Supplementary Table 3: Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach ratings

|  |  |  |  |
| --- | --- | --- | --- |
| **Study, year, Country** | **Study type (Grade rating)** |  | **GRADE tool** |
| Arteaga. 2020 | Observational (Low) | Strong Association (+) | Moderate |
| Berastegui Garcia. 2020 | Observational (Low) |  | Low |
| Duchnowski. 2020 | Observational (Low) | Inconsistency (-), Strong Association (+) | Low |
| Gleason. 2017 | Observational (Low) |  | Low |
| Hosler. 2019 | Observational (Low) | Imprecision (-) | Very Low |
| Mahanna-Gabrielli. 2020 | Observational (Low) |  | Low |
| Miguelena-Hycka. 2019 | Observational (Low) |  | Low |
| Rothrock. 2019 | Observational (Low) | Indirect (-) | Very Low |
| Sanchez. 2020 | Observational (Low) | Inconsistency (-) | Very Low |
| Sokas. 2020 | Observational (Low) | Dose response (+) | Moderate |
| Susano. 2020 | Observational (Low) |  | Low |
| Tejiram. 2021 | Observational (Low) |  | Low |
| Valdatta. 2019 | Observational (Low) | Imprecision(-), indirectness (-) Strong Association (+) | Very Low |
| Wang. 2018 | Observational (Low) |  | Low |
| Yin. 2020 | Observational (Low) |  | Low |
| Chen. 2021 | Observational (Low) | Dose response (+) | Moderate |
| Torrez-Perez. 2021 | Observational (Low) |  | Low |
| Pedemonte. 2021 | Observational (Low) |  | Low |

The quality of evidence was assessed across the domains of risk of bias, consistency, directness, precision, and publication bias. The overall quality was rated as 1 of the 4 following levels of evidence, high (additional research is very unlikely to change our confidence in the estimated effect), moderate (additional research is likely to have an important impact on our confidence in the estimated effect and may change the estimate), low (additional research is very likely to have an important impact on our confidence in the estimated effect and is likely to change the estimate), or very low (very uncertain about the estimated effect). The quality of evidence was rated up for large magnitude of effect (strong association), dose-response gradient, and residual confounding decreasing the magnitude of effect and rated down for imprecision, inconsistency, indirectness, and publication bias. The final certainty was rated very low to high, in terms of how similar the true effect is to the estimated effect.