

### **Annex 1. Data extraction methods**

The abstracts of the articles included in the reviews and EGMs were analysed to determine if the article explicitly described examples of best practice and included either a quantitative or a qualitative description of a beneficial output or outcome in one or more domains of the PHCPI conceptual framework. Evidence included descriptive qualitative research and case studies that described an overall benefit; programme evaluations with a pre–post design or those lacking an explicit study design; interventions using experimental or quasi-experimental designs; and literature reviews (including scoping reviews, narrative syntheses and systematic reviews with or without meta-analyses). Perspective articles, opinion pieces and articles primarily focused on identifying gaps in knowledge were excluded.

For those articles that provided evidence of benefit, a full text review was conducted and the inputs, service delivery processes and outputs/outcomes described were coded to one or more PHCPI domains. A random sample of 10% of articles were coded by a second reviewer (DPr). Coding disagreements were discussed for each article and consensus reached on the principles that should be applied when coding articles to the domains of the PHCPI framework. Subsequently, the first reviewer (MS) coded the remainder of the dataset. Because the source reviews were completed in early 2019, we also conducted supplementary ad hoc searches of the literature based on expert opinion from members of the PHCRC network to identify additional articles published to December 2020 describing best practices prior to COVID-19. PHCPI inputs and outputs were tabulated to understand the density of interventions and outcomes by geographical region, population focus and study type. Documents were coded in a spreadsheet and frequency analyses were conducted using R software.