To: IPPF Membership November 3, 2020

From: Lee Green, Sam Greenberg, Redstone

Open comments on proposed stream 1 formula

- compiled feedback and responses –

In October, Redstone released a draft of the proposed Stream 1 formula to the Federation for an open comment period. For the sake of transparency, we collected and are sharing all of the written comments we received, all anonymized. Where appropriate, we have provided our responses to the comments. This memo is organized into the following sections:

1. Formula changes based on the comments
2. Comments expressing general support for the formula
3. Comments on specific formula components (need metrics, performance, etc.)

We were able to accommodate many of the comments, though we could not incorporate all of them due to lack of relevant data, conflicting guidance from different stakeholders, or needing to adhere to the original GA resolution.

Thank you all once again for the valuable input throughout this process – the many comments we have received from MAs and Secretariat staff have helped make this formula stronger and better able to serve the membership.

\* \* \*

1. Formula changes based on the comments

Based on feedback received, we made several adjustments to the proposed formula:

* **Changed the Outcome 1 performance metric from ER3 (youth and women’s groups that took a public action in support of Sexual and Reproductive Health and Rights) to ER1 (policy initiatives and/or legislative changes to which IPPF advocacy contributed):** We heard from numerous respondents that ER1 better captures the ultimate impact of the advocacy work, and should be our chosen metric.
* **Updated performance calculations to look at an MA’s cumulative performance over the prior three-year cycle:** We heard concern from MAs about basing performance awards solely on the most recent year. MAs mentioned that in some years, unforeseen circumstances (e.g., natural disaster, loss of a major donor) make performance numbers dip. As a result, every time the formula is run, it will look at an MA’s cumulative performance from the prior years in the cycle.
* **Determined that all country need indicators would be based on the most recent year’s estimates:** Given the three-year cycles, some MAs wondered if country need indicators would reflect the average need levels from the prior three years, which would make the formula somewhat outdated and not reflect current realities. We therefore clarified that all need metrics would be based on the most recently available data (for this reason as well, we chose need metrics that are regularly updated, ideally every year).
* **Clarified how inequality levels within each country will impact allocations:** Given feedback from the Federation, the formula incorporates the GINI coefficient as a measure of wealth inequality. Some questioned how inequality would be factored in (if at all), so we have clarified that for upper middle income countries, all else being equal, a country with higher levels of inequality would receive more funding to recognize the poverty within its country.
1. Comments expressing general support for the formula

We received the following general comments on the formula. Responses from Redstone, where relevant, are noted in italics before or after the relevant comments:

* It is fabulous to see a framework fundamentally based on need, and that addresses upfront a range of the current anomalies in funding. I am highly supportive of your approach. I am particularly impressed with the range of indicators you are using around SRHR, which is similar to how we define SRHR need by region here. Notwithstanding the need for good in-country infrastructure to support data collection, it is the most transparent and defensible approach.
* Let me congratulate you for this impressive draft for the new allocation formula. SRHR is a complex space and trying to tie various threads in such a way is challenging. On behalf of my colleagues, I would like to appreciate the detailed work done on this draft.
* Good job. Very comprehensive and equitable as can be.
* I think the proposal is a really good one and I feel confident about it… It seems a conceptually simple and easy to understand model.
* It is noted that the allocation process will occur every three years, to enable longer term planning. Does this also imply that MAs will now be required to submit a three-year business plan?

*Yes, MAs will submit a three-year plan to show how they would spend the money allocated to them by the formula.*

1. Comments on specific formula components

We received the following feedback on specific components of the formula, which we have organized according to the formula’s five main components: country need, MA funding sources, funding floor, performance award, and funding shift adjustments.

Responses from Redstone, where relevant, are noted in italics before or after the relevant comments:

Country need

Data accuracy and frequency

*While no data set is perfect, we have consulted with SRHR experts and reviewed all existing global databases to find the most accurate data. These provide an objective sense of relative country need. We intentionally chose data sets that are updated frequently (ideally every year) so data can be as current as possible.*

* Maternal mortality rate: As from country experience this indicator does not reflect the reality since the data are misinterpreted or underreported…. [maternal deaths] are not counted in the national statistics as maternal deaths by the reasons for dying is due other health-related problems (most related to cardiovascular health problems).
* In some countries the data is not always up to date. The metrics are sometimes measured every 5 years…. due to COVID may not have the data for a number of years. E.g. the multi cluster survey where we get our SRH/HIV data from.
* UN is mentioned as the source for most of the data, while WHO is only mentioned as source of cervical cancer incidence. Is the priority given to UN data?

*We did not have a preference for UN data over WHO data, it just happened that the UN had more of the data points we needed than the WHO.*

Country income

*We received several questions on how country income is addressed. We use the World Bank classifications, which while imperfect, are a widely respected and objective metric for country income level and updated annually to stay current. Given the high levels of inequality in some wealthier countries, we incorporate the GINI coefficient as well based on feedback from the Federation. All data will be drawn from the most recent year available, so the data is as up to date as possible.*

* How is this calculated based on the fact that there is unequal distribution of wealth in a number of countries?
* Income inequality in a country can be capture with the poverty index that exist for many countries, as well as GINI index (socio-economic inequality)
* Please make sure to also consider the unequal distribution of wealth
* Will the calculations consider average incomes/GINI coefficient over a three year period, since the allocation cycle will also be three yearly?
* How are we going to measure country income? Are we going to continue to use the World Bank definitions which no longer aligns to the current situation in country, particularly in light of COVID-19

*The formula will first be run to inform 2022 allocations; by then the impact of COVID will begin showing up in these metrics*

Additional need indicators

*We received the following requests for additional data points. However, since they are not reliably available for the majority of countries where IPPF works, we are not able to include them at this time (e.g., unsafe abortion rates, marginalized populations). For fairness, the formula only includes metrics available for the majority of IPPF countries. In addition, we designed our need metrics to focus on the ultimate outcomes that affect IPPF’s clients (e.g., maternal mortality, STI transmission). Some of the proposed metrics are intermediaries to the outcomes (e.g., CSE, size of youth population); we expect these to influence the ultimate need, which we capture in our metrics (e.g., if the large youth population leads to high number of pregnancies or high levels of STI transmission, the formula will be impacted by those).*

* I think here should be other indicator to capture maternal health which in neonatal mortality rate which is high problem in many countries especially in the middle-income countries as well

*Neonatal mortality rate (as well as infant mortality rate) are correlated extremely highly with maternal mortality rate, which we have included. Adding these to the formula would not meaningfully change the results.*

* Marginalized populations: Include Migrants and refugees.

*Unfortunately, reliable data is not available on the needs of marginalized populations in each country. However, given their importance to IPPF all MAs submitting a three-year plan will have to document how their strategy meets the needs of marginalized populations, which will impact funding decisions.*

* There is also need to take into account each country’s economic and social situation.

*A country’s economic situation is factored in using its income level and level of inequality (GINI coefficient). Social dynamics as they pertain to gender and sexual and reproductive rights are measured using the five indexes of gender empowerment (including access to land, finances, jobs, justice system, political participation, marriage and divorce practices, etc.).*

* In proposed need metrics we would like to see abortion as a standalone area of work in view of its impact on morbidities and mortality and also the sensitive nature of work around this area respecting country specific legal frameworks

*The indexes on gender empowerment include data on abortion access, which was the most objective and widely-available data we found to measure this.*

* Youth is another important area of work. There is only one indicator -adolescent birth rate – to determine need for work in this area. We feel this indicator does not seem adequate to justify the broad spectrum of work necessitated for Young people to cover all dimensions of Youth SRH. In countries with a high proportion of young people, there are also other SRH issues to be addressed, besides teenage pregnancies, but these are not reflected.

*While we wish more indicators existed on youth need, they do not currently exist for enough countries; we therefore use the available data as the best proxy for this.*

* CSE is very important to IPPF work so here I would like to have metric that really capture access to CSE. It could be countries that have mandatory CSE in schools. I understand it is not a population-based data (TOP demographic statistics) but I think that adolescent birth rate does not ultimately define young people health and rights in regards to SRH. So I would add here indicators related to CSE. Maybe difficult but worth to try.

*The formula’s metrics are intentionally outcome-focused – they measure negative impact on the clients IPPF seeks to serve (e.g., maternal mortality, STI transmission). The formula is intentionally agnostic about what causes those adverse outcomes or what can reduce them (such as CSE). If a country were to have extensive CSE but still high negative results for young women, the formula would still want to give it more funding given those bad outcomes.*

* Not only population size but also population distribution e.c age distribution.

*As with the above, the formula is intentionally outcome-focused. For instance, if there is a large youth population but very low negative outcomes, the formula focuses on the lack of negative outcomes, and not the size of the youth population (and vice versa).*

* There seems to be a declining trend around the incidence of HIV. Perhaps the weighting for indicators related to HIV/AIDS can be reduced and some weightage can be added for a standalone indicator around abortion and an indicator of youth SRH needs beyond adolescent birth rates

*The OECD’s Social Institutions and Gender Index (SIGI) includes indicators for abortion access and is included in the formula. Beyond that there are not additional good indicators on these topics that are reliably updated for the majority of countries where IPPF works.*

Fewer indicators

*We received the following feedback to have fewer indicators; however, in prior discussions and surveys we received significant requests to expand, not contract, the number of indicators.*

* STIs: Cut to 2 indicators, I would delete the ART treatment access
* There are too many metrics under gender empowerment and rights, and some are not really relevant to us. I would change the area of work into gender empowerment and inequality and propose 2 indicators – Gender Inequality Index and poverty index that can also capture the characteristics of vulnerable groups in the society
* 46 indicators are too much to capture country needs related to SRHR. I suggest maybe to reduce number of indicators and for each area of work to have 2 core metrics/indicators. It will be easier to define the formula

Additional questions

*We received the following additional comments on need metrics, responded to individually below:*

* How are you going to determine 0 or 1 – what is the benchmark? What is the definition of highest need and the same for lowest need? Even if you have metrics, and for instance you use ratios, what is the cut-off point?

*For each metric, we normalize all eligible countries’ rates. The country with the lowest rate on a given metric gets a 0, and the country with the highest rate gets a 1. We then line up all the other countries in between and given them a score between 0 and 1 based on their relative standing. We do this to every need metric, so they are all lined up on a 0 to 1 scale for consistency. We then add each country’s scores from all of its metrics (based on the weighting defined earlier).* *Given that every country gets a point between 0 and 1 for each metric, we don’t need to set any cutoff points – as need metrics move up or down a country gets a little more or a little less, with no steep cliffs or drop-offs.*

* We understand that total need score will be adjusted based on the country's relative population size. Does this imply that an MA with a higher need score and smaller population size may stand to gain more resources, as compared to an MA with a lower need score and larger population size?

*Each country’s need score is multiplied by its population factor to determine the grant size. The exact size of each grant depends on how high the country need is and how large the country is. We use the square root of population rather than the absolute population number to temper the effects of very large populations. This approach is modeled on common practices of other international bodies and ensures that all funding does not get concentrated in the handful of most populous countries.*

MA funding sources

* It would be helpful to better understand the specific metrics used for this calculation, because now it is not clear how this will be calculated.

*The formula looks at each MA’s total income from all funding sources (restricted grants, private donations, government funds, fee for service, etc.). It then calculates the ratio between the MA’s total income and the preliminary size of its unrestricted core grant (from the calculations above). It identifies MAs with relatively low ratios (ones that have raised disproportionately little relative to their need). MAs then get modest adjustments to their core allocations based on their other funding sources, so that MAs with minimal other funding do not get left behind.*

* If an MA documents the intention to raise funds (A fund raising policy/strategy in place) but is unable to raise the committed amount of funds in a given year due to various external factors (such as changed regulations, laws, donor priorities, humanitarian crises), will this negatively impact fund allocation? Can you please help us understand what a ‘modest’ adjustment to allocation based on other funding sources can be like?

*MAs will not get penalized for missing their fundraising targets if they have been making a strong effort to do so, and factors beyond their control make it hard for them. For countries that have especially low fundraising relative to their need (As calculated above) they may get a small boost so they are not left behind. This would typically not change the grant amount by more than 5-10%.*

Funding floor

*The funding floor guarantees a certain minimum to all eligible MAs, regardless of their size. It does not look at how small a country’s population is to determine its minimum grant size. Each cycle IPPF can set what this minimum grant size is. This means that small island states, or other countries with low need and population numbers, who might otherwise have received a very small grant, will get boosted up to this minimum grant size. Large MAs who would receive much more than this minimum amount will not need this step, since they already receive larger grants.*

* What is considered “a small population”? Should it be included as definition in the formula.
* A similar re-classification has to be done especially for the Caribbean because we have some of the same realities of the Pacific Island States and in order to ensure that no one is left behind.
* The minimum grant amount should be known. It can have different categories due to population size of countries but we as MAs should know what the minimum grant amount is. For example for countries bellow 2 million people the minimum grant is 30,000 USD, between 2 mil – 10 mil is 50,000 and so on…Than the minimum should be adjusted to country income, so those who have lower GDP will get the same amount but those with higher GDP will get less amount
* This, we are concerned, can be tricky for a large MA like us, since our size is large and thus our fixed costs will be larger as compared to fixed costs of smaller MAs. If it’s a certain minimum grant for all, this may seem inequitable for us.

Performance award

Calculation methodology

*We heard concern from MAs about basing performance awards solely on data from the most recent year before the formula was run. MAs mentioned that in some years, unforeseen circumstances (e.g., natural disaster, loss of a major donor) make performance numbers dip. As a result, every time the formula is run, it will look at an MA’s cumulative performance from the prior years in the cycle. Once the formula is run, the performance award it gives out will remain constant for the next three years, to give MAs more predictability about income levels and enable longer-term planning.*

* This side of the Atlantic is prone to natural disasters and economic downfall therefore will experience performance shocks as countries seek to readjust and therefore they should not be penalize in times of crisis.
* If performance awards will be determined by year on ‐year growth, would these be averaged out over a three year to determine allocations? We are concerned that this may create a situation wherein an MA misses out a performance award for a given year in spite of a good overall performance for that year, since funding allocation for that year would have been pre‐determined at the beginning of the three‐year allocation process, that was done earlier.
* Does absolute magnitude of impact cover the entire federation or only the specific region that the MA belongs to?

*Given that the new allocation approach and formula will treat all MAs across all regions in the same manner, the new formula will look at MA performance compared to all other MAs.*

Performance across multiple Outcomes

*The formula recognizes that MAs often work across Outcomes 1, 2, and 3, and is designed to reward MAs for contributions across all three Outcomes. Since different MAs focus more or less on certain Outcomes, it weights each Outcome based on MA spending: the more an MA spends on an outcome, the more the performance there impacts its overall performance score. Each MA is still incentivized to do well across all 3 outcomes, since progress in any Outcome can boost its score. These calculations look at actual, not planned, spending by Outcome.*

* Please let us know if we have got this right: if an MA spends the most on Outcome 3, then its performance score will largely depend on CYP. Thus, if it has done well on Outcomes 1 and 2, but not as well on Outcome 3, its overall performance score will be lower, since greater weighting will be given to CYP because of higher proportion of spends under Outcome 3. Better performance on Outcomes 1 and 2 will only make a small positive impact on the score.

*Yes, this is correct.*

* Since MAs focus is on integration of services, they will operate in practice in all three outcomes. So performance should not focus only on one main outcome but a combination. For example, if one country wants to increase SRHR services to youth but they also know that the awareness about SRHR including rights is very low. So if you want to increase the use of SRHR services you have to work at awareness levels which include CSE training, then provide the services and also the results should be advocated to other stakeholders therefore there is a cross cutting across outcomes which should also be recognized.
* It is mentioned that performance will be assessed against outcome wise relative spends. What would be the funding implications if we were to deliver more than what we promised for a given outcome, but have an underspend within that outcome?

Contextual factors impacting performance

*We recognize that there are many complex dimensions to performance (e.g., quality of service, marginalized populations served, geographic diversity reached), as well as elements of performance that cannot be easily captured by numbers. We were asked to develop a formula based on objective data-points to determine allocations, and must rely on objective data that is available. MA plans and progress will still be reviewed by regional Secretariat staff, who can flag and advise on ways to address these dimensions of performance. In rare cases where an MA has significant performance issues not captured by the numbers, regional Secretariat staff and members of the technical review team can flag these for possible grant reductions.*

* We understand and appreciate that every effort has been made to keep the formula for determining performance awards objective and not influenced by subjective judgments. Can you additionally consider equitable ways to measure performance based on the types of population served, types of services delivered and types of geographies, in view of the examples cited above?
* Personally, I’m not keen on the Awards component. There are so many tricky issues associated with this and it introduces the possibility of ‘gaming the system’  and deeply perverse incentives to fund in the short term to optimise funding rather than strengthening services and creating systemic change. However, this remains the lesser of evils and can be managed through sophisticated evaluation systems.
* Abortion is a critical area of work that demands action related to advocacy as well as action towards improving access. Following the Global Gag Rule a significant proportion of our collective incomes have been restricted due to our rights-based stand to continue working in this area. This is also a resource intensive service, bound by the laws of the land. We feel, there ought to be a special consideration for work done by MAs around abortion; while assessing our performance. We request you to consider including abortion as a standalone area of work, with an abortion specific indicator to measure performance, in addition to CYP. In the given metrics, abortion work runs the risk of not getting counted. safe abortion services to the full extent of the law, and quality post-abortion care can be considered as important, measurable indicators since these interventions contribute to preventing deaths due to unsafe abortions
* SRHR lends itself to poverty alleviation. MAs work around livelihoods initiatives for poor and marginalized groups that certainly include but may not be limited to catering to women. There is a need for an indicator to measure this quantum of work under Outcome 2, such as, a measure of employable skills imparted within poor and marginalized populations served.
* Re the performance award, relative growth and total impact are not the only indicators that should be used to assess performance on outcomes 1-3. Performance award should also take into consideration the environment within which the MA is operating to sustain its services. User preferences have changed over the years and should be a measure that is taken into account. For instance, preference for condoms over IUDs and also the fact that we have shifted our focus to a more comprehensive approach to Reproductive health which takes into account the integration of services as opposed to a focus only on family planning, therefore the metrics have to take this into account. Demographics have changed and for many of us we are at replacement level. In countries where we have fertility below 2.0, those governments are now concerned about increasing the population. Therefore, how we measure performance of the MAs cannot be only linked to CYP, it has to be more than that.
* We may require more resources to deliver some interventions, such as service delivery in hard to reach terrains as compared to service delivery in urban settings. Similarly, some MAs may deliver certain high need, critical services such as a second trimester abortion, that may cost a lot more than delivering, say, an immunization intervention. In absolute numbers, although second trimester services would be fewer in number as compared to immunization services, the overall cost of delivering second trimester abortions would be much more than the overall cost of immunization services. How will this affect the performance award?
* Outcomes must reflect SRHR realities in our specific contexts.

Specific Expected Results indicators

*We heard several requests to modify or improve the specific Expected Results metrics. As of now, we are limited by the actual Expected Results IPPF collects from MAs – we cannot design a formula based on metrics that do not yet exist. We have intentionally designed the formula to be able to accept additional variables so that as IPPF updates the metrics it collects, they can be added easily to the formula.*

*On CYPs, we had many long discussions with SRHR experts on the most appropriate metric for service delivery, acknowledging that each metric has shortcomings. We settled on CYPs since it can best capture the impact of IPPF’s contraceptive services, despite its flaws. We believe it will be important to improve this metric to ensure it does not lead clinics to push longer-acting contraceptive methods at the expense of client choice or dignity. We heard suggestions that the IPES score, or client satisfaction scores, could act as a check on ER8 to balance these concerns. We built a way for the formula to incorporate the IPES score, but given that there are currently concerns with how it is calculated, we suggest not turning this feature “on” until the IPES scores better captures the choice available to any of the MA’s clients. The same could be done for client satisfaction score, or any other new metric IPPF starts collecting in the future.*

* For Outcome 1: Should be defined according to ER1, number of policy and legislation change, since that is most impactful (NOT youth/women networks – this is a output rather outcome or impact)
* For Outcome 1: ER 3- It is not clear if the indicator is for the number of groups or the number of persons, and/or the number of times they participated?  For one and the same action, you could have a lot of participants or just one, and the result might be the same. Moreover, actions involving governmental organizations or politicians with decision-making powers are not taken into account, yet the impact generated by these might be greater.
* For Outcome 2: ER 4 - Since a CSE program can have 20 hours as 50 or 100 hours, we suggest measuring the indicated by the number of hours per youth that were used in training or education in CSE
* For Outcome 3: CYPs: I know IPPF’s work is concentrated to contraceptive services, but there are other SRH services that should be captured here, what about number of STI/HIV testing – since HIV is the country needs was captured as very important but here it is not reflected. So in the country needs HIV incidence and contraceptive prevalence are considered equally, but when it comes to performance award it is only contraception.

*We explored including number of services offered (to capture HIV services) but were told that MAs vary significantly in how they collect and report service statistics.*

* For Outcome 3: We are very concerned about CYP as the calculations are not equitable across contraceptive methods. In a young country the unmet need for spacing is higher among underserved communities, so our priority would be to meet this unmet need. However, it would get us a higher CYP as sterilizations and IUDs would always be counted higher. 1300 cases of Condom acceptors would be equal to one Sterilization and you would agree with us that it is equally important to ensure that young people are able to use a contraceptive method to regulate their fertility as well as couples who are married are able to choose a long term method. What we can suggest is that CYP be balanced with another indicator which can measure the contraceptive uptake by type or number of acceptors. This would then balance the long term methods versus spacing methods

*We agree that it will be useful to have an additional metric to balance out CYPs. At the moment we were advised that none of the Expected Results would do this well, but hope that another indicator (client satisfaction, IPES score) could be added to do this in future years.*

Funding shift adjustments

*The formula will be run once at the beginning of the cycle and determine each MA’s allocation for the next three years. The changes in allocations will be phased in gradually, so that in year 1 the MA gets a small increase/decrease, in year 2 it gets the same increase/decrease, etc. until it reaches its new allocation amount. While the amount will change a bit each year, the MA will know the amounts for all three years in advance.*

* If there is going to be an increase it would be incremental over a period of time. Is this the interpretation of the point?

*Yes, that is correct*

* Not clear exactly how this work. If we receive funding for 4 years, each MA should have a baseline that can be calculated as average from the past 4 years, and this not to be changes in the course of next 4 years. In the next cycle it will be reduced or increased according to their performance, and not reduce or increase year by year.
* It would be interesting to see the differences between what MAs receive now and what they will receive with the formula in order to understand if they are going to make an adjustment or transition model for example year 1 (70% old and 30% new formula), year 2 (50% each), year 3 (30% and 70%) and then if 100%. This is key if they have identified any MAs for which there is a decrease in funding.

*The changes will be introduced linearly over several years, so that each year an MA gets the same amount of increase or decrease until it reaches its new allocation level (e.g., get another $20,000 added each year).*

In addition to the above, we received several questions about the policies and guidelines surrounding the formula’s implementation (e.g., roll-over of funds between years, reporting guidelines). As these are beyond the scope of the formula itself, they are not addressed here.

\* \* \*

Thank you again for your valuable feedback on this important work.