What this is: Alex's rough notes on how to frame red teaming, some questions to ask, and process tips, with help from Claude. It's based on some questions we could ask of our syphilis screen and treat grantmaking, but the broader principles should generalize.

Framing (question you're asking and deliverable)

- Question is: What are the top ways we could be wrong about syphilis screen and treat programs, and what research would you prioritize over the next year to learn more about this?
- Deliverable is a prioritized list of the issues most likely to change our grantmaking to syphilis.

Tips for what to look into during red teaming

- Focus on big-picture issues that could fundamentally change our approach to syphilis grantmaking, not just minor CEA adjustments. Things like (h/t Claude from reading previous red teaming guidance)
    - Alternative approaches: Are there alternative approaches to address syphilis in pregnant women that we should explore? If we started with "reducing congenital syphilis is an effective intervention, what's the best way to achieve this," would we end up somewhere different?
    - Counterfactual scenarios: What would happen if we didn't fund syphilis screening and treatment programs? Would other funders step in? Would governments find alternative ways to address this issue?
    - Unintended consequences: Are there potential negative outcomes from large-scale syphilis screening programs that we haven't considered? For example, could it impact healthcare-seeking behavior for other conditions?
    - Funding landscape: Why don't other major global health funders prioritize syphilis interventions? What does this imply about the intervention's value or challenges?
    - Integration with other interventions: How does syphilis screening and treatment interact with other maternal and child health interventions we fund? Are there synergies or conflicts we should consider?
    - Long-term impact: How might funding syphilis interventions affect health systems and behaviors in the long term? Are we building sustainable capacity or creating dependence?
    - Scalability and generalizability: How confident are we that the results from pilot programs or studies will translate to large-scale implementation across different contexts?
    - Cost-effectiveness comparisons: How does the cost-effectiveness of syphilis interventions compare to our other top programs intuitively? If it's significantly higher or lower, why might that be?
    - M&E: How do we know coverage is increasing (both more screening and more treatment)?

    - Granular details: How does the test work? How does treatment work? How could this fail?
    - Macro-level impacts: If we take our estimated impact of syphilis interventions and extrapolate to a national or global scale, does the predicted reduction in adverse birth outcomes align with observed trends or seem plausible?
    - Learning opportunities: Are there unique learning opportunities from funding syphilis interventions that could inform our broader strategy or other areas of grantmaking?
    - Theory of change robustness: What are all the steps in the theory of change for syphilis interventions, and are there any links we're missing or assuming too much about?
    - Which countries: Why are we funding in the countries we fund? Why not others?
    - Optimization concerns: Are there areas where we might be getting "optimizer's curse" in our analysis of syphilis interventions? Are we potentially overestimating impact in high-burden areas due to noisy data?
    - Future trends: How might emerging technologies or medical advances impact the relevance or effectiveness of syphilis interventions in the near future?
    - Findings from red teaming other programs: Do our takeaways from other red teaming apply here?
- CEA things to check (h/t Claude from reading previous red teaming guidance):
    - Baseline prevalence: How accurate and up-to-date is our data on syphilis prevalence in target populations? Consider potential regional variations and trends over time.
    - Test accuracy: Are we accurately accounting for false positives and false negatives in syphilis screening tests? How might this affect overall cost-effectiveness?
    - Treatment efficacy: Are we using the most current data on the efficacy of syphilis treatment in preventing adverse outcomes? Consider potential differences in efficacy for early vs late treatment.
    - Counterfactual coverage: What proportion of women would receive syphilis screening and treatment without our intervention? This could vary significantly by region or health system capacity.
    - Indirect benefits: Are we fully capturing indirect benefits, such as prevented infections in sexual partners or future pregnancies?
    - Cost variations: How much do implementation costs vary across different contexts? Are we accounting for potential economies of scale or diseconomies of scale in our calculations?
    - Disability weights: Are the disability weights we're using for various outcomes (e.g., stillbirth, congenital syphilis) appropriate and consistent with other interventions we evaluate?
    - Retention and follow-up: Are we accurately modeling the proportion of women who complete the full course of treatment and follow-up care?

    - Health system strengthening effects: Should we include any adjustments for potential positive externalities on the broader health system from implementing syphilis screening programs?
    - Timeframe of benefits: Are we correctly modeling the duration of benefits, especially for outcomes like prevented congenital syphilis which may have lifelong impacts?
    - Interaction effects: If syphilis screening is implemented alongside other maternal health interventions, should we adjust for potential synergies or diminishing returns?
    - Secular trends: Are we adequately accounting for background trends in syphilis prevalence, healthcare access, or other relevant factors that might affect the intervention's impact over time?
    - Age weighting: Given that this intervention primarily affects newborns, should we consider any moral weights related to averting death or disability at the start of life?
    - Leverage and funging: Are we accurately estimating the degree to which our funding might leverage additional resources or potentially displace other funding sources?
    - Sensitivity to key parameters: Which parameters in the CEA have the largest impact on the final cost-effectiveness estimate? Are we sufficiently certain about these high-leverage inputs?
- Consider both potential downsides and upsides we may be overlooking. A lot of times red teaming skews negative, but finding out we’re underfunding syphilis is useful, too.