

An online debate on “Performance-based financing in low- and middle-income countries: still more questions than answers” – contributions on the discussion group

Compiled by Bruno Meessen, on September 9th, 2012

One of the two editorials of the August issue of the Bulletin of the World Health Organization was dedicated to Performance-Based Financing (PBF). Fretheim et al. reported the main findings of their Cochrane review published a few months earlier. In the weeks that followed a discussion developed on the online discussion forum of the PBF Community of Practice (<http://groups.google.com/group/performance-based-financing>). This document reports the different individual contributions. It complements a summary of the discussions, which is available online at: <http://www.healthfinancingafrica.org/>

Background information

The PBF online discussion group gathers more than 650 international experts interested in performance-based financing applied to the health sector. Members are practitioners, health cadres, researchers, technical assistants, consultants, students, etc. Facilitation of the group is rather loose: messages – in French or English – are posted instantly; a debate can be launched by any member. The discussion group is owned by the PBF community of practice, a self-managed community affiliated with the Harmonization for Health in Africa (<http://www.hha-online.org/hso//>). The discussion group is only one knowledge management activity among many others of the CoP.

Mail sent by Emmanuel Ngabire (Rwanda), August 9, 2012

Dear all,

I would like to draw your attention to the article aforementioned. Many of you may know the studies in question and I would appreciate if you can share us your comments. The opinion of their authors (especially those who are members of our community of practice) are the most welcome.

Here below you have the extract but you can read the full article on the link:

<http://www.who.int/bulletin/volumes/90/8/en/index.html>

I have also downloaded it attached it hereto for those who prefer a pdf copy.

I look forward to hear your comments.

Sincerely,

Emmanuel

From Stefaan Van Bastelaere (Belgium), August 9th, 2012

Dear Emmanuel

Thanks for sharing this interesting editorial.

Two observations :

1. The authors reduce PBF to a strategy that generates “quantities”, which in my opinion is unfair. I think we all have been contributing in scientific articles on our experiences, and what we systematically observed is that PBF is (and must be) first of all a strategy to motivate Health Professionals and to improve quality of the delivered services... Those aspects were fully documented in a lot of papers, **proving that the impact on quality and motivation was evident**. Apparently we failed to communicate this to the outside world. The authors reduce PBF to a quantity-generating strategy and easily conclude that PBF is not convincing...

2. Yes we have a methodological problem: there is a problem in setting up RCT because of ethical reasons: If you know something works, it is unethical not to apply it to all levels. It is what we experienced in Rwanda in 2005, once the strategy proved to be effective, it became a national strategy and we lost our “control” groups.

As this is an editorial of the bulletin of the WHO (!!!) I think we must rapidly react, because this editorial is really harming us.

From Bruno Meessen (Belgium), August 9th, 2012

Dear Stefaan, Emmanuel and everyone,

I guess that this editorial will trigger a large debate within our community.

I invite all of you to read it and make your own opinion. And then feel free, as Stefaan, to react and share your view on this google group. Do you agree? Disagree? What would be the best reply to these authors?

We should first seize this paper as an opportunity to self-reflect on our practice. Then, indeed we will probably respond more formally to the authors. I have already been contacted by a few of you.

As a reminder, remain serene and respectful in this debate.

Best,
Bruno

From Longin Gashubije (Burundi), August 9th, 2012

It is interesting to note that the opinions of the authors of the PBF is limited to "the transfer of money or material goods conditional on Taking a Measurable Action gold Achieving a predetermined performance target". I think PBF is more than the transfer of money when well implemented, it allows to transform the whole health system.

For the case of Burundi, one can even say that the reason why the randomized study has not been possible is that during the experiment, the PBF has quickly reveals interesting for almost all actors from the bottom until the top (Donors, Government, providers, population). So instead of waiting for the results of the study of 5 years, the PBF has quickly extended from 2 Provinces (in 2006) and 9 (in 2009) and later 17 provinces (in 2010) within 5 years.

The study found despite the disruption a significant improvement of indicators (quantitative and qualitative) that had never moved before the implementation of the PBF. In the absence of a randomized clinical trial , we

can consider that the Province in which the PBF has been in operation since mid 2006, other factors are negligible in the evaluation indicators.

Assisted deliveries by trained personnel evolved in a positive and meaningful way in these Province (I Know well Cankuzo Province where between 2006 and 2010 the birth increased from 40% to 87%).

In my humble opinion, the studies were conducted in different contexts but none showed a decrease in assisted deliveries. The results of the study by Paulin Basinga et al in Rwanda show a rather positive development indicators as elsewhere in Burundi. For the case of Burundi, the combination of subsidized care and PBF could constituted bias.

However, if one considers the Health facilities where PBF was implemented alone, the results are very positive. It could be better to redefine the PBF and make another quasi-experimental study in countries where the PBF has recently been implemented.

We have unfortunately few studies in Burundi prior to PBF which limits the understanding of changes and creates confusion when trying to show the contribution of PBF.

From Paul-Claudiel Rubeya (Burundi), August 9th, 2012

Hello,

I think as people of land, we can not deny the positive results brought about by the PBF in our in our health systems. However, we must recognize that we must seek to minimize the perverse effects of this strategy. Moreover, we must conduct studies in the requirements and associate, if possible, some of them who still doubt the results already published. Thank you.

From Atle Fretheim (Norway), August 10th 2012

Hi Bruno (long time no see – OCB-board a decade ago...) and everyone else on this google group.

I wrote the editorial and co-authored the Cochrane review that is referred to in the editorial on PBF in the latest WHO Bulletin. And I just recently joined this google group (thanks for the invite, Emmanuel).

I really welcome criticism of what we wrote and I will do my best to respond to comments and questions that are raised on this forum, as timely and clear as I can. An important part of my job as a researcher is to question things, and I certainly appreciate it if people question what I say or write!

I have little or no personal, professional or political conflicting interests related to this topic. I started looking into research on PBF simply because the Norwegian Agency for Development (Norad) commissioned me and a colleague to review the topic a while back, since Norway is investing lots of money on performance based financing schemes.

Please note that the Cochrane review contains tons more information than what is presented in the editorial piece from the WHO Bulletin. I attach a copy of the Cochrane review – it's a tough read, but may be worth it to some of you (at least to skim through). And note that the editorial is mainly an opinion piece about methodological challenges with evaluating health system interventions (e.g. PBF), while the Cochrane review addresses the "does PBF work?"-question – at least partly.

If it were possible to engage in a debate on how PBF-schemes might be evaluated – in a practical and feasible, but robust way – I would find that interesting and helpful. Many of the members of this google group have plenty of “hands on” experience with implementing such programs, and could probably provide important contributions to such a discussion.

Cheers,
Atle F.

From Bruno Meessen (Belgium), August 10th 2012

Dear Atle,

Thanks for joining the group and engaging with us. I have been hoping this move.

Your Cochrane review has been read by several of us. As your paper in the Bulletin.

I am sure that the whole PBF community appreciates efforts done to rigorously document PBF development in LICs. We are confident that it will confirm the value of the strategy. We understand the importance of impact evaluation and understand the need that some other researchers scrutinize the rigor of the impact evaluation studies done by others. The concept of ‘separation of functions’ is dear to us and we appreciate that it applies to what we do as well.

Having said this, I will not lie to you: several of us have been frustrated by the Cochrane Review process. Many PBF practitioners are already struggling with impact evaluators, who do not always fully understand what PBF is. Researchers describe PBF as “P4P”, establish inappropriate links with experiences in HICs, they miss the holistic nature of the strategy. Often, they lack the culture of operating health systems in LICs. At least, they come to countries to engage with us when they have to nest their impact evaluation in the intervention (a real constraint for PBF implementation itself, but this is another discussion).

Then your group, far from the field, assess work done by these people (and some of us, I agree). In your aggregating/synthesizing process, you lose a lot of the context and make oversimplification. Your systematic review lumps together things very different. You miss the systemic outcomes. Of course, you acknowledge that in the review and the paper, but some harm has probably been done in the meantime. A pity that we have been in contact before.

For instance, several of us (cf the mail by Longin yesterday) do not endorse anymore the definition you use in the WHO Bulletin. We find it too narrow and focusing on the provider payment contract only. PBF is in fact a package reshaping the whole nexus of contracts. Now, there is a new intention of impact evaluators to design studies to isolate the contribution of each sub-component of the intervention. It is interesting, but again it betrays to some extent the PBF strategy. We feel sometimes that PBF and Africa are becoming a playground for empirical scientists to test ideas. In fact, many questions can be answered by mere logic.

We understand that we do not control anything and we believe that it is right that it is so. We recognize that multiple perspectives are also enriching PBF and its documentation. But yes, many practitioners are frustrated by the fact that their voice is not heard enough by scientists and impact evaluators. They are really looking for a balanced dialogue. This will only consolidate our common efforts.

My strong belief is that the top priority is that Sophie, you and any person you want to identify come to visit what we believe a well-designed and implemented PBF experience (e.g. Burundi). You will be able to make your

opinion on how holistic the intervention tries to be, appreciate the height of the challenges for robust evaluation... and hear about the side-effects of robust evaluation research designs.

Any human has also preconceptions and blinders on reality. This visit will also be a unique opportunity for you to help us to identify possible side-effects that we overlook. You will have to cover the costs of your visit, but we will find a team of national experts to welcome you and organize a full tour around the country. As a researcher, I also believe that as a profession, we have to give to empirical evidence its real place. This is a major question to be answered by those doing systematic reviews. The latter indeed convey a rather mechanical view of decision-making and policy process. Of course, informed-evidence policy is better, but the most important today in many LIC is to have well-thought, well-funded, well-implemented policies. Evidence will help in this process, but experience, expert opinion, political sensitiveness will be as much crucial. Furthermore, evidence is not limited to robust impact evaluation. I believe that while there is a hierarchy in rigor of the evidence, there is also a hierarchy in its relevance. Systematic reviews score less well on this other scale. I anticipate a debate among global health scholars at this level, because many scholars are uncomfortable with the current trend with “evidence-based policy”. Note that this view is not inspired by personal involvement in the ‘PBF program’; it is much more related to my studies on user fee removals and my activity to advise countries to design and implement correctly financial access policies. Sophie knows what I mean.

My personal feeling was that the Cochrane review came too early (in fact I was surprised by the number of studies you found rigorous enough). The good thing is that it allows us to have a thorough discussion now, quite at an early stage of the impact evaluation program. This will only strengthen what we try to do. I put Christel and Monique in cc to this important discussion. It indeed concerns the World Bank as well. So yes, we can have a discussion together, but at this stage of the process, we believe that there is first a need to rebalance the dialogue.

Note that while I am a scientist and from the North myself, I believe there is also a dimension of listening to stakeholders in Africa.

I encourage other members of this group to share their view on how this debate should be structured or to already contribute. This is a great opportunity for knowledge!

Best,
Bruno

From Sophie Witter (United Kingdom), August 10th, 2012

Hi Bruno,

As a health systems person, I can fully understand your frustrations with the Cochrane process. It has its strengths and limitations, like all methods. Its strength is that it provides a robust framework for assessing the impact of interventions on health outcomes. The weaknesses – it filters out more unorthodox study designs and process studies. However, this is not particular to PBF – this is an accepted approach to establishing evidence on impact and is well understood globally. It doesn’t claim to be trying to do anything else. It certainly doesn’t replace discussions about political processes, implementation modalities etc. Nobody – certainly not our study authors – would argue against a place for all sorts of different materials and debates on PBF, as for other health system strategies.

It may be a bit early for the evidence on PBF, but these reviews are updated periodically, so that should not be a problem in the longer term. There are a lot of new studies underway, as you know, so we expect more impact evaluation results in the coming years.

I think that PBF is a particularly tricky topic for systematic reviews as it has been interpreted and implemented in very different ways and very different contexts. These are all issues we highlight in the review. We call for more research on systemic effects and the relationship to different contexts. I don't think we would disagree on much there.

Happy to continue the dialogue.

Best wishes, Sophie

From Por Ir (Cambodia), August 10th 2012

Dear all,

After following up your reactions to the editorial and some answers from Sophie, the first author of the Cochrane review based on which the editorial was written, I would like to share a few words on this.

I am surprised to see many reactions to this editorial, but not to the Cochrane review itself when it was published. Moreover, most of the reactions tend to criticize that the authors undermined the potential (or even evidence of) impact of PBF. To me, the review was well written with very transparent Cochrane recommended methods and well balanced findings and conclusions. So, if we want to challenge the findings, we better challenge the Cochrane review methods (mainly for health system research), but not the authors.

One of possible limitations of Cochrane and other systematic reviews is that they are often based on available studies, especially those with strong methods and published, whereas in the real world there is no systematic studies on the intervention/strategy of interest, and many studies often fail to get published for a number of reasons: weak methods, no positive findings (as authors are often among the implementer and tend to show good results) or no capacity or interest in publication... Considering this, absence of evidence on the positive effects may not necessarily mean that the intervention/strategy of interest does not work, and the increasing evidence on positive effects may simply mean that there are more studies on this intervention/strategy, but not necessarily proportionally more schemes/programs that work. However, this is not an issue for clinical research for which one robust study with confirmed outcome of interest can allow making a strong case of evidence base.

We can also pose this question to the currently increasing evidence on the positive effects of PBF whether there are proportionally more PBF schemes/programs that work or just because there are more studies, and by change, more PBF schemes/programs that work have been picked up while at the same time many PBF schemes/programs which do not work are ignored. I cannot answer this question, but I can only confirm that there are insufficient evidence/studies on the PBF schemes/programs that do not work or negative effects (side effects) of PBF.

Regards,

Por

From Joanne Harnmeijer (Netherlands), August 17th, 2012

Dear Bruno

Following on your remark: “My personal feeling was that the Cochrane review came too early (in fact I was surprised by the number of studies you found rigorous enough)” I would like to add some observations:

Of the nine studies selected six were unpublished and of these three were mere evaluations. The majority of the studies did not claim to be a “rigorous” impact study nor were they submitted to peer reviewed journals. Not unexpectedly, the reviewers labeled the quality of the evidence presented in the studies as weak or very weak. What *is* surprising is that studies were admitted in the systematic review as if they satisfied the Cochrane criteria, which they clearly did not *and did not claim to either*. Also the 9 studies largely presented old data, of pre 2010. Lastly, I think it would have been appropriate for this review to acknowledge and use the results of the 2011 review on the same topic

(<http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD009255/abstract;jsessionid=92F31137C0965A07EAF8ED5C89AC89DE.d03t01>).

In my opinion the systematic review we are discussing here is turning things upside down and in doing so risks to be a disservice. One undesirable effect could be that current and future studies are increasingly rigorous to the point of intervening with the actual interventions, for the sake of academic rigour (read: RCT), but at the cost of regular OECD-DAC evaluation criteria (relevance, effectiveness and yes: impact ..).

Regards,

Joanne

From Sophie Witter (United Kingdom), August 17th, 2012

Dear Joanne,

For the systematic review, being published was not a criterion for inclusion, nor were any date restrictions set (all of this was laid out in the protocol which was published in 2009). That is normal for Cochrane reviews, which focus on gathering all relevant evidence, subject to meeting study design conditions and assessing risk of bias.

I understand your concern about an over-emphasis on impact evaluations to the detriment of other types of studies (which was also Bruno’s concern in his last email). However, I think that the risk is very low. There is a vast literature of all sorts of kinds now on PBF – theoretical, practical, etc. I don’t anticipate that changing. Why should it? There is a need for many different sorts of knowledge, in this as in other areas of health sector development.

All the best,

Sophie

From Joanne Harnmeijer (Netherlands), August 17th, 2012

Dear Sophie

Thank you for your fast reply.

I looked up the entry criteria for your systematic review. They include CBAs and ITS. I think my point still holds: authors of the studies (reports; papers) you decided to include as CBAs, or ITS, did not attempt to present their data as such. The label, if you wish, was yours. It is here that I sense a tension. (Occasionally one might have access to the raw data to decide if the studies would, after all, qualify, but that goes beyond the nature of a systematic review and has its own difficulties.)

Allow me a clarification: I wasn't referring to interference with studies (indeed, there is none), but rather to interference with the actual interventions, in current and future attempts to design and implement "better" studies, that is: of sufficient rigour. I sense a tension here. As Bruno hinted "... they have to nest their impact evaluation in the intervention: a real constraint for PBF implementation itself".

I would thus thoroughly agree with Atle's remark "If it were possible to engage in a debate on how PBF-schemes might be evaluated – in a practical and feasible, but robust way – I would find that interesting and helpful."

Sincerely,

Joanne

From Sophie Witter (UK), August 17th 2012

Dear Joanne,

Thanks for the clarification. I can't comment on the design of interventions, but I agree that they should be set up to be effective and sustainable, not just to be measurable!

We have a session scheduled at the Beijing symposium on ways of evaluating PBF, which aims to tackle the issues raised below (with a particular focus on evaluating impact on health systems). Your contribution would be welcomed, if you are able to attend, or through the CoP group before/after.

Bruno, Jurrien and others are participating.

All the best,
Sophie

From Atle Fretheim (Norway), August 17th 2012

A key question that some of you have raised is:

Should we have been more strict, and limited our review to only a few (none?) of the studies we included?

A common critique of Cochrane reviews is the opposite: "The inclusion criteria were too narrow and that the authors ignored many relevant studies" is typical. Critics saying "You should have excluded many more studies" is quite rare – and therefore more interesting, and fun!

In order to debate whether we included too many studies (or not), we need to address at least these two questions:

- a) What types of studies we should have opened up for?
- b) Was our classification of studies reasonable?

For a):

We adhered to what is more or less common practice in Cochrane reviews on health systems interventions: randomised controlled trials (RCTs), > non-randomised controlled before-after studies (CBAs) and interrupted time series (ITS) analyses are recommended study designs to include.

In our review the inclusion of such studies may have led to results that were more confusing than informative – perhaps. Still, I think it makes sense to include “weaker” designs, particularly when there are few or no RCTs that have been conducted on the topic. Happy to debate!

For b):

I am happy to discuss whether or not it made sense to label study X “RCT”, “CBA”, “ITS”, or “none of these”. But I don’t think the authors’ intentions or how they labeled their own study matters when we (review-authors) assess a study for inclusion in a systematic review. I think the opposite: That it would be even better if the reviewers don’t have that information!

Important point to make: We do not criticize authors of the original studies. It is not our intention, anyway. On the contrary! They may have conducted the best study possible given the circumstances. I would like to add that we are very thankful to many of the original authors. Several of them responded promptly to our e-mails and even sent us their full datasets.

I think Cochrane reviewers in general struggle much more than we did when trying to engage study authors and getting access to unpublished data. So, we are very thankful, indeed!

Cheers,

Atle,

From Joanne Harnmeijer (Netherlands), August 17th, 2012

Dear Atle,

Nice discussion, although a little intense.

My point was not so much that more studies ought to have been excluded, but that to measure impact one needs a design and thus an intention. Where I see the tension in your review is that the studies are first categorized as CBAs (or ITS) - that is: they "qualify" - and then their quality (as CBA, or ITS) is scored as low or very low.

Yet I think current and future researchers, (or, if you wish: evaluators), can learn much from your description of the studies' characteristics. As the majority (6) of the nine selected studies were, in your judgement, CBAs, and yet were either scored as high risk for bias (5) or unclear risk for bias (1): what can one do differently if the intention is to do a robust study? Same question for the 2 ITS studies. Your review nicely lists the common pitfalls.

Shortcuts to do studies that are both feasible and robust and do not interfere with the actual intervention may follow from there. But all this has to come from intention and design, and can not, or very rarely, succeed as an afterthought. I thus disagree with you on the point of labelling: researchers must know what data they need for what type of study. Only then can they decide if "robustness" is an option.

Best, Joanne

NB From now on I will keep quiet on this topic as I am new to this group and do not wish to come across as someone who thinks she has all the answers (I don't ..).

From Eric Bigirimana (Burundi/Cameroon), August 21st, 2012

Hi everyone, Hi Bruno, Hi Atle,

As regards the exchanges based on the Cochrane review, there are several points to be debated upon and I think this Atle and al. publication refers us to a lot of questions concerning assessment of PBF impact.

As regards the exchanges based on the Cochrane review, there are several points to be debated upon and I think this Atle and al. publication refers us to a lot of questions concerning assessment of PBF impact.

1) Atle and al. make an analysis of studies carried out on the PBF and conclusions drawn on the results obtained in different countries. I suppose that this criticism is based on the scientific precision applied at different stages of the research to measure the impact of PBF in low and middle income countries.

According to the Cochrane review, these researchers do not question the strategy itself. Their works do not show that the PBF strategy is not effective, but rather, that results produced up to date have not really proven its efficiency.

2) As Bruno indicated in his message to Atle et al., I am surprised by the number of studies Atle et al. claim to have revised and found rigorous enough. On my part, I think that the PBF is not sufficiently being supported by a research programme capable of drawing out the necessary conclusions in Africa. I would like institutions that support the implementation of the PBF in various countries to consider reserving a reasonable budget in order to document enough results gotten from the different experiences.

However, it has been noted that studies carried out often focus on quantitative aspects to the detriment of the qualitative meanwhile interventions are widely diverse depending on context and processes used in different countries. This causes a relevant problem as contextual aspects are overlooked whereas they are of utmost importance in explaining the success or failure of a PBF programme.

During the Study tour visits carried out by our BREGMANS teams in the three countries cited by Atle et al., (Burundi, Democratic Republic of Congo and Rwanda), the participants realised the importance of taking context into consideration to be able to know how to adapt the PBF to local realities. For example, the results accorded to the PBF are highly influenced by the situation of the health system in place and how it operates in the various countries. In addition, they also discovered that context can either act for or against the PBF, hence the need to adapt it to the circumstances found in each country. In turn, the PBF influences the different levers of the health system and vice versa. Interactions are observed between PBF and the distribution of human resources, management of information and NHIS, governance, management of drugs, and other financing mechanisms in health sector, etc. It is therefore abnormal to focus documentation only on quantitative aspects (mainly based on health care delivery) while ignoring a wave of information which could come from qualitative analysis, notably as concerns analysis on the PBF impact.

Even though the PBF does not pay the process indicators and compensates the results gotten, documenting the processes which led to the obtained results, either positive or negative, is indispensable. These procedures will greatly provide information on the factors that determine the success or failure of the PBF in the different contexts.

3) I understand the strict requirements made by Atle et al. review for as unbiased studies as possible. I hope that the researchers working on the PBF ought to take note of this in future.

Given the aspects listed in point 2), it is necessary to question the relevance of the Cochrane review in the presence of complex social interventions such as those in health sector. We know that the Cochrane review is

better indicated in empirical studies, especially clinical trials to evaluate drugs efficiency. It mostly concerns experimental trials where confounding factors are strictly controlled with the aim of establishing causality links between a proposed remedy and results obtained. This is different in social interventions such as public health. In these fields, where there is the influence of culture, behavioural aspects, socio-economic conditions and others, controlling confounding factors is not certain. Moreover, concerning the PBF in Africa, it is juxtaposed with other strategies in the domain. The contexts and situations are often in perpetual overlapping. This does not give the Cochrane review an appropriate opportunity to determine the direct causality link between a strategy taken in isolation and the results obtained. Although statistical techniques of seeing the contribution of each strategy exist, the factors which can be confounded keep increasing and increase the limits of interpretation.

In place of the **Cochrane review**, I will prefer the **Realist review** which searches for evidence while asking the following questions:

- What makes the PBF to work?
- For which population (target group) does the PBF work?
- In what circumstances does the PBF work?

According to David Gough et al. (2012)p.9 (see publication reference down below), Realist synthesis is a mixed methods review [9] that examines the usefulness of mid-level policy interventions across different areas of social policy by unpacking the implicit models of change, followed by an iterative process of identifying and analyzing the evidence in support of each part of that model. The difference between this method and more 'standard' systematic review methods is that the search for empirical evidence is more of an iterative, investigative process of tracking down and interpreting evidence. Realist synthesis will also consider a broad range of empirical evidence and will assess its value in terms of its contribution rather than according to some preset criteria. The approach therefore differs from the predominantly a priori strategy used in either standard 'black box' or in theory driven aggregative reviews.

The Realist review is then found midway between positivism which is strongly linked to facts which can be verified and reproduced through the methods used and results obtained (Objectivism inspiring the Cochrane review) and constructivism which is linked to mechanisms and processes which led to the outcomes in a given context (subjectivism inspiring the Narrative review).

In the case of the PBF in Africa, it would combine the analysis of the often quantitative data base and the analysis of the qualitative information found in countries which have tested the PBF approach. By this last point, I clearly understand Bruno's concern that researchers should reach the field to acquire experiences lived by experts and local stakeholders. I wish to mention that these experiences lived are not or are hardly documented within the purview of the PBF.

As Bruno did, I request researchers teams such as Atle et al., to continue such exchanges of views in our google group. This will even help the researchers of the South to improve our research capacities and reinforce the relationship between research and the choice of health policies.

For further information on Realist synthesis you can read:

Pawson R, Greenhalgh T, Harvey G & Walshe K (2004) Realist synthesis: an introduction: RMP Methods Paper 2/2004. Centre for Census and Survey Research, University of Manchester, Manchester.

Pawson, R. (2001) Evidence based policy: II. The promise of 'realist synthesis'. ESRC UK Centre for Evidence Based Policy and Practice: Working Paper 4.

www.evidencenetwork.org/cgi-win/enet.exe/biblioview?420

Gough D, Thomas J, Oliver S: **Clarifying differences between review designs and methods**, Systematic Reviews 2012, 1:28 doi:10.1186/2046-4053-1-28

From Sophie Witter (United Kingdom), August 22nd, 2012

I think that the realist evaluation approach is very interesting and agree that it would be well applied to PBF. For the Cochrane methods, if there are enough robust studies, then you can look for context patterns. Unfortunately, if you only have a few (as was the case for the PBF review), then that is ruled out. But I would just note that the methodology in itself is not unable to take into account contextual differences. We look forward to reading some realist evaluations of PBF by your group!

Best wishes,

Sophie

From Eric Bigirimana (Burundi/Cameroon), August 24th, 2012

Dear Sophie,

Thank you for your reply. Certainly, Realist synthesis approach could fit better with the PBF strategy. It takes context patterns into consideration and then responds to Cochrane review limitations. BREGMANS teams are working on new tools that can handle this kind of issues. We'll share information with other members of the group.

I agree with you, we need enough robust studies to draw conclusions on PBF impact. That's why I'm emphasizing on research funding which I find not supported enough for PBF in Africa. My point is that funding institutions should increase budget for research activities. That can shed light on various mechanisms involved in the purview of PBF.

Best regards,

Eric

From Robert Soeters (Netherlands), August 22nd 2012

Atle Fretheim's acknowledgement on the fact that Cordaid and SINA Health transparently shared non-published original study data with the Cochrane Review is much appreciated. We were also satisfied to note that 3 out of 4 of their World Health Bulletin WHO editorial references to original data came from us implying that the authors believed that those studies fulfilled the Cochrane Review criteria.

We shared these data with the Cochrane reviewers to demonstrate our general desire for transparency. Yet, as PBF practitioners we expected in response to receive constructive criticism on how to improve PBF and to advance the research agenda together instead of the generalized negative judgment on PBF. For PBF implementers, the reality on the ground since 2000 has been that there were few realistic research proposals from the academic world while we agreed at the same time that PBF could only advance when we would match new PBF interventions with a research component. We therefore engaged at our own initiative in intervention control studies and in action research on specific issues.

This initial strategy worked well in the sense that there was valuable feed-back and suggestions for improvement on our research from several academic persons but also from participants to workshops during which the results of the studies were being presented, analyzed and lessons learned. As the result, over several years we developed the PBF best practices, theories and instruments and we then put this together in PBF

course books followed by hundreds of participants, operational country manuals and toolkits. It resulted in a comprehensive set of common sense conclusions on how to implement PBF.

Already over 30 countries now largely apply this set of PBF best practices. So apparently there is something that resonates well in practice. For most health practitioners in low- and mid income countries it is clear that the status quo of doing things must change and that we have to move away from the traditional input strategies with a hierarchical government dominated health system. PBF is not a magic bullet nor a one trick approach but has become an ever changing and improving health reform strategy that is appealing to thousands of health providers, decisions makers and politicians.

It would have been helpful if our transparency of sharing data had been matched by the Cochrane review authors by first sharing their findings and analysis with the original authors in order to check on the definition of what is PBF, contextual factors and then on the appropriateness of their research methodology. The authors decided otherwise and published a paper without double-checking them with the field workers from which they obtained the data.

The result is a biased set of recommendations and some conclusions, which are communicated out of context. Several commentators already in the CoP PBF discussion during the last days have addressed the nuances that were overlooked in this way.

In summary:

1. The Cochrane review used a too limited definition of PBF that did not take into account the comprehensiveness of PBF as a healthreform approach. They mainly looked at a number of reproductive health indicators while PBF also includes quality of care, equity, decentralization, autonomous health facility management and sustainability and has selected indicators for those wider health reform elements.
2. The suggestion that there was little or no evidence for PBF in DR of Congo is not in line with the study published in Health Affairs, which was again confirmed in 2012 by workshops in Bukavu and Kinshasa whereby those directly involved with the PBF project confirmed that the findings were plausible and consistent.
3. The Burundi data analysis also showed significant better results in the PBF intervention provinces but somehow the Cochrane review came to a different conclusion. For that reason Cordaid recently submitted the same data set to the scrutiny of a different independent academic institution and their first assessment is that quality of care in Burundi significantly improved while they are still working on the other indicators. We therefore suggest that the Cochrane review may suffer from interpretation bias.
4. The WHB editorial suffers from selection bias whereby the authors zoomed in on two indicators (antenatal care and deliveries). Alternatively, if other criteria were chosen such as in Burundi quality of care (significant improvements) and in general sustainability (approximately 54% of PBF is financed by the government) another independent evaluator would have come to significantly better results in the PBF intervention area.
5. The authors did not take into account that the reproductive health indicators they selected for their study also happened in Burundi and the DR of Congo to be the main priority of the Emergency NGOS in the control areas to which they targeted their resources while in the PBF area a much wider set of objectives was being aimed at.
6. The Cochrane review failed to analyze important contextual factors in the DR of Congo. The authors should have known that in DR of Congo the control districts had four times more resources than the PBF districts (Soeters et al, Health Affairs 2011). Yet the overall results in the PBF districts were equal or better than in the control area so as a result the cost-effectiveness was firmly in favor of PBF. This convinced both local and national government to go ahead with PBF but this contextual factor was not flagged by the Cochrane review.

A solid critical discussion about PBF is highly welcome but while making sweeping statements that there is no or little evidence of PBF the Cochrane Review authors should also have explained what in their view should be

the alternative for the PBF health reform agenda. This is important in the context and flow of governments seeking new health reform approaches.

In short there is a need for a different approach on to how to conduct PBF research. We the PBF practitioners (and after intense consultations during the last 2 weeks) would hereby like to share a few suggestions for a more appropriate research agenda in ongoing and new PBF projects:

1. Research should involve a comprehensive set of predetermined indicators and taking into account the full characteristic of what is PBF instead of only zooming in on a few indicators.
2. This comprehensive research approach should be matched by a set of SMART research indicators to capture the much wider PBF health reform objectives. For example, instead of generally stating that we aim at sustainability we may propose an indicator such as that from total revenues, households should approximately contribute 8-12% for health expenditure either directly or through some form of fixed annual contribution. Instead of simply stating that PBF embraces good quality we should also investigate the price of quality and we know from our PBF projects in low-income countries that a health center covering 10,000 inhabitants should generate at least \$ 50,000 per year per person in cash revenues. This is approximately \$ 5.00 per capita per year but, currently, countries with traditional input systems only generate between \$ 0.50 and \$1.00 and therefore whatever a health center tries to achieve it will be too limited in scope and of very poor quality.
3. Research methodologies should not harm ongoing trends and flows in health reforms by proposing unrealistic designs for example for control areas just for the sake of conducting research. While some researchers may not believe there is evidence for PBF, the fact is that for implementers, decisions makers and politicians there is enough evidence simply because many PBF characteristics make sense, it is immediately applicable, and proposes a full toolkit of instruments and best practices that can be adapted to local circumstances. Not applying this is harmful for the development of the health system unless if researchers propose a similar comprehensive alternative for PBF. To the contrary, suggesting changes on sub-issues such as how to improve quality or sustainability are very welcome.
4. Full random sampling of intervention and control areas is probably not appropriate in the real world and can better be replaced by matching PBF intervention districts with purposely selected PBF control districts that resemble as much as possible the intervention districts. This can easily be done during the gradual roll out of PBF in a given country whereby at first a number of districts are included in the PBF intervention and that there are control districts that are not (yet) included in PBF.
5. When reviewing in how far a project can be included in PBF research it should be established in how far the project can be labeled as PBF. There is a quickly growing consensus in low- and mid-income countries on the definition of PBF and it should be avoided that studies are included of projects that do not qualify as PBF. It is unhelpful to study a project that is not PBF and then draw conclusions as if it is PBF. Cases in point were studies on Uganda (Palmer et al) and Zambia (Cochrane Review).
6. The context in the control areas should be described with the purpose to analyze to what the PBF intervention is being compared. This helps both practitioners and researchers to arrive at pragmatic, constructive recommendations on how to proceed with pointers to successes, challenges and failures. Sweeping conclusions are therefore not helpful when they are not put in context or are not fully analyzed with descriptions of the PBF area and the control areas and taking into account cost-benefit analysis.
7. Most of the above research approaches Cordaid already adopted for the 2010-2013 PBF study in the Central African Republic and of which the results are expected next year. We propose that more projects (including already ongoing ones) follow the above suggestions.
8. We would welcome if researchers make efforts to understand PBF with its comprehensive health reform characteristic by either visiting study areas or alternatively at least involve the grass roots implementers in their analysis and who produced the data. This to check for context factors and a good understanding of a given project – not for influencing their judgements.

We have started working on a paper with more details on the proposed research methodologies and indicators to be selected for such studies and taking into account the above suggestions. We invite the academic world to constructively engage on the PBF health reforms and thereby to improve the research agenda.

Warm regards to all, Robert Soeters

From Atle Fretheim (Norway), August 22nd 2012 – for an unknown reason, this mail did not reach the discussion forum

Robert (and other interested parties),

Yes, we are truly thankful that you so transparently shared your data with us. It was indeed a clear demonstration of desire for transparency. I salute you for that.

Now, let's move from praising you to trying to respond to your extensive criticism.

First, you write:

“Yet, as PBF practitioners we expected in response to receive constructive criticism on how to improve PBF and to advance the research agenda together instead of the generalized negative judgment on PBF.”

I think this criticism is unreasonable, considering what we wrote specifically about this in our review. The following quote from the Cochrane review is also relevant to those who criticised us for not being sufficiently interested in process evaluation, ignoring qualitative aspects, not addressing the “why” and “how”-questions, and/or not considering the wider health system impact of PBF:

“Implications for research

There is a need for high-quality research into PBF, looking at a range of modalities, scales of magnitude, levels of implementation within the health system, types of services and providers, and contexts. More robust study designs are needed to detect reliably the small to moderate, though potentially important, effects that PBF may be expected to have. And although there may be practical challenges with implementing randomised experiments of PBF, recent studies have demonstrated that it can be feasible to conduct such trials. Robust effectiveness evaluations should be complemented by in-depth process evaluations to uncover the mechanisms by which the intervention may or may not work, and to probe the motivational effects which are intended to be at the core of the intervention.

....

The equity impact of PBF schemes is another of the many under-researched areas - what are the differential impacts on different social groups? As noted, evaluations should take a broad perspective and consider wider health systems effects, intended or unintended. No organizational impacts were reported in the included studies, despite the view that PBF can increase managerial autonomy.”

“It would have been helpful if our transparency of sharing data had been matched by the Cochrane review authors by first sharing their findings and analysis with the original authors in order to check on the definition of what is PBF, contextual factors and then on the appropriateness of their research methodology. The authors decided otherwise and published a paper without double-checking them with the field workers from which they obtained the data.”

It is not common practice to ask original study authors to read through a Cochrane review, but we could have elected to double check various things - sure. As far as I recall we did write to several of the authors asking for some clarifications. I agree that if misunderstandings or errors could have been avoided, we should have contacted authors. Otherwise I am not convinced that it would be a very useful approach.

As of yet, no one has pointed to factual errors in our review (we are keen to learn about errors if you find them).

We did not find any studies on contextual factors that we deemed relevant. However, our search strategy may have failed. We would very much like to see process evaluation reports, qualitative explorations etc. conducted alongside/embedded in the impact studies that were included in the review (and by that I mean written articles or reports, using some reasonable methodology e.g. surveys, interviews etc). If such reports do not exist, you

can not criticise us for not including them! If they do exist, we want to look at them, and possibly include them in an updated version of the review. If you can show that our search for such reports has been sloppy, I will apologise!

“The Cochrane review used a too limited definition of PBF that did not take into account the comprehensiveness of PBF as a healthreform approach. They mainly looked at a number of reproductive health indicators while PBF also includes quality of care, equity, decentralization, autonomous health facility management and sustainability and has selected indicators for those wider health reform elements.”

We used a mainstream definition of PBF – or so we thought. I have no strong feelings attached to PBF-definitions and I am not sure that it matters much anyhow: We looked for studies on practically all the indicators/outcomes you mention, but didn’t find them. Where are they?

“The suggestion that there was little or no evidence for PBF in DR of Congo is not in line with the study published in Health Affairs, which was again confirmed in 2012 by workshops in Bukavu and Kinshasa whereby those directly involved with the PBF project confirmed that the findings were plausible and consistent.”

I disagree with your interpretation of the DRC-data. I find it less convincing than you do. We have in detail explained why in the Cochrane review. I am happy to debate the specifics points – if there are specific points in our assessment where you disagree. I am not familiar with the discussions in the DRC-workshops you mention. I think it is encouraging that those working “hands-on” in PBF-programmes are convinced that they are successful, but this does not constitute “evidence” in the conventional sense. I am not particularly interested in initiating a “What is evidence?”-debate, but I am willing to take part if you push for it.

“The Burundi data analysis also showed significant better results in the PBF intervention provinces but somehow the Cochrane review came to a different conclusion. For that reason Cordaid recently submitted the same data set to the scrutiny of a different independent academic institution and their first assessment is that quality of care in Burundi significantly improved while they are still working on the other indicators. We therefore suggest that the Cochrane review may suffer from interpretation bias.”

Interesting. Can you send me the report from the independent academic institution? No doubt – we may well have made mistakes in our review (but as I have pointed out: No one has made us aware of any errors so far). I don’t see how interpretation bias can have changed an estimate from positive to negative. My guess is that we (or you!) have made a mistake. Would be interesting to clarify what the cause of the discrepancy is.

“The WHB editorial suffers from selection bias whereby the authors zoomed in on two indicators (antenatal care and deliveries). Alternatively, if other criteria were chosen such as in Burundi quality of care (significant improvements) and in general sustainability (approximately 54% of PBF is financed by the government) another independent evaluator would have come to significantly better results in the PBF intervention area. “

You have liberal interpretation of the term “selection bias”! I chose to focus on the utilization outcomes that were reported in more than one trial. That’s a choice, and in my mind a sensible one. Those outcomes illustrate nicely the diversity of the results in general. Hand picking results that show how bad or good PBF is – now that might qualify as “selection bias”. It’s not fair to say that we did that, I think. We just showed that they are “ all over the place”. They are!

“The authors did not take into account that the reproductive health indicators they selected for their study also happened in Burundi and the DR of Congo to be the main priority of the Emergency NGOS in the control areas to which they targeted their resources while in the PBF area a much wider set of objectives was being aimed at.”

You may have a point here. I’d have to dig into your reports (and as you know, they are very long and often in French, so it takes me some effort to read through them – I’m not very keen on doing that now). Still, if your point is that important interventions were going on in control areas that were not going on in the intervention

areas, then the basis for believing in the control groups as representing a fair comparison is lost (thus the request for RCTs, which would reduce – though not eliminate – the risk of this happening).

“The Cochrane review failed to analyze important contextual factors in the DR of Congo. The authors should have known that in DR of Congo the control districts had four times more resources than the PBF districts (Soeters et al, Health Affairs2011). Yet the overall results in the PBF districts were equal or better than in the control area so as a result the cost-effectiveness was firmly in favor of PBF. This convinced both local and national government to go ahead with PBF but this contextual factor was not flagged by the Cochrane review.”

You may have a point here too. This is the type of contextual information that may be of some value, perhaps (I’m not sure). But speculation that the PBF-intervention would have been proven more effective had the funding been similar, is not very convincing – I think.

“A solid critical discussion about PBF is highly welcome but while making sweeping statements that there is no or little evidence of PBF the Cochrane Review authors should also have explained what in their view should be the alternative for the PBF health reform agenda. This is important in the context and flow of governments seeking new health reform approaches. “

I disagree. Our job was to assess the evidence base for PBF, not advise policymakers what to do. We implicitly suggest that PBF should be developed further, but that more evaluation is needed before we can feel confident about its usefulness as an approach for health reform. A very logical conclusion based on what we found in our review, no?

“In short there is a need for a different approach on to how to conduct PBF research. We the PBF practitioners (and after intense consultations during the last 2 weeks) would hereby like to share a few suggestions for a more appropriate research agenda in ongoing and new PBF projects:

Research should involve a comprehensive set of predetermined indicators and taking into account the full characteristic of what is PBF instead of only zooming in on a few indicators. “

Couldn’t have put it better myself. Totally in agreement!

“This comprehensive research approach should be matched by a set of SMART research indicators to capture the much wider PBF health reform objectives. For example, instead of generally stating that we aim at sustainability we may propose an indicator such as that from total revenues, households should approximately contribute 8-12% for health expenditure either directly or through some form of fixed annual contribution. Instead of simply stating that PBF embraces good quality we should also investigate the price of quality and we know from our PBF projects in low-income countries that a health center covering 10,000 inhabitants should generate at least \$ 50,000 per year per person in cash revenues. This is approximately \$ 5.00 per capita per year but, currently, countries with traditional input systems only generate between \$ 0.50 and \$1.00 and therefore whatever a health center tries to achieve it will be toolimited in scope and of very poor quality. “

Fine. Although I personally think that patient-important outcomes are the most important (but anyone can have their own opinion on what they think is important and what PBF is trying to achieve).

“Research methodologies should *not* harm ongoing trends and flows in health reforms by proposing unrealistic designs for example for control areas just for the sake of conducting research. While some researchers may not believe there is evidence for PBF, the fact is that for implementers, decisions makers and politicians there *is* enough evidence simply because many PBF characteristics make sense, it is immediately applicable, and proposes a full toolkit of instruments and best practices that can be adapted to local circumstances. Not applying this is harmful for the development of the health system unless if researchers propose a similar comprehensive alternative for PBF. To the contrary, suggesting changes on sub-issues such as how to improve quality or sustainability *are* very welcome. “

Our disagreement is so massive on this point that a discussion seems meaningless. I am not convinced about the effectiveness of PBF – after looking at the data. Thus, I want more convincing evidence. You are already convinced and therefore don’t understand why we need more evidence. Still, you may want to contribute to more and better evidence in order to convince others about the usefulness of PBF?

“Full random sampling of intervention and control areas is probably not appropriate in the real world and can better be replaced by matching PBF intervention districts with purposely selected PBF control districts that resemble as much as possible the intervention districts. This can easily be done during the gradual roll out of PBF in a given country whereby at first a number of districts are included in the PBF intervention and that there are control districts that are not (yet) included in PBF.”

In principle – and also in practice, at least on some cases – a gradual roll-out of a programme can be done in a randomised way (there are examples). But I know that there may be difficult to do this – for a variety of reasons. I would be disappointed to learn that the researchers/evaluators themselves represent a barrier to trying to carry out randomised trials of PBF. Shouldn't we aim for that, when feasible? Regarding the ethics argument: I don't think it is more ethical to use “matching” to select control districts than to use the flip of a coin. On the contrary.

“When reviewing in how far a project can be included in PBF research it should be established in how far the project can be labeled as PBF. There is a quickly growing consensus in low- and mid-income countries on the definition of PBF and it should be avoided that studies are included of projects that do not qualify as PBF. It is unhelpful to study a project that *is not* PBF and then draw conclusions as if it *is* PBF. Cases in point were studies on Uganda (Palmer et al) and Zambia (Cochrane Review).”

If you (or others) could explain why you don't think the Zambia study that we included is a PBF-study, please do!

“The context in the control areas should be described with the purpose to analyze to *what* the PBF intervention is being compared. This helps both practitioners and researchers to arrive at pragmatic, constructive recommendations on how to proceed with pointers to successes, challenges and failures. Sweeping conclusions are therefore not helpful when they are not put in context or are not fully analyzed with descriptions of the PBF area and the control areas and taking into account cost-benefit analysis. “

Context should be more extensively described in general – I agree. I am puzzled by your use of the term “sweeping conclusions” (several times), since our conclusions are mainly an expression of great uncertainty, and little more. But I am not interested in a semantic discussion, so it's not a big deal.

“Most of the above research approaches Cordaid already adopted for the 2010-2013 PBF study in the Central African Republic and of which the results are expected next year. We propose that more projects (including already ongoing ones) follow the above suggestions. We would welcome if researchers make efforts to understand PBF with its comprehensive healthreform characteristic by either visiting study areas or alternatively at least involve the grass roots implementers in their analysis and who produced the data. This to check for context factors and a good understanding of a given project – not for influencing their judgements. “

I would re-phrase: Those who evaluate PBF programmes should ensure that they also include context factors in their reporting/description of the programme under evaluation. This to enable authors of systematic review to integrate such information in the analysis.

“We have started working on a paper with more details on the proposed research methodologies and indicators to be selected for such studies and taking into account the above suggestions. We invite the academic world to constructively engage on the PBF health reforms and thereby to improve the research agenda.”

I worry that relying on non-randomised comparison groups (e.g. Districts A and B introduced PBF and are compared with districts C and D that don't have PBF) will cause more confusion than clarification, and thereby prevent the field from moving forward as fast as it could. As demonstrated in most of the studies we included (most were non-randomised comparisons), the control areas were so different from the intervention areas that drawing any conclusions based on the data at all, was questionable. This was a major reason for us to downgrade the quality of the evidence, and in the end conclude that we have no little about the impact of PBF. The

inability to draw firm conclusions from PBF-studies will likely be repeated over and over again unless more rigorous approaches are used widely when PBF is evaluated (in the editorial we suggest randomised studies and/or interrupted time-series for measuring the effect of PBF-programmes).

Cheers,
Atle

On August 25th, after the announcement by a member of the approbation by the Burundian parliament of an IDA grant of US\$ 14,8 million for PBF, the non-facilitated discussion spontaneously took another direction (sustainability of PBF, in French first).