32.1% drop-out; Evaluators performing post-tests were aware of baseline scores |
| Harrison et al (2019) 24 | Y | Serious | High drop-out rates; No information on whether or not evaluators were aware of baseline scores at post-test |
| Liu et al (2019) 29 | Y | Moderate | Evaluators at post rehab-assessment were aware of baseline test scores, however this will not have influenced the main outcomes of the study (gait-parameters) since no instructions were given with regards to the outcomes during the tests |
| Mesquita et al (2016) 28 | Y | Moderate | 24% drop-out  Evaluators at post rehab-assessment were aware of baseline test scores |