Appendix H. Detailed Strength of Evidence Assessment Table

Table H1. Detailed strength of evidence assessment

| Key Question or Population | Outcome | Comparison | Risk of Bias for the evidence-base | Consistency | Precision | Directness | Overall Rating | Key Findings and Comments |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *Key question 1* |  |  |  |  |  |  |  |  |
| All facing decisions in no worse than early cancer | Knowledge about the condition or the options | Using vs. not using DAs | Low to moderate | Somewhat inconsistent (high between study SD) | Mostly precise | Direct | High | - 38 trials (12,484) patients in analysis  - SMD: 0.23 (0.09, 0.35)  - Outcome is a surrogate of decisional quality (as concept) |
|  |  | Between DAs, according to delivery formats\* | Low to moderate | Mostly consistent | Somewhat imprecise | Indirect (based on hierarchical regression) | Low | - [see above for number of trials and patients]  - No statistical evidence for a difference between DAs with and without attributes; however 95% CrI are wide |
|  |  | Between DAs, according to their content\*\* | Low to moderate | Mostly consistent | Somewhat imprecise | Indirect (based on hierarchical regression) | Low | - [see above for number of trials and patients]  - No statistical evidence for a difference between DAs with and without attributes; however 95% CrI are wide. |
|  |  | Between DAs, according to other attributes\*\*\* | Low to moderate | Mostly consistent | Somewhat imprecise | Indirect (based on hierarchical regression) | Low | - [see above for number of trials and patients]  - No statistical evidence for a difference between DAs with and without attributes; however 95% CrI are wide |
|  | Congruence of choice and values, informed choices, accurate risk perception | Using vs. not using DAs | Low to moderate (few studies report results) | Mostly consistent | Imprecise | Direct | Low | - 11 trials (4455 patients) for congruence/informed choices; 8 trials (2316) patients for risk perception  For all listed outcomes:  - No quantitative synthesis  - Using DAs better than not using in most studies  - Outcomes are surrogates of decisional quality (as concept)  - Magnitude of clinically important effects unclear |
|  |  | Between DAs, by formats, contents or other attributes | Low | Undefined | Imprecise | Undefined | Not rated | - [see above for number of trials and patients]  - Not feasible to assess because of limited number of trials |
|  | Decisional conflict scale | Using vs. not using DAs | Low to moderate | Mostly consistent | Mostly precise | Direct | Moderate | - 28 trials (7,923 patients) in analysis  - WMD: -0.5.3 (-8.9, -1.8) on a 0-100 scale  - Clinically important difference unclear; the observed WMD is likely small |
|  |  | Between DAs, according to delivery formats\* | Low to moderate | Mostly consistent | Somewhat imprecise | Indirect (based on hierarchical regression) | Low | - [see above for number of trials and patients]  - No statistical evidence for a difference between DAs with and without attributes; however 95% CrI are somewhat wide |
|  |  | Between DAs, according to their content\*\* | Low to moderate | Mostly consistent | Somewhat imprecise | Indirect (based on hierarchical regression) | Low | - [see above for number of trials and patients]  - No statistical evidence for a difference between DAs with and without attributes; however 95% CrI are somewhat wide |
|  |  | Between DAs, according to other attributes\*\*\* | Low to moderate | Mostly consistent | Somewhat imprecise | Indirect (based on hierarchical regression) | Low | - [see above for number of trials and patients]  - No statistical evidence for a difference between DAs with and without attributes; however 95% CrI are somewhat wide |
|  | Proportion undecided | Using vs. not using DAs | Low to moderate | Consistent | Imprecise | Direct | Low | - 4 trials (2483 patients)  - All trials show statistically significant results that the proportion undecided is lower in DAs |
|  |  | Between DAs, by formats, contents or other attributes | Low | Undefined (sparse data) | Imprecise | Undefined | Not rated | - [see above for number of trials and patients]  - Not feasible to assess because of limited number of trials per outcome definition |
|  | Communication with provider, participation in decisionmaking, satisfaction with decisionmaking, actual/ intended choices | Using vs. not using DAs | Low to moderate (relatively few studies report results) | Somewhat consistent or undefined, depending on outcome | Imprecise | Direct | Insufficient | - 1 trial (256 patients) for communication; 8 (2173) for participation in decisionmaking; 4 (1131) for patient satisfaction; 48 trials for actual/intended choices  For all listed outcomes:  - No quantitative synthesis  - Outcomes are surrogates of decisional quality (as concept)  - Magnitude of clinically important effects unclear |
|  |  | Between DAs, by formats, contents or other attributes | Low | Undefined (sparse data) | Imprecise | Undefined | Not rated | - [see above for number of trials and patients]  - Not feasible to assess because of limited number of trials per outcome definition |
|  | Anxiety | Using vs. not using DAs | Low to moderate | Consistent | Precise | Direct | High | - 14 trials (2958 patients) in analysis  - STAI WMD: -0.1 (-1.0, 0.7)  - Clinically important difference unclear; indications that the observed WMD is small |
|  |  | Between DAs, according to delivery formats\* | Low to moderate | Mostly consistent | Imprecise | Indirect (based on hierarchical regression) | Low | - [see above for number of trials and patients]  - No statistical evidence for a difference between DAs with and without attributes |
|  |  | Between DAs, according to their content\*\* | Low to moderate | Mostly consistent | Imprecise | Indirect (based on hierarchical regression) | Low | - [see above for number of trials and patients]  - No statistical evidence for a difference between DAs with and without attributes |
|  |  | Between DAs, according to other attributes\*\*\* | Low to moderate | Mostly consistent | Imprecise | Indirect (based on hierarchical regression) | Low | - [see above for number of trials and patients]  - No statistical evidence for a difference between DAs with and without attributes |
|  | Depression, emotional distress, decision regret, quality of life | Using vs. not using DAs | Low to moderate (relatively few studies report results) | Somewhat consistent or undefined, depending on outcome | Imprecise | Direct | Low | - 8 trials (1075 patients) for decision regret, 4 (777) for quality of life, 17 (not all analyzable) for depression  For all listed outcomes:  - No quantitative synthesis  - Outcomes are surrogates of decisional quality (as concept)  - Magnitude of clinically important effects unclear  - No indication for difference |
|  |  | Between DAs, by formats, contents or other attributes | Low | Undefined | Imprecise | Undefined | Not rated | - [see above for number of trials and patients]  - Not feasible to assess because of limited number of trials on the same outcome definition |
|  | Resource use, length of consultation, costs, litigation rates | Using vs. not using DAs | Unclear | Undefined | Imprecise | Undefined | Not rated | - 1 trial (314 patients) for resource use, 3 (417) for length of consultation, no trials on litigation rates  - Not feasible to assess because of limited number of trials or no evidence |
|  |  | Between DAs, by formats, contents or other attributes | Unclear | Undefined | Imprecise | Undefined | Not rated | - [as above] |
| Separately for populations at average risk, high risk, or with early cancer | Knowledge | Using vs. not using DAs, (evidence for differential effects by population group) | Low | Generally in agreement with respective outcome | Somewhat imprecise | Direct (amounting to a subgroup analysis) | Moderate | - 38 trials (12,484) patients  - No statistical evidence for a difference between DAs with and without attributes; however 95% CrI are wide |
|  | Decisional conflict, anxiety | Using vs. not using DAs, (evidence for differential effects by population group) | Low | Generally in agreement with respective outcome | Precise | Direct (amounting to a subgroup analysis) | Moderate | - 28 (7,923) for decisional conflict, 14 (2958) for anxiety  - No statistical evidence for a difference between DAs with and without attributes; 95% CrI are somewhat wide (decisional conflict) or narrow (anxiety) |
|  | All other outcomes | Using vs. not using DAs, (evidence for differential effects by population group) | Unclear | Undefined | Imprecise | Undefined | Not rated | - Not feasible to assess because of limited number of trials or no evidence |
| *Key question 2* |  |  |  |  |  |  |  |  |
| *All providers or prospective participants* | All aforementioned outcomes | Using vs. not using interventions to promote use of DAs | Low | Undefined | Imprecise | Direct | Not rated | - 3 cluster randomized trials with 5, 120, 220 clusters, one study on financial incentives and one on an academic detailing intervention.  - No empirical data for most aforementioned outcomes; or from at most one study  - This question was used to contextualize the first key question: The overall goal is to promote shared decisionmaking; promotion through DA use is not the only approach. |

\*Audiovisual material, software or website, printed material, in-person education, option grid, decision board.   
\*\*Explicit values clarification, probability of outcomes (generic), probability of outcomes (personalized), others’ opinions, coaching in decisionmaking (human mediated), guidance in decision making (non-human-mediated), decision analytic model  
\*\*\*Developed based on theory, needing a human to deliver, having both explicit clarification of values and presenting personalized probabilities of outcomes, tailored to target population, used by patient and provider, used by patient only, includes human for logistical support, includes support group, includes patient navigator.