

How to use this 10-point checklist

- Read through your AI-generated plan and for each question, mark:
 - Yes (meets standard)
 - Unsure (needs review)
 - No (problem identified)

KEY: Mostly Yes answers: Plan is likely usable with minor refinement

More than 3 Unsure answers: Review sections carefully and revise

More than 3 No answers: Do not use without expert input

1) Clarity of Purpose and Scope

Does the plan clearly explain what is being evaluated and why?

- ☐ Is the purpose of the evaluation stated clearly (e.g., improve program, assess impact, inform funding)?
- ☐ Are the main evaluation questions specific and understandable (not vague or generic)?
- ☐ Do the questions match the program's goals or intended outcomes?
- ☐ Is the scope realistic (not trying to answer too many things at once)?

Red flag: Questions sound generic (e.g., "Does the program work?")

2) Alignment with Program Reality

Does the plan reflect your actual program and context?

- ☐ Does it accurately describe your program activities and target population?
- ☐ Does it consider real-world constraints (time, budget, data availability)?
- ☐ Does it reflect how your program actually operates (not an idealized version)?
- ☐ Are important stakeholders or groups included?

Red flag: The plan doesn't reflect how your program works in practice.

3) Logic and Causal Reasoning (Theory of Change)

Does the plan explain how the program is supposed to produce results?

- ☐ Is there a clear explanation of how activities lead to outcomes?
- ☐ Are key assumptions and contextual factors clearly stated?
- ☐ Do the evaluation questions test meaningful relationships (not just descriptions)?

Red flag: The plan jumps from activities to outcomes without explaining how or why.

4) Appropriateness of Methods

Are the suggested evaluation methods suitable for your questions?

- ☐ Do methods match the questions, available resources, and timeline?
- ☐ Are data sources realistic and available?
- ☐ Does the plan avoid unnecessarily complex methods?

Red flag: The plan recommends advanced methods without explaining feasibility.

5) Validity of Measures and Indicators

Are you measuring the right things in the right way?

- ☐ Are indicators clearly defined (not vague like “success” or “impact”)?
- ☐ Do the indicators actually reflect the outcomes you care about?
- ☐ Are both outputs (activities) and outcomes (results) considered?
- ☐ Are qualitative aspects (e.g., participant experiences) included where relevant?

Red flag: Indicators are unclear, generic, or not tied to real outcomes.

6) Bias, Fairness, and Inclusion

Does the plan consider equity and avoid biased assumptions?

- ☐ Are different groups (e.g., demographic or geographic) considered?
- ☐ Does the plan meaningfully incorporate diverse groups and voices?
- ☐ Are potential sources of bias acknowledged or mitigated?

Red flag: Evaluation outcomes are not reviewed for disparities

7) Transparency and Explainability

Can you clearly explain how the plan was developed and how it works?

- ☐ Can you explain the logic of the evaluation plan in your own words?
- ☐ Are key assumptions and decisions documented?

Red flag: You do not understand why certain questions, methods, or indicators were chosen.

8) Feasibility and Practicality

Can this evaluation realistically be carried out?

- ☐ Is the timeline reasonable?
- ☐ Are required data, tools, or staff available?
- ☐ Is the level of effort appropriate for your organization and information need?
- ☐ Are roles and responsibilities clear?

Red flag: The plan is too complex or resource-intensive to implement.

9) Ethical and Data Considerations

Does the plan respect privacy and ethical standards?

- ☐ Does it protect sensitive data (e.g., personal or confidential information)?
- ☐ Are consent and ethical considerations mentioned where relevant?
- ☐ Does it avoid using data in ways that could harm participants?

Red flag: No mention of data protection, confidentiality, or human subjects safeguards.

10) Human Oversight and Review

Has the plan been reviewed critically by a human?

- ☐ Have you questioned and validated the AI output (not just accepted it)?
- ☐ Has someone with relevant expertise reviewed the plan (if available)?

Red flag: The plan is being used “as-is” without review or validation.