2 Method for developing the guideline

2.1 Guideline contributors

In order to develop the hypertension guideline, WHO established three groups:

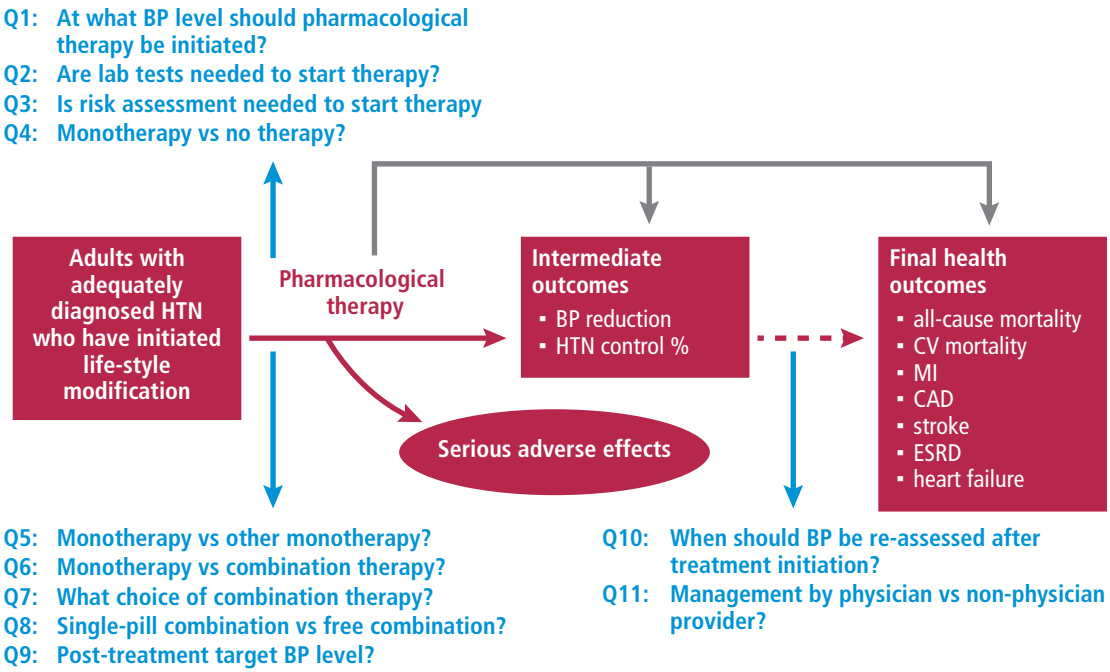
1. An internal WHO Steering Group to coordinate the guideline development process.
2. A Guideline Development Group (GDG), composed of hypertension (HTN) physicians, nephrologists, cardiologists, pharmacists, nurses, researchers, academics, policy-makers and representatives of patient groups, to review the evidence and develop recommendations. WHO selected the members of the GDG based on relevant expertise but it also considered appropriate representation by region and sex.
3. An External Review Group (ERG), composed of technical experts, representatives of HTN patient groups and ministries of health from low-resource countries, to provide peer review of the guideline and ensure recommendations are aligned with current global needs.

Annex 1 lists the contributors in each group. Annex 2 describes the process for declaring and managing conflicts of interest.

2.2 Analytical framework and PICOs

An initial GDG meeting was held in Geneva in July 2019 to determine the scope and PICO (population, intervention, comparison, outcome) questions of the guideline. The GDG first developed an analytical framework (Fig. 1) that demonstrates the impact of interventions on intermediate and final outcomes and displays the order of the key questions to better visualize and place them along the patient-flow pathway. Following this, and a preliminary scoping review and discussion between the steering group and methodologist, PICO questions were developed. These were considered, deliberated on, further refined and voted on during the meeting.

Fig. 1 Analytic framework for antihypertensive medication treatment



2.3 Outcome importance rating

Members of the WHO Steering Group, in consultation with the GDG and methodologist, developed a list of treatment outcomes most relevant to the care of individuals with HTN. The GDG then rated each outcome on a scale from 1 to 9 and indicated whether it considered each outcome critical (rated 7–9), important (rated 4–6) or not important (rated 1–3) for decision-making. The mean scores are provided in Annex 3.

2.4 Reviews of evidence

The WHO Steering Group, with the participation of the GDG, determined the scope of the guideline and identified eleven questions in the format of population, intervention, comparison, and outcomes (PICO) to guide the search for systematic reviews (Annex 4). Eleven overviews of reviews informed the guideline development process. A systematic search was carried out in PubMed, Embase, The Cochrane Library, and Epistemonikos to identify existing systematic reviews that answered the PICO questions published in 2015 or after. Suitable systematic reviews were then evaluated based on the following criteria:

- their methodology as appraised by the AMSTAR (Assessing the Methodological Quality of Systematic Reviews) tool;
- how directly they matched the PICO questions;
- whether they reported sufficient information to allow for an assessment of the certainty of the evidence (e.g. tables with characteristics of included studies, risk of bias assessments at the study level, results of meta analyses in forest plots);
- whether they reported evidence in subgroups of interest (e.g. patients with diabetes mellitus (DM), cardiovascular disease (CVD), chronic kidney disease (CKD) etc); and
- the date of the most recent review to ensure the most updated evidence was used.

The Systematic Review Team prioritized the most useful reviews for each question, comparison, outcome and subpopulation within questions, and included as many reviews as necessary to address each question comprehensively. (Web Annex A provides details of the search terms and strategy used.) No reviews were updated.

Two questions (PICO 2 and 10, see Annex 4) were not addressed in an existing systematic review, and evidence from primary studies was therefore used. In this case, the Systematic Review Team reviewed the list of studies used in existing guidelines, queried the GDG and content experts, and searched for primary studies.

A total of 159 systematic reviews and 17 additional primary studies were included. Most of these were traditional systematic reviews in which the authors conducted pairwise meta-analysis, whereas nine analysed individual patient data. The Systematic Review Team also identified several published network meta-analyses. (See Web Annex A for full details.)

Most of the included reviews were found to be of high certainty when assessed using the AMSTAR tool. Details of the selection process, and the reviews and studies included for each PICO question, are presented in Web Annex A.

Another overview of reviews was conducted to identify systematic reviews to inform other decision criteria in the evidence-to-decision framework, including people's values, resources, acceptability, feasibility and equity, presented in Web Annex A. This review was enriched by primary studies identified by GDG members. This review focused on evidence informing HTN management in low- and middle-income settings.

2.5 Certainty of evidence and strength of recommendations

The GDG rated the certainty of evidence and developed the recommendations using the GRADE (Grading of Recommendations Assessment, Development and Evaluation) approach (4). When making recommendations, GRADE defines the certainty of a body of evidence as "the extent of our confidence that the estimates of an effect are adequate to support a particular decision or recommendation" *(5)*.

Members of the GDG, with the help of the methodologist, developed evidence profiles to summarize relative and absolute estimates of effects, and an assessment of the certainty of the evidence. One evidence profile for each comparison within a PICO question was constructed, using the online Guideline Development Tool GRADEpro (https://gradepro.org). When there were systematic reviews addressing the relative effects in specific subpopulations, separate evidence profiles were constructed for each subpopulation.

According to the GRADE approach, the certainty of the evidence can be high, moderate, low, or very low. Bodies of evidence from randomized controlled trials – which comprised almost the totality of those included in this guideline – start the assessment as high-certainty evidence but can be down-rated due to considerations of risk of bias, inconsistency, imprecision, indirectness, and publication bias. Table 1 presents an explanation of the different levels of certainty of the evidence *(5)*.

Table 1 Certainty of evidence and its implications

Certainty level	Definition
High	We are very confident that the true effect lies close to that of the estimate of effect.
Moderate	We are moderately confident in the effect estimate. (The true effect is likely to be close to the estimate of effect, but there is a possibility that it is substantially different.)
Low	Our confidence in the effect estimate is limited. (The true effect may be substantially different from the estimate of the effect.)
Very low	We have very little confidence in the effect estimate. (The true effect is likely to be substantially different from the estimate of effect.)

The strength of the recommendations reflects the degree of confidence of the GDG that the desirable effects (e.g. beneficial health outcomes) of the recommendations outweigh the undesirable effects (e.g. adverse effects). The strength of recommendations in this guideline was graded into two categories:

A strong recommendation is one for which the GDG was confident that the desirable effects of adhering to the recommendation outweigh the undesirable effects.

A weak or conditional recommendation is one for which the GDG concluded that the desirable effects of adhering to the recommendation probably outweigh the undesirable effects, but was not confident about these trade-offs.

2.6 Deciding upon recommendations

The GDG met virtually for four sessions in February 2021. Systematic reviews and GRADE tables were presented at the meeting. Formulation of recommendations and their strength were facilitated by the chair and supported by the methodologist, with the aim of achieving unanimous consensus. While the plan was to use a simple majority vote, full consensus was reached on all recommendations.

The GDG used evidence-to-decision tables to guide the process of developing recommendations. These tables addressed the certainty of evidence, the balance between desirable and undesirable effects, values, resource use and cost-effectiveness, equity, acceptability and feasibility. These tables are available in Web Annex B, and the full evidence profiles in Web Annex A. The WHO Steering Group noted remarks made by members of the Expert Review Group and considered incorporating these into the final guideline. Some evidence-to-decision frameworks were consolidated to provide recommendations that are more practical and implementable from an end-user perspective; hence, the 11 questions led to eight recommendations.

2.7 Funding

The development of this guideline was financially supported by the US Centers for Disease Control and Prevention and the World Health Organization.