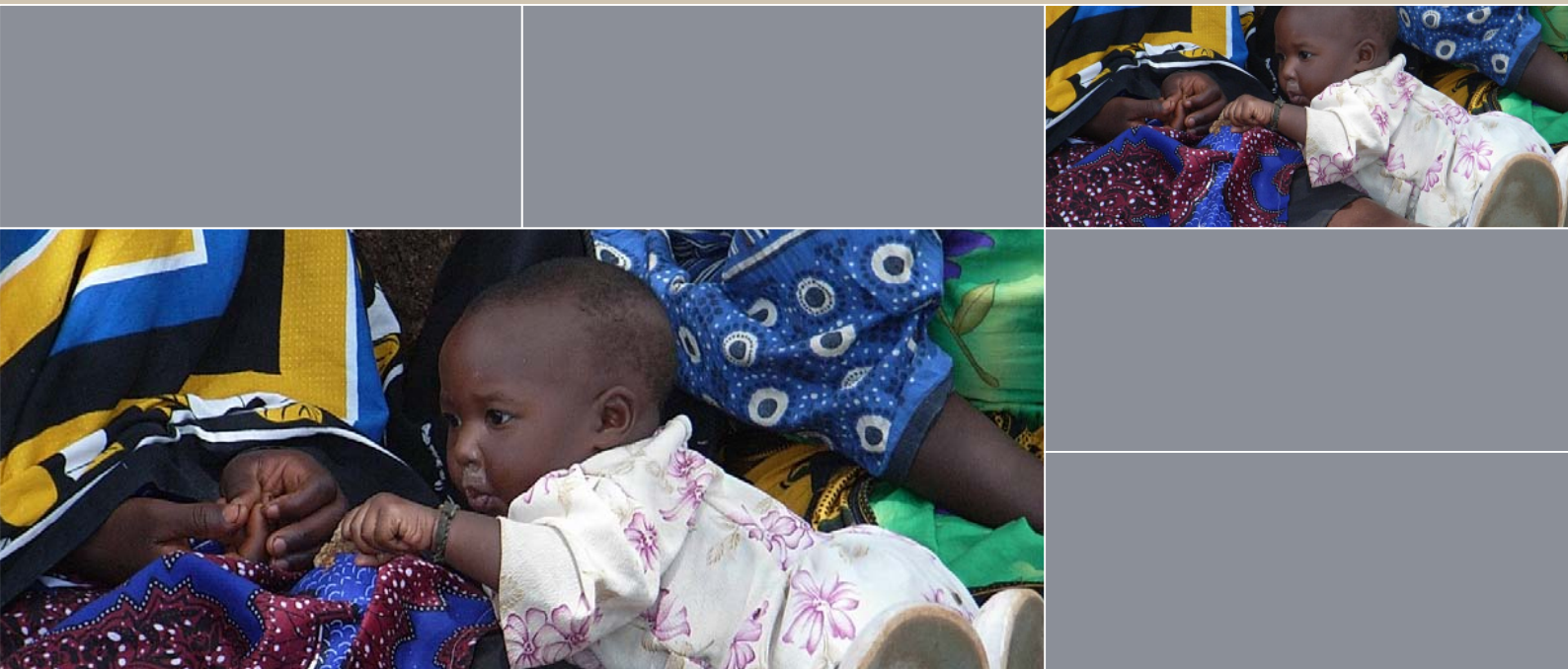




Evaluability Study of Partnership Initiatives

Norwegian Support to Achieve Millennium Development Goals 4 & 5

Report 9/2010 – Study



Norad

Norwegian Agency for
Development Cooperation
P.O.Box 8034 Dep, NO-0030 Oslo
Ruseløkkveien 26, Oslo, Norway

Phone: +47 22 24 20 30

Fax: +47 22 24 20 31

Photos: From Tanzania. Eivind Lund, Norad

Design: Agendum See Design

Print: 07 Xpress AS, Oslo

ISBN: 978-82-7548-543-2

Evaluability Study of Partnership Initiatives

Norwegian Support to Achieve
Millennium Development Goals 4 & 5

February 2011

Mott MacDonald Limited/HLSP

Beth Plowman
Henry Lucas

Contents

List of Abbreviations	v
Executive Summary	ix
1. Study Background	3
1.1 The Partnership Initiatives (PI)	3
1.2 Study objectives, scope and methods	3
1.3 Assumptions and limitations	5
1.4 Acknowledgements	5
2. Assessment of Evaluability	6
2.1 Preliminary observations	6
2.2 Team’s approach and assessment	8
2.2.1 Evaluating the impact of the Partnership Initiatives “programme” versus specific interventions.	9
2.2.2 Ensure timely and consistent Partnership Initiatives documentation	10
2.2.3 Systematic Implementation-level Monitoring	11
2.2.4 Use of Comparison Groups	11
2.2.5 Range of evaluation options	13
3. Main Recommendations	15
3.1 General Recommendations	15
3.2 “Best Case” Options for each Partnership Initiative	19
Appendix A: Terms of Reference	25
Appendix B: Persons Interviewed	28
Appendix C: Revised Work Plan and Timeline	29
Appendix D: Bibliography	30
Appendix E: Norway - Pakistan Partnership Initiative Profile	37
Appendix F: Norway - Pakistan Partnership Initiative Monitoring & Evaluation Plan	41
Appendix G: Norway - Pakistan Partnership Initiative Evaluation Options Appraisal	53
Appendix H: Norway - Malawi Partnership Initiative Profile	64
Appendix I: Norway - Malawi Partnership Initiative Monitoring and Evaluation Plan	67
Appendix J: Norway - Malawi Partnership Initiative Evaluation Options Appraisal	70

Appendix K: Norway - India Partnership Initiative Profile	75
Appendix L: Norway - India Partnership Initiative Monitoring & Evaluation Plan	80
Appendix M: Norway - India Partnership Initiative Evaluation Options Appraisal	96
Appendix N: Norway - Nigeria Partnership Initiative Profile	108
Appendix O: Norway - Nigeria Partnership Initiative Monitoring and Evaluation Plan	111
Appendix P: Norway - Nigeria Evaluation Options Appraisal	117
Appendix Q: Norway - Tanzania Partnership Initiative Profile	121
Appendix R: Norway - Tanzania Partnership Initiative Monitoring and Evaluation Plan	124
Appendix S: Norway - Tanzania Partnership Initiative Evaluation Options Appraisal	131

List of Abbreviations

AHHA	Department of Global Health and AIDS, Norad
ANC	Antenatal Care
DFID	Department for International Development (of the UK)
DLHS	Demographic and Life Health Survey
EVAL	Norad Evaluation Department
GoI	Government of India
GoM	Government of Malawi
GoN	Government of Nigeria
GoP	Government of Pakistan
GoT	Government of Tanzania
HMIS	Health management information system
JSC	Joint Steering Committee of NIPI
JSY	Janani Suraksha Yojana
KFW	German Development Bank
MDGs	Millennium Development Goals
MFA	Ministry of Foreign Affairs
MNCH	Maternal, newborn and child health
MOU	Memorandum of Understanding
M&E	Monitoring and evaluation
NIPI	Norway-India Partnership Initiative
NGO	Non Governmental Organisation
NOK	Norwegian kroner
Norad	Norwegian Agency for Development Cooperation
NMPI	Norway- Malawi Partership Initiative
NNPI	Norway-Nigeria Partnership Initiative
NPPI	Norway-Pakistan Partnership Initiative
NRHM	National Rural Health Mission
NTPI	Norway-Tanzania Partnership Initiative
PI	Partnership Initiative
PMG	Program Management Group of NIPI
PRRINN-MNCH	Partnership for Reviving Routine Immunization in Northern Nigeria; Maternal, Newborn and Child Health Initiative
RBF	Results-based financing
RNE	Royal Norwegian Embassy
UN	United Nations
UNAIDS	United Nations Program on HIV & AIDS
UNDP	United Nations Development Programme
UNICEF	United Nations Children Fund

UNFPA United Nations Fund for Population Activities
UNOPS United Nations Office for Project Services
WHO World Health Organisation

Executive Summary



Executive Summary

Study Background

In September 2007, the Norwegian Government under the leadership of its Prime Minister launched the Global Campaign for the Health Millennium Development Goals (MDGs). The Global Campaign embodies a multi-country commitment to find better ways of achieving value for money and ensuring that the most vulnerable groups have access to essential services. Norway's support focuses particularly on MDG 4 to reduce child mortality and MDG 5 to improve maternal health and includes the Partnership Initiatives (PI) to support their achievement. The Partnership Initiatives take the form of bilateral cooperation agreements with countries with high child mortality rates including India, Pakistan, Malawi, Nigeria and Tanzania. Managed by the Royal Norwegian Embassies (RNE), the Partnership Initiative has committed NOK 1,225,000,000 (US\$ 205 million) through agreements with a range of partners in these countries.

The means of evaluating progress towards the MDGs is receiving close scrutiny and will continue to draw attention as the 2015 achievement date for targets draws near. It is therefore apropos that Norad would seek an Evaluability Study for the Partnerships Initiatives (PI) – the Norwegian channel of bilateral support to these global goals. The purpose of this study is to assess the extent to which the Partnership Initiatives can be evaluated in a reliable and credible manner and to make recommendations and propose action plans for impact evaluations to be conducted in the PIs at a later stage.

Between August and December 2010, a two person team conducted interviews, reviewed documents and analyzed baseline data sets. For each of the five Partnership Initiatives, the Team prepared PI Country Profiles and the Evaluation Options Appraisals. The PI Country Profiles serve to consolidate basic variables for each Partnership Initiative in a comparable format (e.g. duration, objectives, funding, partners, geographic focus and criteria for selection). The Appraisals present the Team's consideration of a basic set of parameters for impact evaluation in each PI country. As part of that appraisal, a set of recommendations and options for evaluation were developed for each PI. These options and recommendations were subsequently reviewed by Norad and by the Royal Norwegian Embassies, and feed back was provided to the Team based on which this final report has been prepared.

In the initial appraisal of options for evaluation the Team made a number of observations that helped to form the basis of the study's approach. Among those observations:

- There is considerable diversity between the five PI country programs in terms of implementation status and arrangements, partners, timelines and technical emphases. Any proposed impact evaluation design needs to be tailored to each PI and its context. As a consequence, comparability of outcomes between PI countries will be limited.
- As far as the Team could determine, the Partnership Initiative lacks an overarching strategy or framework document that defines its statement of purpose, objectives and the means of achieving them. This makes it difficult to develop and use common evaluation criteria on principles or priorities across countries participating in PIs (e.g. there is no common reference point or definition for the frequent references to innovation and risk-taking within Partnership Initiative documents).

Study approach

The Team sought to build on recent, relevant experiences in global health programme evaluation. Recent experience in evaluating large-scale child health initiatives suggests that traditional designs (intervention vs. comparison areas) are increasingly limited in their ability to isolate and evaluate specific program effects given the many parallel initiatives underway. The Team therefore emphasized a set of basic principles and approaches throughout the Options Appraisals. These principles and their relevance to the Partnership Initiatives include:

- *Evaluating the impact of the PI “programme” versus specific interventions.* For a number of reasons, it will be difficult to evaluate the impact of the Partnership Initiative as a comprehensive programme in some countries. However within each PI, there are clearly interventions and innovative elements that should be subject to rigorous evaluation to demonstrate their effects on maternal and child service delivery and outcomes. In other words, the evaluability of specific PI interventions is important and attainable while the evaluability of an individual PI “programme” as a whole is less feasible.
- *Ensure timely and consistent PI documentation.* Systematic documentation describing key elements and intensity of a PI over time, as well as features of the overall context that may affect program objectives, is essential to good evaluation. This documentation should also include the PI’s logical model/ framework/causal chain (i.e. inputs, activities, outputs, outcomes) depicting how the PI is expected to achieve its purpose. Few PI were found to be conducting the on-going systematic documentation needed to support evaluation.
- *Systematic implementation-level monitoring.* The Team emphasizes this function as the lynchpin to maintaining an evidence base able to convincingly link processes with outputs and outcomes. This form of monitoring would provide regular, valid measures of specific PI inputs and processes including training, supervision, and delivery channels. In addition, implementation-level monitoring can also provide needed information on the timing, intensity and quality of programme implementation. The Team found that few PIs are currently tracking implementation in a manner that would support impact evaluation.
- *Use of comparison groups.* One major challenge to impact evaluability is the difficulty or impossibility of selecting an “untouched” comparison group in settings where levels of development assistance for health have risen sharply over the last decade. Increased development assistance has brought in new

interventions and stakeholders that often act as confounding variables to impact evaluability. Some PI countries have selected neighbouring districts as comparison groups in their baseline data collection where the said confounding variables should be taken into account. The Team suggests in the main report and in the PI country annexes several design and analytical approaches to enhance the use of comparison groups in PI evaluation efforts.

Range of evaluation options

The Team considered a range of options for the evaluation of the Partnership Initiatives in each of the five countries. Very briefly, the three categories are:

- the “*gold standard*” evaluation design, which entails random assignment of either individuals or groups to either a treatment or control group. This option would require that key variables are defined in advance of the intervention to facilitate the random assignment procedure. Efforts are made to limit or control the influence of external factors – including implementation of similar interventions occurring in the control area.
- a *global health – normative* approach can encompass a number of differing designs but, at a minimum, this approach requires a justifiable effort to account for external factors. A common variant of the approach entails “before and after” measurements in a programme intervention area and a comparison area – where presumably the programme intervention will not be implemented. A number of design features can strengthen the rigor of this approach.
- a *minimal option* focused on achievement of pre-established performance targets or criteria without accounting for external influences (e.g. via comparison groups).

Recommendations

In order to improve the chances of evaluability across all Partnership Initiatives, the Team recommends that those responsible for the Partnership Initiatives:

- a) *Clarify logical models and revisit assumptions* – the Partnership Initiatives operate largely without clearly articulated logical models. In some cases, only general explanations are provided on how PI resources and partners are expected to work together through defined processes to achieve a set of intended outputs and outcomes. This degree of generality offers little guidance to those designing monitoring and evaluation procedures. In terms of timing, it would be ideal for new PIs to be supported by experienced evaluation staff in order to a) develop and/or update their logic model after the completion of a baseline survey, b) analyse other available data and c) establish a set of measurable programme targets. In the case of PIs which are further in terms of implementation, the development or review/ update of a logic model could be considered as part of mid-term review efforts.
- b) *Recognize and address issues surrounding attribution* - PI documents often do not fully account for the complex donor landscape in-country or attempt to locate the role of the PI within this larger context. While there are exceptions, it would benefit the Partnership Initiative as a whole if these issues were clearly and consistently incorporated into their guiding documents.

- c) *Adhere to internationally-accepted standards for monitoring MDGs 4 and 5* – At country and global levels, MDGs 4 and 5 are tracked using a consistent set of internationally-agreed upon indicators. Three of the internationally-agreed MDG 4 and 5 indicators are recommended for use across Partnership Initiatives: the proportion of infants immunized against measles, the proportion of births attended by skilled health personnel, and ANC coverage. On a positive note, program priorities of all Partnership Initiatives are fairly well-aligned to these measures. However, there is wide variation in the actual indicators selected which raises some questions. Primarily, would it be more beneficial for Norwegian support for MDGs 4 and 5 to consistently use a set of internationally-agreed indicators across all supported countries?
- d) *Revisit and refine the monitoring and evaluation plans* –in several cases PI (M&E) monitoring and valuations plans were outlined as an initial element of the program document but then never further elaborated or made operational. Almost every PI has an over-abundance of indicators, many of which are neither measurable, attributable nor adequately defined. PIs which are mid-stream in implementation should take the opportunity to update their initial M&E plans while newly developed PIs should carefully develop their M&E plans with a view towards streamlining the monitoring process and reducing the burden of collecting unnecessary data.

In terms of PI-specific recommendations, the Team has outlined a “best case” option for each. The “best case” options vary considerably from country-to-country but each require that the PI managers take some form of immediate action to ensure the most robust final evaluation of their initiative.

Norway-Pakistan Partnership Initiative (NPPI): The *global health – normative* option provides the best case to meet internationally-accepted standards and provide evidence for maternal, newborn, and child health policy dialogue and program development. The Team proposes that efforts focus particularly on a) contracting out of maternal, newborn, and child health services to the private sector; and b) generating demand through the use of vouchers/incentives. This option would entail evaluation in the 10 project districts where baseline data has already been collected. The evaluation should emphasize measures of program duration and intensity of implementation for either individual women or villages, and associate that exposure with the desired outcomes. The design should also capitalize on NPPI’s sequential introduction across districts (as activities are not initiated in all ten districts simultaneously) in a “pipeline” analysis”. In a “pipeline” analysis, groups which are targeted to receive the intervention but not yet covered can serve as an internal comparison group. This method would allow for an internal comparison group within the ten districts. Finally, this approach could also include consistent and systematic monitoring of implementation across agencies involved in the project.

Norway-India Partnership Initiative (NIPI): A “minimal” option evaluation might best serve the NIPI programme by triangulating a wealth of available data sources (DLHS-3, NIPI baseline, UNFPA and UNICEF coverage surveys, NRHM routine data, health information system data as well as the upcoming DLHS-4) in order to a)

determine with known certainty whether NIPI targets were achieved (y/n) and b) whether statistically significant change occurred in specified outcome measures over time. This option could include the planned NIPI mid-line and end line survey data, in a greatly reduced template focused on a smaller number of key indicators.

It is further recommended that specific, targeted interventions are evaluated for their effect on desired outcomes (e.g. to what extent do Accredited Social Health Activist (ASHA) activities increase the utilization of home-based newborn care practices?). This is best conducted through a set of relatively small, well-designed studies and could help guide National Rural Health Mission (NRHM) investments and direction. These investments would provide relevant information about specific interventions supported through NIPI and better reflect its contribution.

Norway-Malawi Partnership Initiative (NMPI): The NMPI is well-positioned to conduct a global health - normative option and can still incorporate important design elements from its inception. The Team would encourage the evaluation designers to take careful note of existing international experiences and approaches for evaluation of results-based financing – notably the required timeframe and required resources. As part of the initial monitoring and evaluation design, a strategy and plan for documentation should be developed and contracted. As the NMPI sites were chosen for their better than average conditions, the PI partners should assiduously document the conditions required for successful performance (e.g. service accessibility and quality) with an eye towards how those conditions will be made available in other areas and program effect replicated.

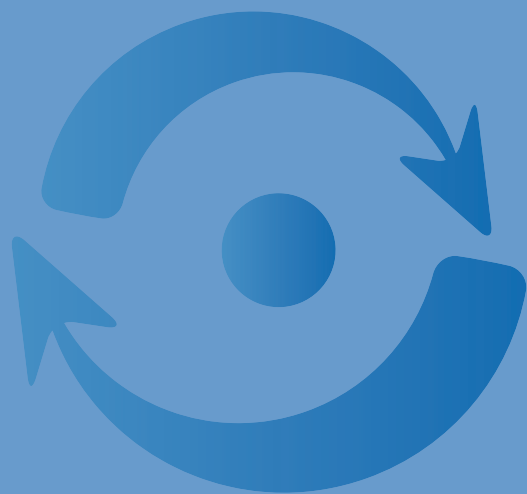
The Team would also recommend caution about M&E goals and timelines which seem far too ambitious for a pilot project expected to be evaluated after 2 to 3 years. Program designers are encouraged to moderate their expectation for the M&E component to a more achievable set of aims. Finally, the program designers are strongly encouraged to work closely with the “RBF for Health impact evaluation” network (part of a Norwegian funded initiative based at the World Bank to support results-based financing innovation in eight countries). There is a readily available set of tools and materials to assist with the further development and refinement of the evaluation approach.

Norway-Tanzania Partnership Initiative (NTPI): The majority of the NTPI investment is directed through the pooled funding mechanisms and assessed according to jointly agreed procedures. However, as with the PI in Malawi, the NTPI is well-positioned to conduct a global health - normative option around the pay-for-performance (P4P) component. As this component and its evaluation are still being developed, it is possible to incorporate important design elements from the beginning. The Team recommends that the program be realistic and focused on the evaluation design and requirements for data collection. For example, since the timeline for the pilot has been compressed to just 18 months the evaluation design should accommodate this truncated implementation schedule. The idea of using a “facilities readiness” measure is a very good one but will need sufficient time to be developed, tested and garner buy-in from national decision-makers (as there are clear implications for the national P4P programme). The NTPI is encouraged to work

closely with the RBF for Health impact evaluation network (referred to earlier in the case of Malawi) and to draw on the readily available tools and materials to assist in the design of the evaluation. Finally, since a few selected districts are benefitting from a Health management information system (HMIS) improvement scheme, the NTPI ought to describe minimum requirements for HMIS operation during any proposed scale-up of the pay-for-performance pilot, recognizing that not all districts will have benefitted from the improved HMIS.

Norway-Nigeria Partnership Initiative (NNPI): The Partnership Initiative in Nigeria is already following a global health – normative approach to evaluation. The aim of that evaluation is to measure change over time in specified outcome indicators in both cluster and non-cluster areas within each state. The Team’s concern, is that ‘success’ at this level, given the dysfunctional nature of services prior to the intervention, will need to be carefully interpreted in terms of scaling up to state level. We would strongly recommend that this is a case where the use of implementation data to link inputs to intermediate outputs/outcomes in order to analyse the implementation process could be of greater value than a simple assessment of project success.

Main Report



1. Study Background

1.1 The Partnership Initiatives (PI)

In September 2007, the Norwegian Prime Minister played a lead role in the launch of the Global Campaign for the Health Millennium Development Goals (MDGs). The Global Campaign embodies a multi-country commitment to find better ways of achieving value for money and ensuring that the most vulnerable groups have access to essential services¹. Norway's support focuses particularly on MDG 4 on reducing child mortality and MDG 5 on improving maternal health and includes the Partnership Initiatives (PI) to support their achievement. The Partnership Initiatives take the form of bilateral cooperation agreements with countries with high child mortality rates. Managed by the Royal Norwegian Embassies (RNE), this cooperation includes significant Norwegian support for the effort to reduce child and maternal mortality in these countries.

1.2 Study objectives, scope and methods

The purpose of this study is to assess the evaluability² of Partnership Initiatives (PIs) in India, Tanzania, Nigeria, Pakistan and Malawi. The evaluability study is intended to produce recommendations and proposed action plans for impact evaluations to be conducted in the five PIs at a later stage. For each of the five Partnership Initiatives, the study aimed to:

- Map the basis for Norway's support to the PI using project documents and assess the adequacy and quality of such documentation for the purposes of impact evaluation at end of the initial implementation period (typically 5-6 years)
- Identify contextual factors likely to influence PI impacts, both intended and unintended, and assess the availability, adequacy, and quality of contextual factor data/information currently available, going back at least 5 years in time as possible
- Review currently-available baseline data/studies and assess their appropriateness in terms of reliability and validity, with reference to the following questions:
 - What are the “right” impacts to be measured/verified?
 - Are the “right” impacts verifiable to acceptable standards given the existing baseline information?
 - Where information is lacking or incomplete, can data be (re)constructed?
 - What counterfactuals need to be considered?

1 The Global Campaign for Health Millennium Development Goals. Accessed at: <http://www.norad.no/en/Thematic+areas/Health+and+aids/Maternal%2C+child+and+women%27s+health/Global+campaign+for+the+health+MDGs>.

2 As per the OECD/DAC Evaluation Glossary, *evaluability* is the extent to which an activity or a program can be evaluated in a reliable and credible fashion. The assessment of evaluability involves an early review of the proposed activity in order to determine whether its objectives are adequately defined and its results can be verified.

- Summarize important findings and conclusions and develop recommendations for future efforts using the following categories: (1) a gold standard option, (2) a bare-minimum option satisfying internationally-accepted standards for impact evaluation, and (3) if/as appropriate, an option representing the best combination of good evaluation-good value for money given the present situation.
- Develop impact evaluation design frameworks for the various options/contexts identified.

The Study Terms of Reference are shown in Appendix A and the final timeline for the Study appears in Appendix C. This Report a) outlines the main approaches taken by the Team; and b) presents an assessment of the items above; and c) makes recommendations to enhance the evaluability of each PI.

The contract for this Evaluability Study was implemented by HLSP/Mott MacDonald following competitive tendering. HLSP assembled a two person team, consisting of Beth Plowman (Team Leader/health and evaluation expert) and Henry Lucas (Data Analyst), to carry out the study. The Team were supported and the contract was managed by a small team from HLSP comprising Nicolas Avril (Contract Manager), Matt Cooper (Project Officer) and Javier Martinez (Quality Assurance). The Team Leader was responsible for organizing and conducting interviews, reviewing documents and drafting study materials including the Study Report. The Data Analyst reviewed baseline survey materials and other documents, compiled tables on existing indicators, contributed to the Study Report, acquired baseline data sets and analyzed those data on basic quality parameters.

The study relied primarily on a desk review inclusive of program design documents and memos, appraisals, agreements with partners, work plans, annual progress reports, monitoring plans, website content, reports of baseline data collection, and, in one case, a Mid-Term Review Report. Standards for the conduct of impact evaluations have been drawn from the published literature and other sources and used accordingly throughout the appraisals. A complete list of documentation reviewed appears in Appendix D.

The two-person study team visited Oslo (16-20 August 2010) to gain a greater appreciation of the Partnerships Initiatives, its aims, scope, structure and challenges. The Norwegian Agency for Development Cooperation (Norad) provided the Team with documentation compiled from all five countries and facilitated access to the Norwegian state archives. Interviews, both in-person and by phone/Skype, were conducted with individuals within Norwegian Agency for Development Cooperation (Norad) as well as the Ministry of Foreign Affairs, Royal Norwegian Embassies and, in a few cases, implementing partners (Individuals interviewed appear in Appendix B). In two cases, baseline data sets were received and assessed for data quality.

The Team shared an inception report with Norad in December 2010, and based on feed back received a first draft version of this report was circulated for comments and corrections within Norad and the Royal Norwegian Embassies in the five countries. This Final report has been prepared following feed back received on the first draft from Norad and from the RNEs.

The Study Team developed two main products to underpin the assessments of evaluability. These are the PI Country Profiles and the Evaluation Options Appraisal (Appendices E through S). The PI Country Profiles serve to consolidate basic variables for each PI in a comparable format. The Profiles include information on PI duration, objectives, funding, partners, geographic focus and criteria for selection. The Appraisals present the Team's consideration of a basic set of parameters for impact evaluation in each PI country. In addition, for each PI, an indicator table is presented. These tables represent a compilation of indicators drawn from available program documentation – they do not represent the Team's recommendations on appropriate indicators for evaluating PI performance or impact. In all cases, the Team believes that these indicator tables could be far more focused and streamlined.

1.3 Assumptions and limitations

Partnerships are often complex, dynamic endeavours where local insights and knowledge are crucial to fully understand evaluability. In this sense the Team was limited by its reliance on a desk review process where travel to the five countries did not take place. As a result, the study may have been limited at times by the extent to which the key questions are answerable through available materials or key informant interviews.

As PI oversight is largely located within the norwegian Embassies (although this varies by country), the team depended on the norwegian embassies placing a priority on the PI Evaluability Study and responding to requests for information in a timely manner. Overall, requests for interviews were met promptly and norwegian embassies informants and implementers were candid in their responses and generous with their time. The acquisition of baseline datasets posed slightly more difficulty with some delays encountered and partially fulfilled requests for both meta- and micro- data.

1.4 Acknowledgements

The Team and HLSP would like to thank all those who assisted with the Evaluability Study. Special thanks go to Siv J. Lillestøl, Senior Adviser in the Department for Evaluation, Norad for facilitating the Team's work with each of the Royal Norwegian Embassies and for her patience dealing with delays in Study completion. The Team would also like to thank Cliff Wang, Senior Adviser in the Global Health and AIDS Department, Norad, for his support and insights throughout the Study. His support in amassing and sharing the documentation from the five Partnership Initiatives made the Team's work far more manageable. The Team also appreciates the commitment of staff within Norad, the Ministry of Foreign Affairs, Royal Norwegian Embassies and implementing partners who generously gave their time to this Study.

2. Assessment of Evaluability

2.1 Preliminary observations

Based on the first activity phase, the Team made a number of observations that helped to form the basis of the study's approach. Among those observations:

1. There is **considerable diversity** between the five PI country programs in terms of implementation status and arrangements, partners, timelines and technical emphases. Table 1 below summarizes the five PI included in the study. For example, the PIs in Tanzania and Malawi essentially serve as pilot projects for performance-based funding approaches and incorporate research components to demonstrate the effect of those interventions. In contrast, the PIs in India, Pakistan and Nigeria operate primarily as support to the large-scale implementation of maternal, newborn and child health programs, with limited 'innovation' components. Any proposed impact evaluation design will need to be tailored to each PI and its context. As a consequence, comparability of outcomes between PI countries is limited.

Table 1: Overview of the five Partnership Initiatives

PI	Dates	Level of funding	Nature of partnership	Geographic scope
India	September 2006- 2013 ⁱ	NOK 500 million	Agreements with UNOPS ⁱⁱ , UNICEF, WHO	Four states with three districts in each through UNOPS ^v , fifth state supported through UNICEF
Pakistan	2008/09 ⁱⁱⁱ - 2013	NOK 250 million	Agreement with UNDP for the UN's Pakistan One Fund to administer; three UN agencies (UNICEF, UNFPA and WHO) implement	Ten districts in Sindh province representing 29 % of the provincial population
Nigeria	June 2008 - 2010	NOK 250 million	Agreement with DFID to expand an existing project, implemented through a consortium	Four states in Northern Nigeria

PI	Dates	Level of funding	Nature of partnership	Geographic scope
Tanzania	2007-2011	NOK 225 million	MOU between RNE and the GoT to participate in the pooled funding mechanism (“basket financing”) and to support separate activities ^{iv}	Basket fund disburses to all areas; performance-based pilot in several districts in Coastal province.
Malawi	2011-2013	NOK 30 million	Agreement with KfW, as lead donor, is being finalized; Ministry of Finance also a signatory. RNE may sign separate agreement with GoM.	Performance-based funding pilot in the Central Eastern Zone.

ⁱ Agreement, originally slated to end in 2010, extended to 2013 and may be further extended to 2015.

ⁱⁱ Agreement signed January 2009

ⁱⁱⁱ UNOPS, in turn, sub-contracts with State Health Societies in Bihar, Madhya Pradesh, Orissa and Rajasthan.

^{iv} Eighty percent of funding through basket funding. Twenty percent for feasibility study and piloting of performance-based approaches, strengthening HMIS, support for NGOs.

^v UNICEF and WHO work in additional areas consistent with their national programs of cooperation.

In general, it can be noted that later-starting PIs (Nigeria, Malawi and Tanzania) are designed with a stronger, clearer technical basis and rationale compared to some of the early PIs (notably India). As part of its work, the Team identified lessons learned from the PI experience in regards to monitoring and evaluation to improve and strengthen the PI programs still in the formative stage.

2. The PI programs have a **dual nature**. Although fundamentally development cooperation programs, they have their origins in a political initiative. During the initial week of discussions, many potential evaluation questions, reflecting both aspects of the programs, were proposed. Some of the questions and issues cannot be effectively addressed in the Evaluability Study. To illustrate, interviewees expressed interest in knowing: (a) how to value the political commitment of having highest level politicians travelling to PI countries and visibly advocating for MDGs 4 and 5; and (b) the comparative effectiveness of using the PI to channel development monies to countries versus other channels (such as multilaterals). Answering these questions in a meaningful manner requires additional forms of information collection and analyses and is, essentially, beyond the scope of the study.
3. The Partnership Initiative documents make frequent reference to a range of concepts such as innovation and risk-taking. Unfortunately, as far as we can determine, **these concepts have not been described in any overarching strategy or framework document for the PI**. While it may be reasonable to use such terms as “innovation” and “risk-taking” in describing an overall approach to be adopted in an intervention, their lack of definition poses a problem for evaluation which relies on clearly articulated statements of purpose,

objectives and the means of achieving those objectives. While such statements are unavailable in an overarching document, the team sought to construct an understanding of innovations for each country-specific PI.

2.2 Team's approach and assessment

The means of evaluating progress towards the MDGs is receiving close scrutiny and will continue to draw attention as the 2015 achievement date for targets draws near. It is therefore apropos that Norad would seek an Evaluability Study for the Norwegian-funded Partnerships Initiatives (PI), a program aimed at supporting implementation of MDGs 4 and 5 in five countries.

For some evaluation experts, recent experience with large-scale program evaluations points to the need for a change of focus and more innovative approaches^{3 4 5}. In light of multiple new international health assistance actors and resources, they regard

.... recent experience in evaluating large-scale child health initiatives suggests that traditional design (intervention vs. comparison areas) seldom allows valid attribution in the current development context.

traditional evaluation designs which rely on baseline and impact studies in both intervention and comparison areas as no longer appropriate. These designs are seen as increasingly limited in their ability to isolate and evaluate specific program effects given that parallel initiatives are often being undertaken simultaneously. For example, when development partners support similar interventions in different geographic areas, evaluations which aim to compare intervention and comparison areas may be seriously compromised. A depiction of the overall environment in which PIs operate and are evaluated appears in Figure 1 below. The intended outcomes (coverage and impact) are influenced by a range of factors which cannot be fully controlled nor accounted for.

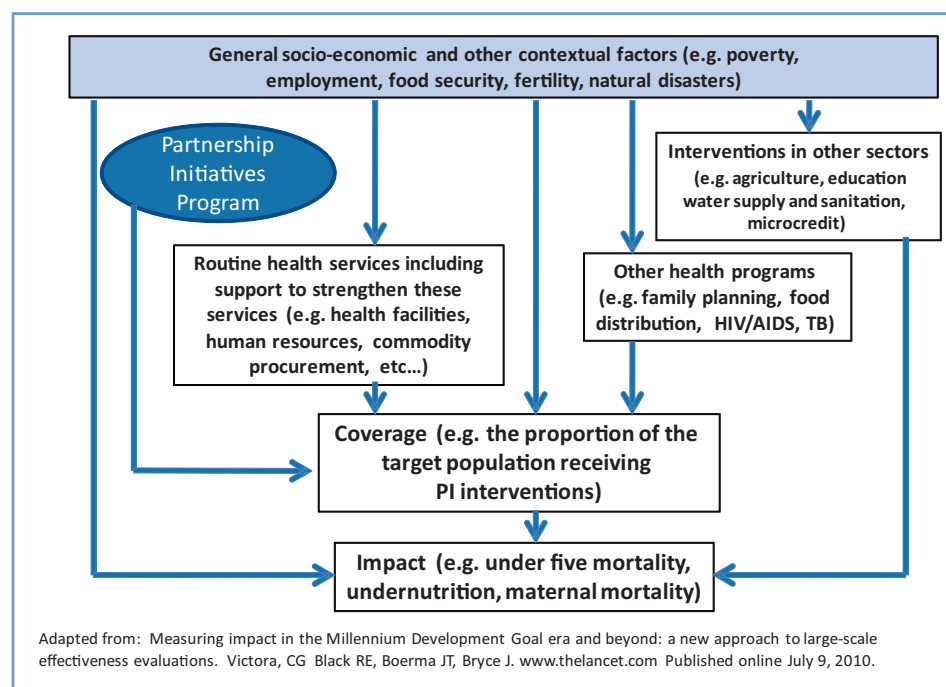
In order to build on these recent and relevant experiences, the Team has emphasized a set of basic principles and approaches throughout the attached Options for Evaluation. These principles and corresponding assessment of their applicability in the evaluation of the PIs appear below.

3 Measuring impact in the Millennium Development Goal era and beyond: a new approach to large-scale effectiveness evaluations. Cesar G Victora, Robert E Black, J Ties Boerma, Jennifer Bryce. www.thelancet.com Published online July 9, 2010.

4 The Accelerated Child Survival and Development programme in west Africa: a retrospective evaluation Jennifer Bryce, Kate Gilroy, Gareth Jones, Elizabeth Hazel, Robert E Black, Cesar G Victora. *Lancet* 2010; 375: 572–82. Published Online. January 12, 2010.

5 Evaluating child survival programmes. Cesar G Victora, Robert E Black, Jennifer Bryce. *Bull World Health Organ* 2009; 87:83.

Figure 1: Schematic of factors operating in parallel effecting maternal and child health



2.2.1 Evaluating the impact of the Partnership Initiatives “programme” versus specific interventions.

The observations above on the increasing difficulty in utilizing traditional evaluation designs (i.e. utilizing baseline and impact studies in both intervention and comparison areas) have implications for the Partnership Initiatives. For a number of reasons, it will be difficult to evaluate the impact of the Partnership Initiative as a comprehensive programme in some countries. Typically, a programme impact evaluation would encompass a set of interventions working in tandem to attain a specific development objective⁶. In at least two cases (i.e. Tanzania and Malawi), the Partnership Initiative comprises both funding to a basket or pooled funding mechanism and funds for results-based financing pilots. Clearly, this distribution of resources, while responsive to country setting, cannot be subject to a single, comprehensive programme impact evaluation. In India and Pakistan, elements of the Partnership Initiative are implemented by differing agencies sometimes in different geographic areas. Here again, it would be difficult to design and conduct a single evaluation of programme impact. However within each PI, there are clearly interventions and innovative elements that should be subject to rigorous evaluation to demonstrate effects on maternal and child service delivery and outcomes. These evaluative activities are required to generate an evidence base for replication. In sum, the evaluability of specific PI interventions is attainable while the evaluability of an individual PI “programme” as a whole is less feasible. Indeed, the results-based funding pilots of the Tanzania PI and Malawi PI are being designed in this manner.

⁶ Glossary of key terms in evaluation and results based management. OECD. Development Assistance Committee. Undated.

2.2.2 Ensure timely and consistent Partnership Initiatives documentation

Systematic documentation describing the key elements and intensity of a PI over time, as well as features of the overall context that may affect program objectives, is essential to good evaluation⁷. This documentation should also include the PI's logical model/framework/causal chain (i.e. inputs, activities, outputs, outcomes) depicting how the PI is expected to achieve its purpose. Beginning this documentation process early in the life of the PI is important, even where there is a high probability that substantial design revisions will be required over time. The internal validity of the evaluation design can be greatly strengthened by routinely updating decisions and knowledge on key issues including sample selection and minimum effect size, potential "spillover" effects (i.e. PI activities reaching other areas), contagion (other development partners engaged in similar activities in either PI or comparison areas) and other relevant changes in program context. Structured documentation can also provide important contextual information such as significant delays due to natural occurrences/disasters (e.g. flooding in Sindh province in Pakistan in 2010). Table 2 includes a selected set of variables which represent the documentation efforts planned or underway within each PI.

Table 2: Summary of current status of documentation efforts within Partnership Initiatives

	Logical model or framework	Systematic documentation conducted	Criteria for site selection	Presence/ role of other dev. partners (i.e. potential contagion) described	Sampling parameters (e.g. minimum effect size)
India	Example is appended to Strategy document but not finalized	Materials prepared for PMG and JSC and mtg. minutes provide high-level "snapshot", No detailed implementation-level documentation was provided.	States – reference to high mortality; among NRHM "high focus" states Districts – no criteria found	Significant involvement of other partners in same states is not accounted for.	Description in Baseline Survey Report provides sufficient basis for further examination.
Pakistan	Elements exist in Program Document but not coalesced	Not being conducted	Reference to high mortality in the selected districts.	Important initiatives working in the same area with similar purpose are not mentioned.	Description in Baseline Survey Report provides sufficient basis for further examination.

⁷ Standard guidance is available to facilitate the documentation process. See: *Guidelines for documenting program implementation and contextual factors in independent evaluations of the Catalytic Initiative. Working Paper, v. 1.1.* Document prepared by Kate Gilroy, Elizabeth Hazel, Jennifer Callaghan, Jennifer Bryce and the IIP-JHU Catalytic Initiative evaluation group. Institute for International Programs Johns Hopkins Bloomberg School of Public Health.

	Logical model or framework	Systematic documentation conducted	Criteria for site selection	Presence/ role of other dev. partners (i.e. potential contagion) described	Sampling parameters (e.g. minimum effect size)
Nigeria	A combined log frame was created for DFID and Norwegian joint programme	Annual Reviews provide good document base for implementation experience.	Extensive description of other actors and efforts over time.	Project Concept Note and Desk Appraisal provide insight on the other donors and planning, positioning of projects with similar aims.	Yes
Tanzania	No	NA (too early)	Under process	Yes	NA (too early)
Malawi	Included in Decision Document	NA (too early)	Learning site selected based on proximity to capitol and infrastructure	Yes	NA (too early)

2.2.3 Systematic Implementation-level Monitoring

The Team emphasizes this function as the lynchpin to maintaining an evidence base able to convincingly link processes with outputs and outcomes. This form of monitoring would provide regular, valid measures of specific PI inputs and processes including training, supervision, and delivery channels⁸. By tracking program implementation in a stepwise manner, this design element can signal when follow-up (end line) assessments are warranted based on implementation status and actual duration as opposed to more arbitrary five-year project timeframes. In addition, implementation-level monitoring can also provide needed information on the timing, intensity and quality of programme implementation. The Team found that no PI, with the possible exception of Nigeria, is currently tracking implementation in a manner that would contribute to impact evaluation (albeit in Tanzania and Malawi, significant elements of the PI are just getting started).

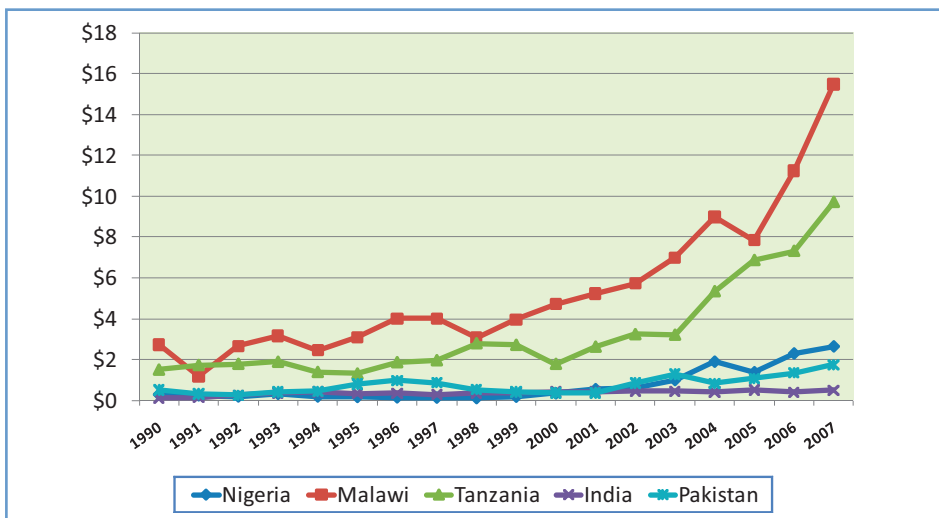
2.2.4 Use of Comparison Groups

A challenge in evaluating the impact of the PI in a given country is to determine what would have happened to the target population (i.e. women and children) if the PI program had not existed -- otherwise known as the counterfactual. While the counterfactual can be examined at several levels (e.g. resource availability, service delivery) the Team will focus here on the use of comparison groups.

⁸ The Team believes that program effort or performance scores may be an appropriate means for PI programs to create quantitative measures of implementation strength based on information generated through this documentation process. Such composite measures can reflect the intensity of program implementation (i.e. the "dose") in dose-response analyses.

Some PI countries have selected neighbouring districts as comparison groups in their baseline data collection. Challenges to such intervention-comparison designs and their increasingly limited ability to isolate and evaluate specific program effects were described above. As indicated, one major challenge is the selection of an “untouched” comparison group when the overall level of development assistance for health has risen sharply over the last decade. In countries such as Malawi and Tanzania, even if the PI intervention isn’t implemented in the so-called control districts, there is likely some significant level of investment by another development partner (See Figure 2 below).

Figure 2: Per capita development assistance for health, Partnership Initiative countries, 1990 to 2007,



Source: Institute for Health Metrics and Evaluation.

The Team suggests several design and analytical approaches to enhance the use of comparison groups in PI evaluation efforts. At a minimum, PIs using comparison groups in their evaluations must account for the presence of other actors working in comparison areas, including the interventions they support and the levels of additional investment. An evaluation design option is the use of a “pipeline approach” in which the staggered introduction of PI activities effectively creates a comparison group among those who are targeted to be reached by/participate in PI activities but have not yet done so⁹. By rolling out PI activities in this fashion, comparisons between groups and/or districts can be made to assess the associations between levels of program implementation and outcome variables.

Another, and possibly complementary, evaluation strategy in environments with multiple development partners and maternal and child health initiatives underway is through the use of a “difference-in-difference” approach to analysis. Also known as double difference, this involves comparing *changes* over time observed in the PI target population with those seen in the comparison group. The strength in this approach is its’ recognition that in real-life settings, other factors and/or programs will often create changes in one or both groups (i.e. comparison groups do not remain “untouched”)¹⁰.

⁹ Khandker SR. Handbook on impact evaluation: quantitative methods and practices. 2010. The World Bank.

¹⁰ This analytical approach was used in the evaluation of the UNICEF-funded Accelerated Child Survival and Development Program in Ghana. Source: Final Report. Retrospective Evaluation of ACSD: Ghana. Submitted to UNICEF on 7 October 2008. Institute for International Programs Johns Hopkins Bloomberg School of Public Health Baltimore, MD.

2.2.5 Range of evaluation options

The Team considered a range of options for the evaluation of the Partnership Initiatives in each of the five countries (see Option Appraisals in Appendices E – I). This section provides a brief introduction to the three categories included in the appraisals¹¹. A summary of these approaches appears in Table 5.

Table 5: Range of options considered for Partnership Initiative evaluability

	“Gold standard”	Global health - Normative	Minimal
Type of inference	Probability	Plausibility	Adequacy
Questions to be answered	<ul style="list-style-type: none"> • Are effects measured due to the implemented program? • Can the difference between programme and control areas be attributed to the programme with a small, known probability of confounding effects? 	<ul style="list-style-type: none"> • Are measured effects likely due to the program rather than other influences? • Can confounding factors be ruled out as contributors to the observed change? 	<ul style="list-style-type: none"> • How did the programme perform compared to previously-established targets or criteria?
Design considerations	<ul style="list-style-type: none"> • Controlled trial usually involving random assignment of individuals or clusters to either a treatment or control group • Typically designed before the intervention takes place (i.e. ex ante evaluation) to allow for randomization and control of spill-over effects 	<ul style="list-style-type: none"> • Uses clusters or groups with and without the program to attempt to account for non-program factors; typically measured before and after program implementation • Without randomization, control groups are created through other means such as propensity score matchingⁱ. • Analytical methods including dose-responseⁱⁱ, pipelineⁱⁱⁱ and difference within differences^{iv} analyses also allow for comparison of acceptable rigor. 	<ul style="list-style-type: none"> • Assessment before and after program implementation without comparison group or area; can determine whether established targets were met or not

- i Propensity score matching is a method in which comparison areas are selected and matched to program areas based on similarities demonstrated through statistical methods.
- ii Dose response analyses create categories for analyses based on program participation (e.g. intensity and duration).
- iii Pipeline approach is an evaluation design in which the comparison group is scheduled to participate in the program but have not yet done so.
- iv Difference within differences analyses is useful when programs operating in the comparison areas result in changes in key indicators. Comparisons are then made based on the difference in the outcome indicators between program and comparison areas.

11 This section draws from Habicht JP, Victora CG and JP Vaughan. Linking evaluation needs to design choices. 1997. UNICEF Staff Working Paper. Evaluation and Research Series. EVL-97-003.

The “*gold standard*” evaluation design entails random assignment of either individual or group to either a treatment or control group. These evaluations are designed in advance of the intervention in order to facilitate the random assignment procedure. Efforts are made to limit or control the influence of external factors – including implementation of similar interventions occurring in the control area. The difficulties in adhering to these conditions were discussed above. Even successful applications of the design encounter some difficulty in maintaining the divide between treatment and control groups¹².

The second option considered throughout the Report is loosely term the *global health –normative approach*. This approach can encompass a number of differing designs but is considered a well-accepted standard for programme evaluation. At a minimum, this approach requires a justifiable effort to account for external factors through a variety of means. A common variant of the approach entails “before and after” measurements in a programme intervention area and a comparison area – where presumably the programme intervention will not be implemented. The overall rigor of the approach can be enhanced through careful selection of the comparison group (e.g. propensity score matching) or even use of an internal comparison group (e.g. pipeline analyses). This normative approach can also be strengthened through careful programme documentation and use of programme strength measures (e.g. dose-response) to account for differing degrees of implementation.

A third option is termed a *minimal option*. The focus in this approach is on the achievement of pre-established performance targets or criteria. These targets are generated used for either programme outcomes (e.g. a 20 % increase in the percent of infants exclusively breastfed for six months) or outputs (e.g. number of ITNs distributed at ANC clinics increased from x to y). This approach does not attempt to account for external influences.

12 e.g. a) Rivera JA, Sotres-Alvarez D, Habicht JP, Shamah T and S Villalpando. 2004. Impact of the Mexican program for education, health and nutrition (Progresa) on rates of growth and anemia in infants and young children. JAMA. Vol. 291 No. 21. b) Basinga P, Gertler PJ, Binagwaho A, Soucat A, Sturdy JR, C.M.J. Vermeersch. January 2010. Paying Primary Health Care Centers for Performance in Rwanda. . The World Bank. Policy Research Working Paper 5190.

3. Main Recommendations

3.1 General Recommendations

Within the recommendations, certain themes emerge across the PI's. These overarching themes are grouped here and are applicable, to some extent, to all PIs.

Clarify logical model and revisit assumptions – The Partnership Initiatives operate largely without clearly articulated logical models. In some cases, only general explanations are provided on how PI resources and partners are expected to work together through defined processes to achieve a set of intended outputs and outcomes. This degree of generality offers little guidance to those designing monitoring and evaluation procedures. Key assumptions that may seriously influence the pace of implementation and/or achievement of results are often left unstated. This lack of a considered, detailed program framework greatly complicates any attempt at evaluation (and accountability).

“The logic model describes how a program should work, presenting the causal chain from inputs, through activities and outputs, to outcomes. While logic models present a theory about the expected program outcome, they do not demonstrate whether the program caused the observed outcome.”

3ie (2009)

Several Partnership Initiatives have logical models which require only additional notation and updating (e.g. Malawi). Other PIs lack an existing logical model (e.g. India and Pakistan) and it would be worthwhile for these PIs to develop one – even at this point in implementation. In terms of timing, it would be ideal for new PIs to develop and/or update their logic model after the completion of a baseline survey, analyses of other available data and establishment of programme targets. For PIs which are further in implementation, the development or review/update of a logic model could be considered as part of mid-term review efforts. Finally, in keeping with the Paris Declaration for Aid Effectiveness, the PIs' logic models should be aligned with existing performance frameworks for the programmes that they support¹³.

Clarify attribution issues – The PIs include longer-term impact measures, including under-five mortality, maternal mortality and neonatal mortality, among their indicators for monitoring and evaluation. In general, longer-term impact measures (i.e. under-five mortality rate, infant mortality rate and maternal mortality ratio) are not recommended for PI monitoring and evaluation. The reasons are two-fold and relate to measurement issues and attribution. In the absence of vital event registration

¹³ The Paris Declaration on Aid Effectiveness (2005) Accra Agenda for Action. OECD.

systems, reliable mortality measurements are derived from household survey data. Infant and under-five mortality estimates require relatively large sample sizes and are therefore typically generated for five year intervals. In addition, mortality rates are not commonly available for administrative units below the first sub-national level. Maternal mortality ratios require even longer time frames (10 years) and are not appropriate for Partnership Initiative M&E. Secondly, changes in mortality are brought about through collective efforts, often across sectors, which rarely allow for the contribution of any single donor or actor to be reliably estimated. To the extent that mortality measures are included in development partners' M&E plans, they should be consistent with overall national targets and timelines and acknowledged as achievable through joint efforts.

In addition, the PI documents often do not begin to account for the complex donor landscape in these five countries or attempt to locate the role of the PI within this larger context. While there are exceptions, it would benefit the Partnership Initiative as a whole if these issues were clearly and consistently incorporated into their guiding documents.

Revisit and refine the monitoring and evaluation plans – In several cases, it appears that PI M&E plans were outlined as an initial element of the program document but then never further elaborated or made operational. While baseline data collection/ compilation has been attempted for all PIs, implementation-level monitoring has, with the exception of Nigeria, been neglected. In some cases, there is no indication as to which partner is to report on what indicators, when or how. In addition, almost every PI has an over-abundance of indicators – many of which are not measurable, not attributable or not adequately defined (see the appendix tables labelled M&E Plans). PIs which are mid-stream (e.g. NPPI) should take the opportunity to ground-truth and update their initial M&E plans, as the NIPI is doing. Newly developed PIs, such as Malawi, should carefully revise their M&E plans with an eye towards streamlining and reducing the monitoring burden.

Adhere to internationally-accepted standards for monitoring MDGs 4 and 5 - Partnership Initiative designers should adhere to donor commitments to reduce the use of performance indicators that are not consistent with partners' national development strategies and to harmonize monitoring requirements with emphases on the use of partner countries' statistical, monitoring and evaluation systems¹⁴. At country and global levels, MDGs 4 and 5 are tracked using a consistent set of internationally-agreed upon indicators. Some but not all of these measures are appropriate for use in Partnership Initiative monitoring and evaluation (i.e. mortality measures are not recommended for use in PI M&E for the reasons described above). Table 3 below presents each of the internationally-accepted MDG 4 and 5 indicators with a corresponding recommendation for its use in the PI monitoring and evaluation.

Three of the internationally-agreed MDG 4 and 5 indicators are recommended for use across Partnership Initiatives: the proportion of infant immunized against measles, the proportion of births attended by skilled health personnel, and ANC

¹⁴ Ibid.

coverage. Program priorities of all Partnership Initiatives are fairly well-aligned to these measures. Indeed, as seen in Table 4, for skilled attendance at both and ANC coverage each PI has already included a variant of the indicators within their monitoring and evaluation plans. Two of five PIs have included measles immunization coverage. On a positive note, Table 4 demonstrates that there are categories of programmatic priority which are consistent across Partnership Initiatives in the five countries. However, the wide variety of indicator selection raises some questions – which, while noted here, are beyond the scope of this Study to examine. Firstly, it would be important to examine whether these varied measures are consistent with the construct of indicators being used by the national programmes supported in each PI country. If not, what is the rationale for departure? Secondly, would it behoove the Norwegian support for MDGs 4 and 5 to consistently use a set of internationally-agreed indicators across all supported countries?

Any effort to create a single set of performance measures for the PIs is advised to a) focus on a small number of indicators such as those highlighted here; b) ensure that PI monitoring and evaluation is consistent with national MDG monitoring efforts; and c) ensure that these outcome measures are firmly linked to a set of process and output measures which reflect the unique contribution of the PI to their achievement.

Table 3: MDG 4 and 5 indicators and recommendations for the Partnership Initiatives

MDG goal and target	Internationally-agreed indicators	Recommendation for Partnership Initiatives
Goal 4: Reduce child mortality		
Target 4.A: Reduce by two-thirds, between 1990 and 2015, the under-five mortality rate	4.1 Under-five mortality rate	• Not recommended for PI M&E ⁱ
	4.2 Infant mortality rate	• Not recommended for PI M&E ⁱ
	4.3 Proportion of 1 year-old children immunised against measles	• Recommended for PI M&E ⁱⁱ
Goal 5: Improve maternal health		
Target 5.A: Reduce by three quarters, between 1990 and 2015, the maternal mortality ratio	5.1 Maternal mortality ratio	• Not recommended for PI M&E ⁱ
	5.2 Proportion of births attended by skilled health personnel	• Recommended for PI M&E
Target 5.B: Achieve, by 2015, universal access to reproductive health	5.3 Contraceptive prevalence rate	• Not applicable
	5.4 Adolescent birth rate	• Not applicable
	5.5 Antenatal care coverage (at least one visit and at least four visits)	• Recommended for PI M&E
	5.6 Unmet need for family planning	• Not applicable

ⁱ Rationale for exclusion of mortality measures - measuring change reliably requires longer-time frame and large sample size; changes are achievable only through collective efforts which rarely allow for the contribution of any single donor or actor to be reliably estimated

ⁱⁱ There is justification to use other antigens, notably DPT3, as a performance measure. DPT3 is regarded as more indicative of immunization systems strengthening whereas measles coverage can fluctuate based on campaigns.

Table 4: Current use of three recommended MDG indicators across Partnership Initiatives

Recommended MDG indicator	Malawi	Tanzania	Nigeria	Pakistan	India
Proportion of 1 year-old children immunised against measles	<ul style="list-style-type: none"> • % Fully immunised infants • # Infants given vaccine X 	<ul style="list-style-type: none"> • Immunization - DTPHb 3 (performance measure for facilities participating in P4P pilot) 	<ul style="list-style-type: none"> • % of infants fully immunised by first birthday • Number of 1-year-old children immunised against measles 	<ul style="list-style-type: none"> • % children 12-23 months immunised against six immunisable childhood diseases • % children less than 1 immunised against measles 	<ul style="list-style-type: none"> • % Children (12-23 months) fully immunized (BCG, 3 doses of DPT, Polio and Measles) • DPT1 to DPT3 drop-out rates
Proportion of births attended by skilled health personnel	<ul style="list-style-type: none"> • # Facility based deliveries • % Women that benefit from transport+** during last delivery • # Women who use shelters before delivery 	<ul style="list-style-type: none"> • Proportion of births assisted by a skilled attendant. • Proportion of deliveries taking place in a health facility. 	<ul style="list-style-type: none"> • % of births attended by a skilled birth attendant (SBA) in targeted CEOC clusters 	<ul style="list-style-type: none"> • % births attended by a skilled health attendant (SBA) • % births attended by a skilled health attendant at a health facility (simply at a facility in HH survey) 	<ul style="list-style-type: none"> • Deliveries in the NIPI states conducted by SBAs • % Home deliveries assisted by a Doctor/ Nurse/LHV/ANM • % Institutional births
Antenatal care coverage (at least one visit and at least four visits)	<ul style="list-style-type: none"> • # 1st Antenatal visit • # Women with 2 Antenatal visits 	<ul style="list-style-type: none"> • Proportion of women booking early for ANC 	<ul style="list-style-type: none"> • % of women receiving ANC 	<ul style="list-style-type: none"> • % pregnant women with at least 1 ANC visit to a skilled birth attendant 	<ul style="list-style-type: none"> • % Mothers who had at least 3 Ante- Natal Care visits during last pregnancy

Finally, although not a MDG indicator, there is some convergence around postnatal care/postpartum care indicators within the PIs. Four of the five PIs have some form of postnatal/postpartum care indicator thus creating an opportunity to advance use of these (relatively new) measures. The postnatal/postpartum measures are as follows:

- Malawi - # Infants with postnatal check up
- Tanzania - Postnatal care attendance rate
- Pakistan - % new mothers receiving postpartum care
- India - % Children had check up within 24 hours after delivery (based on last live birth)
% Mothers who received post natal care within 48 hours of delivery (last child)
Number of newborns and mothers visited up to 6 times during 1st 6 weeks after delivery)

3.2 “Best Case” Options for each Partnership Initiative

A range of recommendations based on evaluability appraisal appear in the appendix. However, for each PI, the Team has identified a “best case” option which also appears below.

Norway-Pakistan Partnership Initiative (NPPI): The *global health – normative* option provides the best case for the NPPI in that it could meet internationally-accepted standards and provide evidence for maternal, newborn, and child health (MNCH) policy dialogue and program development. The Team proposes that efforts focus particularly on a) contracting out of maternal, newborn, and child health services to the private sector; and b) generating demand through the use of vouchers/incentives. This option would entail evaluation in the 10 project districts without the use of control districts.

In those 10 project districts, evaluators should fully account for the inputs of other development partners -- but not try to limit or control those inputs. Baseline sample parameters were designed around the estimated prevalence of skilled attendance at birth. The end line survey sample should be designed to either a) determine with a known confidence whether pre-determined targets had been achieved, or b) measure statistically significant change over time (pre/post) in key variables (the preferred option). This design would analyze the “dose-response” of NPPI-supported MNCH providers/services (i.e. measures of program duration and intensity of implementation) for either individual women or villages and associate that exposure with the desired outcomes. Moreover, this design would maximum NPPI’s sequential introduction across districts (as activities are not initiated in all ten districts simultaneously) in a “pipeline” analysis. This method would allow for an internal comparison group within the ten districts. Finally, this approach could also include consistent and systematic monitoring of implementation across agencies in the project.

Norway-India Partnership Initiative (NIPI): NIPI is best positioned to conduct a “minimal” option evaluation that would seek to triangulate available data sources (Demographic and Life Health Survey (DLHS)-3, NIPI baseline, UNFPA and UNICEF coverage surveys, National Rural Health Mission (NRHM) routine data, health

information system data as well as the upcoming DLHS-4) to a) determine with known certainty whether NIPI targets were achieved (y/n) and b) whether statistically significant change occurred in specified outcome measures over time. This option could include the planned NIPI mid-line and end line survey data, in a greatly reduced form focused on a small number of key indicators. Triangulation of these available data sources at multiple levels could help to identify associations using timelines (of implementation) and trends in desired outputs and outcomes¹⁵. This approach would focus more on whether change has been achieved and less on identifying and substantiating the NIPI contribution. Given the wealth of existing information (both survey and routine), the Team believes that this option is entirely feasible for NIPI and could provide a well-rounded picture of the NIPI experience.

Further, it is recommended that specific, targeted interventions are evaluated for their effect on desired outcomes (e.g. to what extent do ASHA activities increase the utilization of home-based newborn care practices? to what extent does the presence of Yashotas improve specific aspects of facility-based delivery services?). This is best conducted through a set of relatively small, well-designed studies and could help guide National Rural Health Mission (NRHM) investments and direction. These investments would provide relevant information about specific interventions supported through NIPI and reflect its contribution.

Norway-Malawi Partnership Initiative (NMPI): The NMPI is well-positioned to conduct a **global health – normative** option with important design elements in place at its inception. According to the PI documents, a monitoring and evaluation framework will be defined during initial phase of the results-based financing (RBF) program design (starting January 2011). The Team would encourage the designers to take careful note of existing international experiences and approaches for evaluation of results-based financing – notably the required timeframe and required resources. As part of the initial monitoring and evaluation design, a strategy and plan for documentation should be developed and contracted. Documentation becomes particularly important as the study areas were selected based on their accessibility to the capitol and relatively better infrastructure. These circumstances could be seen as compromising the programmes’ replicability. As a learning site with better than average conditions, the PI partners should assiduously document the conditions required for successful performance (e.g. service accessibility and quality) with an eye towards how those conditions will be made available in other areas and program effect replicated.

The available documents cite the main goal of the monitoring and evaluation component as “improving the availability, quality and use of the data needed to inform results-based financing program reviews and planning processes, monitor health outcomes progress, health system performance and ultimately demonstrate impact of the project on selected outcome indicators with reasonable degree of certainty”. This goal is far too ambitious for a pilot project expected to be evaluated

15 For an example of this approach see: System-level determinants of immunization coverage disparities among health districts in Burkina Faso: a multiple case study Slim Haddad, Abel Bicaba, Marta Feletto, Elie Taminy, Moussa Kabore, Boubacar Ouédraogo, Gisèle Contreras, Renée Larocque, Pierre Fournier. BMC International Health and Human Rights 2009, 9(Suppl 1). Published: 14 October 2009.

after 2 to 3 years -- likewise the expectation that a baseline, mid-term and end of project evaluation will be carried out.

Program designers are encouraged to moderate their expectation for the M&E component to a more achievable set of aims. While the concept note drafted for the evaluation is a good start (even the basic depiction below of indicator domains and data collection is notably absent in other PIs reviewed), its next version should reflect the short timeframe and add greater specificity on what existing information will be utilized in the evaluation (potential data sources mentioned project and health facility documents, clinical reporting systems and Health Management Information System (HMIS)). The Team encourages the PI to rely on existing sources such as the HMIS and the instrument jointly developed by the Ministry of Health and JHPIEGO¹⁶ which will be utilized throughout the country (independent of the PI pilot) and will thus serve as an important measurement of maternal and newborn health services.

Finally, the program designers are strongly encouraged to work closely with the RBF for Health impact evaluation network (part of a Norwegian funded initiative based at the World Bank to support results-based financing innovation in eight countries). There is a readily available set of tools and materials to assist with the further development and refinement of the evaluation approach.

Norway-Tanzania Partnership Initiative (NTPI): The majority of the NTPI investment is directed through the pooled funding mechanisms and assessed according to jointly agreed procedures. However, as with the PI in Malawi, the NTPI is well-positioned to conduct a **global health – normative** option around the pay-for-performance (P4P) component with important design elements in place at its inception. The Team recommends that the program be realistic and focused in the evaluation design and requirements for data collection – as pilot’s timeline has been compressed to just 18 months, the evaluation design will have to accommodate this truncated implementation schedule. The idea of using a “facilities readiness” measure is a very good one but will need sufficient time to be developed, tested and garner buy-in from national decision-makers (as there are clear implications for the national P4P programme). The NTPI is encouraged to work closely with the RBF for Health impact evaluation network (part of a Norwegian funded initiative based at the World Bank to support result-based financing innovation in eight countries) and to draw on the readily available tools and materials to assist in the design of the evaluation. Finally, as the selected districts are benefitting from a HMIS improvement scheme, the NTPI will need to be able to describe minimum requirements for HMIS operation during any proposed scale-up of the pay-for-performance pilot – recognizing that not all districts will have benefitted from the Health Management Information System activity.

Norway-Nigeria Partnership Initiative: The Partnership Initiative in Nigeria is already following a **global health – normative** approach to evaluation. The aim of that evaluation is to measure change over time in specified outcome indicators in

16 an affiliate of Johns Hopkins University

both cluster and non-cluster areas within each state. The Team's concern, is that 'success' at this level, given the dysfunctional nature of services prior to the intervention, will need to be carefully interpreted in terms of scaling up to state level. We would strongly recommend that this is a case where the use of implementation data to link inputs to intermediate outputs/outcomes along the project timeline to analyse the implementation process could be of greater value than a simple assessment of project success.

Appendixes



Appendix A: Terms of Reference

Evaluability Study of Norwegian support to the achievement of the Millennium Development Goals 4 and 5: to reduce child mortality and improve maternal health.

Purpose

The purpose of the study is to assess the evaluability of Partnership Initiatives (PIs) in India, Tanzania, Nigeria, Pakistan and Malawi. Partnership Initiatives refer to Norwegian support to the achievement of the Millennium Development Goals 4 and 5: to reduce child mortality and improve maternal health. This is part of Norway's global health portfolio.

The evaluability study shall result in recommendations and proposed action plans for impact evaluations to be conducted in the five PIs at a later stage. Besides being a stand-alone exercise related to the five PIs, this study can also be seen as necessary input to a larger evaluation of Norway's total programme of support to MDG 4-5 attainment, scheduled to take place in 2011.

Scope

For each of the 5 PIs, the Team shall:

1. Map the basis for Norway's support to the PI using project documents (examining project design, results chain, indicators and success criteria, implementation arrangements, methods, approaches, types of activities, monitoring and evaluation systems, stakeholders, intended target groups and geographic coverage, and so on). Then assess the adequacy and quality of such documentation in terms of potential "impact evaluability" of the PI towards the end of the initial implementation period (typically 5-6 years).
2. Identify contextual factors likely to influence intended and unintended PI impacts. Then assess the availability, adequacy, and quality of contextual factor data/information currently available, going back at least 5 years in time where possible.
3. Review currently-available baseline data/studies and their appropriateness in terms of impact evaluation reliability and validity, keeping such questions as the following in mind:
 - What are the "right" impacts needing to be measured/verified?
 - Are these "right" impacts actually verifiable to acceptable standards given the existing baseline information available?
 - Where information is lacking or incomplete, can data be (re)constructed?
 - What counterfactuals need to be considered?

4. Summarise important findings and conclusions related to each of the above tasks.
5. Develop recommendations under main options for going forward using 2-3 categories: i.e. (1) a gold standard option, (2) a bare-minimum option satisfying internationally-accepted standards for impact evaluation, and (3) if/as appropriate, an option representing the best combination of good evaluation-good value for money given the present situation.
6. From the previous point, develop impact evaluation design frameworks for the various options/contexts identified.

Information Sources/Approach:

- Briefing sessions with representatives from Norad's Department for Global Health and AIDS (AHHA) and Norad's Evaluation Department (EVAL).
- Interviews with key stakeholders in the Ministry of Foreign Affairs (MFA), as well as other key informants.
- Existing project documents and baseline studies (will be made available to the Team upon arrival in Oslo.)
- Existing relevant research and evaluations

Reporting:

- Presentation of the tender proposal to be made to interested parties when contract is awarded.
- Inception Report after no more than 3 person-weeks elapsed work. Comments from interested parties.
- Draft Report after no more than 8 person-weeks elapsed work. Comments from interested parties.
- Final Report.

Management

The study will be managed by the Evaluation Department, Norad (EVAL). An independent team of researchers or consultants will be assigned the study according to prevailing regulations on public procurement in Norway. The team leader shall report to EVAL on the team's progress, including any problems that may jeopardize the assignment. The main stakeholders in the study will be asked by EVAL to comment on the following evaluation products: inception report, draft report and final report. Reports will be submitted to EVAL. EVAL will be in charge of all communication with the team and for approving reports.

Interested parties/stakeholders:

- Representatives from Norad's Department for Global Health and AIDS (AHHA), relevant sections in MFA, embassies and the Norwegian Knowledge Centre for the Health Services (K-Centre).

The team requirements

The team shall comprise two experienced individuals who in combination possess the following qualifications:

- higher relevant academic degrees

- advanced knowledge of evaluation standards and principles in the context of international evaluation
- documented competence in conducting impact evaluations using counterfactuals and baseline surveys
- documented knowledge of and experience from work in international health issues with special emphasis on developing countries
- documented knowledge of maternal and infant mortality and surrounding issues related to access and quality of outreach and facility-based health care and treatment services.
- good knowledge of Norwegian development cooperation policies and instruments;
- language qualifications: Fluency in English (speak, read, write), in addition at least one of the consultants should be able to understand (read) Norwegian.

Deliverables and budget

The assignment is estimated to 10 person-weeks. The deliverables in the consultancy consist of following outputs:

- Inception Report not exceeding 10 pages shall be prepared.
- Draft Final Report for feedback from the stakeholders and EVAL.
- Final Study Report maximum 30 pages - prepared in accordance with EVAL's guidelines given in *Annex A-3 Guidelines for Report* of this document.
- Seminar for dissemination of the final report in Oslo to be organised by EVAL.

Appendix B: Persons Interviewed

Cliff Wang	Senior Adviser, Global Health and AIDS Department, Norad
Paul Richard Fife	Head, Health Unit Dept. of Human Development and Service Delivery, Norad
Lene Lothe	Senior Adviser, Global Health and AIDS Department, Norad
Ingvar Theo Olsen	Senior Adviser, Global Health and AIDS Department, Norad
Helga Fogstad	Senior Adviser; Global Health and AIDS Department, Norad
Dr. Kaliprasad Pappu	National Coordinator, NIPI-UNOPS, New Delhi, India
Urvashi Chandra	Advisor- Monitoring & Evaluation, NIPI-UNOPS New Delhi, India
Terje Thodesen	Counsellor, Royal Norwegian Embassy, Islamabad
Tomas Alme	NIPI-UNOPS, New Delhi, India
Hanne Tilrem	First Secretary, Royal Norwegian Embassy, Dar es Salaam, Tanzania
Dr Carolyn Sunners	Health Adviser & Deputy Head, DFID Northern Nigeria Office

Appendix D: Bibliography

Main Report

3ie (2009) 3ie impact evaluation glossary. International Initiative for Impact Evaluation: New Delhi, India

Basinga P, Gertler PJ, Binagwaho A, Soucat A, Sturdy JR, C.M.J. Vermeersch. January 2010. Paying Primary Health Care Centers for Performance in Rwanda. . The World Bank. Policy Research Working Paper 5190.

Bryce J, Gilroy K, Jones G, Hazel E, Black RE, Victora CG. 2010. The Accelerated Child Survival and Development programme in west Africa: a retrospective evaluation. *Lancet*; 375: 572–82. Published Online. January 12, 2010.

Gilroy K, Hazel E, Callaghan J, Bryce J and the IIP-JHU Catalytic Initiative evaluation group Undated. Guidelines for documenting program implementation and contextual factors in independent evaluations of the Catalytic Initiative. Working Paper, v. 1.1. Institute for International Programs Johns Hopkins Bloomberg School of Public Health.

The Global Campaign for Health Millennium Development Goals. Accessed at: <http://www.norad.no/en/Thematic+areas/Health+and+aids/Maternal%2C+child+and+women%27s+health/Global+campaign+for+the+health+MDGs>.

Habict JP, Victora CG and JP Vaughan. 1997. Linking evaluation needs to design choices. UNICEF Staff Working Paper. Evaluation and Research Series. EVL-97-003.

Haddad S, Bicaba A, Feletto M, Taminy E, Kabore M, Ouédraogo B, Contreras G, Larocque N, Fournier P. 2009. System-level determinants of immunization coverage disparities among health districts in Burkina Faso: a multiple case study *BMC International Health and Human Rights*. Vol. 9(Suppl 1).

Khandker SR. 2010. Handbook on impact evaluation: quantitative methods and practices. The World Bank.

OECD. Undated. Glossary of key terms in evaluation and results based management. Development Assistance Committee.

OECD. The Paris Declaration on Aid Effectiveness (2005) and Accra Agenda for Action (2008).

Rivera JA, Sotres-Alvarez D, Habicht JP, Shamah T and S Villalpando. 2004. Impact of the Mexican program for education, health and nutrition (Progresá) on rates of growth and anemia in infants and young children. *JAMA*. Vol. 291 No. 21.

UNICEF. Final Report. Retrospective Evaluation of ACSD: Ghana. Submitted to UNICEF on 7 October 2008. Institute for International Programs Johns Hopkins Bloomberg School of Public Health Baltimore, MD.

Victora CG, Black RE, Boerma JT, Bryce J. 2010. Measuring impact in the Millennium Development Goal era and beyond: a new approach to large-scale effectiveness evaluations. *www.thelancet.com* Published online July 9, 2010.

Victora CG, Black RE, Bryce J. 2009. Evaluating child survival programmes. *Bull World Health Organ*. 87:83.

Partnership Initiatives, by country:

Pakistan:

1. NPPI Project Document. Reducing maternal, neonatal & child mortality in Sindh. August 25 2008. Delivering as One. United Nations in Pakistan. Ministry of Health/Government of Pakistan. Norway-Pakistan Partnership Initiative.
2. Memorandum of Understanding between The Government of the Kingdom of Norway and The Islamic Republic of Pakistan regarding The Norway-Pakistan Partnership Initiative (NPPI).
3. Result Based Financial Mechanisms for Improving Maternal, Newborn and Child Health Outputs: A Feasibility Study for 10 Selected Districts of Rural Sindh. Research Report Commissioned by MNCH Program, Sindh and UNICEF. Funded by Norwegian Pakistan Partnership Initiative [NPPI]. Prepared by Aga Khan University, Karachi March 2009.
4. Standard Administrative Arrangement between the Kingdom of Norway, the United Nations Resident Coordinator and the United Nations Development Programme.
5. NPPI Baseline Survey in Sindh. Report 1: Baseline Household Survey. Submitted to UNICEF Sindh. Prepared by: Tauseef Ahmed PhD. 30 March 2009.
6. GAVI Alliance Annual Progress Report 2009. Submitted by the Government of Pakistan. 15 May 2010.
7. <http://www.mnch.gov.pk/fl.php>. Accessed 29 November 2010.
8. <http://www.sindhhealth.gov.pk/history.htm>. Accessed 29 November 2010.
9. National Institute of Population Studies (NIPS) [Pakistan], and Macro International Inc. 2008. Pakistan Demographic and Health Survey 2006-07. Islamabad, Pakistan: National Institute of Population Studies and Macro International Inc.
10. GAVI Alliance. GAVI Alliance Health System Strengthening (HSS) Applications. Pakistan. May 2007.
11. National Maternal Newborn and Child Health (MNCH) Program. 2006 – 2012. Government of Pakistan. Ministry of Health. November 2006.
12. Norway Pakistan Partnership Initiative (NPPI). Core Action Plan April-December 2009.

13. Norway Pakistan Partnership Initiative (NPPI). Progress Update (May 2010) WHO-UNICEF-UNFPA Activities.
14. Norway Pakistan Partnership Initiative (NPPI). Core Action Plan January -December 2010.
15. <http://www.thardeep.org>. Accessed 29 November 2010.
16. United Nations Children's Fund. Executive Board. Annual session 2008. 3-5 June 2008. Short-duration country programme document. Islamic Republic of Pakistan. E/ICEF/2008/P/L.11
17. Executive Board of UNFPA and UNDP. Annual session 2008. 16 to 27 June 2008, Geneva. Item 4 of the provisional agenda. UNFPA - Country programmes and related matters. United nations Population Fund. Extensions of country programmes in the Asia and the Pacific region. Note by the Executive Director.
18. Sindh Development Review 2008-09. People's Primary HealthCare Initiative (PHHI). Planning and Development Department. Government of Sindh. www.Sindhpn.gov.pk.
19. Ministry of Health (MoH) and Technical Resource Facility (TRF). Draft Aide Memoire. Second Mini Review of the National Maternal, Newborn and Child Health (MNCH) Programme 12-14 April, 2010.

Malawi:

1. Malawi RBF. Pilot on Results-based Financing (RBF) for Maternal and Newborn Health in Malawi – September 2010. Norad.
2. Potential Result-based Financing for reduced Maternal and Child Mortality in Malawi - Background document. Norad.
3. Norwegian-German Initiative to support MDG's 4 and 5. Feasibility Study on an MCH programme in Malawi using the instrument of RBF. Phase 1. FINAL Draft Report. December 2009. David Griffith, Brigitte Jordan-Harder, Alice Maida.
4. Results-based Financing (RBF) for Maternal and Newborn Health in Malawi. Minutes of Mission. Lilongwe, 21 June 2010
5. A concept note for the evaluation of result based funding mechanism of maternal and child health services in Malawi. Prepared by Jobiba Chinkhumba MBBS MSc., Department of Community Health College of Medicine.
6. Ministry of Health. March 2007 Malawi National Health Accounts (NHA). 2002-2004 with Subaccounts for HIV and AIDS, Reproductive and Child Health, Department of Health Planning & Policy Development, Lilongwe, Malawi. Abt
7. Results-based financing for health. Health results-based financing impact Evaluation network. World Bank.
8. Decision Document – Project and Programme Support. Royal Norwegian Embassey Lilongwe.

India:

1. Norway-India Partnership Initiative. Final Report Of The Mid-Term Review. February 2010. Dr Ashok Dutta, Dr. Rani Gera, Dr Antoinette Pirie and Mr Stein-Erik Kruse.
2. Norway India Partnership Initiative. Annual Report 2008. UNOPS. UNICEF. WHO.
3. Minutes of the meeting of the 7th PMG held on 17-04-09.
4. Short summary of NIPI interventions in Bihar. Presented by Ravi Parmar. Executive Director, SHSB. 13-11-09.

5. NIPI Strategy Document. 2008.
6. Norway-India Partnership to achieve MDG4 (NIPI Program Document).
7. UNICEF-NIPI Implementation Status 2009 and Plan for 2010. 8th PMG mtg. 13-11-09.
8. Norway-India Partnership Initiative. Update on WHO Activities. 8th NIPI Programme Management Group Meeting. 13 November 2009.
9. Minutes of the 6th JSC held on 28-08-2008.
10. Draft minutes of the 7th JSC held on 04-06-2009.
11. Minutes of the Meeting of the PMG held on 21-08-2008
12. Short summary of NIPI interventions in Madhya Pradesh. Dr. Manohar Agnani. MD NRHM. Madhya Pradesh. 13-11-09.
13. Short summary of NIPI interventions in Orissa, Shri G. Mathivathanan, IAS. Mission Director. NRHM, Orissa. 13-11-09.
14. Short summary of NIPI interventions in Rajasthan. Dr M L Jain, Director (RCH), Rajasthan. PMG mtg. 13-11-09.
15. Summary agenda notes for 8th Program Management Group meeting. Norway India Partnership Initiative on November 13th 2009. New Delhi.
16. Summary Report of Assignment for Norad AHHA re: NIPI, October-December 2008. Cliff Wang, HeSo.
17. Visit to Bihar State and Nalanda District, November 2008. Travel Report. By Cliff Wang and Olav Hernar.
18. NIPI issues in advance of the mid-term review. Ambassaden i New Delhi, India. Kopi : UD- GIL UD- REG. 19.08.2009.
19. NIPI. Monitoring and Evaluation Strategy and Plan, 2010-2013. DRAFT version 4.
20. International Institute for Population Sciences (IIPS), 2010. District Level Household and Facility Survey (DLHS-3), 2007-08: India. Bihar: Mumbai: IIPS. Accessed at www.rchiips.org.
21. International Institute for Population Sciences (IIPS), 2010. District Level Household and Facility Survey (DLHS-3), 2007-08: India. Madhya Pradesh: Mumbai: IIPS. Accessed at www.rchiips.org.
22. International Institute for Population Sciences (IIPS), 2010. District Level Household and Facility Survey (DLHS-3), 2007-08: India. Orissa: Mumbai: IIPS. Accessed at www.rchiips.org.
23. International Institute for Population Sciences (IIPS), 2010. District Level Household and Facility Survey (DLHS-3), 2007-08: India. Rajasthan: Mumbai: IIPS. Accessed at www.rchiips.org.
24. Baseline Survey On Child And Related Maternal Health Care. Rajasthan. Revised Report. Prepared For: Norway India Partnership Initiative, New Delhi. DRS Pvt. Ltd. GVL Narsimha Rao Development & Research Services Private. August 2009.
25. Baseline Survey On Child And Related Maternal Health Care. State Report: Bihar. Prepared For: Norway India Partnership Initiative, New Delhi. DRS Pvt. Ltd. GVL Narsimha Rao Development & Research Services Private. August 2009.
26. Baseline Survey On Child And Related Maternal Health Care. Madhya Pradesh. Revised Report. Prepared For: Norway India Partnership Initiative, New Delhi.

DRS Pvt. Ltd. GVL Narsimha Rao Development & Research Services Private.
August 2009.

27. Baseline Survey On Child And Related Maternal Health Care. Orissa. Revised Report. Prepared For: Norway India Partnership Initiative, New Delhi. DRS Pvt. Ltd. GVL Narsimha Rao Development & Research Services Private. August 2009.
28. Activities by State Health Societies in Orissa, Bihar, Madhya Pradesh and Rajasthan under Norway – India Partnership Initiative. Program Update, January 2009.
29. Concurrent Assessment of Janani Suraksha Yojana (JSY) in Selected States. Bihar, Madhya Pradesh, Orissa, Rajasthan, Uttar Pradesh. UNFPA. 2009.
30. Coverage Evaluation Survey 2009 (CES 2009). UNICEF.
31. Activities by State Health Societies in Orissa, Bihar, Madhya Pradesh and Rajasthan under NIPI. Program Update. January 2009.
32. NIPI Expenditure Report. (Excel spreadsheet prepared for MTR committee).
33. UNICEF ELA Appendix 1: Project Summary (13 December 2006)
34. UNOPS ELA Appendix 1: Project Summary (23 November 2006)
35. WHO ELA Appendix 1: Program Summary (13 December 2006)

Nigeria:

1. UK-Norway MNCH Initiative in Northern Nigeria. Outline of Joint Programme. May 2008.
2. The UK-Norway MNCH Initiative in the Northern States Of Nigeria. Desk Appraisal. April 2008. Mercy Bannerman. Nordic Consulting Group
3. PRRINN-MNCH Annual Review Draft Narrative Report, Carol Bradford and Maisha Strozier, 10 March 2010 Appendix 11.
4. Project Memorandum. Nigeria Northern States. Maternal, Newborn And Child Health Programme. April 2008.
5. PRRINN-MNCH Baseline Studies 2009. Summary Report.
6. Arrangement On Delegated Cooperation Between UK Department For International Development (Dfid) and the Norwegian Ministry Of Foreign Affairs (MFA) regarding support to Maternal, Newborn And Child Health Initiative In Northern Nigeria. Signed 13 June 2008.
7. Northern States Maternal, Newborn and Child Health Programme. CNTR 200808350. Part B – General and Technical Proposal. Submitted by Health Partners International in Joint Venture with Save the Children UK, and GRID Consulting Ltd.
8. 2009 Annual Report. PRRINN-MNCH.
9. PRRINN-MNCH Annual Review. Narrative Report. Carol Bradford and Maisha Strozier. 24 March 2010. Dfid Human Resource Development Center.
10. Project concept note. Reviving routine immunization, Nigeria. Mis code: 048-555-xxx. 06-09-04.
11. Joint logframe and Final M&E Framework. PRRINN-MNCH

Tanzania:

1. Appraisal of the Norway Tanzania Partnership Initiative Programme. Report to the Norwegian Agency for Development Cooperation. John James, Wanjiku Kamau. August 2007.

2. Performance-Based Financing Report On Feasibility and Implementation Options Final Report. Paul Smithson, Dr Nelly Iteba, Oscar Mukasa, Dr Ali Mzige, Max Mapunda, Gradeline Minja. September 2007
3. Norway Tanzania Partnership Initiative. Programme Document. Submitted by the Ministry of Health and Social Welfare of Tanzania to the Royal Norwegian Embassy.
4. Memorandum of Understanding between the Partners (Government of Tanzania and Donors participating in the pooled funding("basket financing") of the health sector concerning the pooled funding for the Government of Tanzania's Health Sector Programme based on Second Health Sector Strategic Plan and the Health Sector medium term expenditure Framework. July 2003; June 2008.
5. Results-Based Bonus Design, Implementation & Budget "Malipo kwa Ufanisi Bora katika Huduma za Afya" (MUBHA). Final Report. Report of the design team describing detailed model, implementation strategy and funding requirements. Paul Smithson, Dr Rena Eichler, Dr Samson Winani, Dr Aziz Msuya, Ingvar Theo Olsen, Erica Musch. Commissioned by Royal Embassy of Norway, Tanzania. 20th February 2008
6. The United Republic Of Tanzania. Ministry Of Health And Social Welfare. Agenda No. 5.1. Payment For Performance Strategy 2008-2015. December, 2008.
7. The United Republic Of Tanzania. Ministry Of Health And Social Welfare. Agenda No. 5.2 Implementation Guideline. Payment For Performance. December, 2008.
8. Memorandum of Understanding between the Partners (Government of Tanzania and Development Partners participating in the pooled funding("basket") of the health sector concerning the pooled funding for the Government of Tanzania's Health Sector Programme based on Second and Third and the Health Sector medium term expenditure Framework based on the Ministry Of Health And Social Welfare and Prime Minister's Regional Administration and Local Government and the Comprehensive Council Health Plans. July 1, 2008 to June 30, 2015.
9. Centre For Health And Social Development. Payment For Performance Appraisal. Report To Norad And The Royal Norwegian Embassy, Tanzania. Marilyn Lauglo and Goddfrey B. R. Swai. Final Report. 2 April 2009.
10. Programme document. Tanzania P4P for maternal and newborn health pilot. Undated. No author.
11. Draft Programme document. Pilot to improve P4P for maternal and newborn health in Tanzania. Undated. No author.
- 11bis. Pay for Performance in Tanzania. Undated. No author.
12. Informing the Design of a Pay for Performance Initiative Pilot. Report prepared by Gregory Kabadi, Dominic Moshia, Flora Kessy and Josephine Borghi. June 2010.
13. A Review Of The Status Of HMIS Data Sources. By Gregory Kabadi and Dr Dominic Moshia, Ifakara Health Institute. Undated.
14. Paying for performance to improve the delivery of health interventions: an outline of good practices and lessons in relation to the design of the P4P pilot in Tanzania. Flora Kessy (PhD). Draft, 16th May 2010.

15. Learning from current P4P initiatives in Tanzania – An Assessment of possible linkages Report prepared for the Ministry of Health and Social Welfare and Norad. Josephine Borghi. May 2010.
16. Pay for Performance (P4P) in Tanzania – Status September 2010.

Appendix E: Norway - Pakistan Partnership Initiative Profile

Overview	The RNE is investing NOK 250 million in support of the national Maternal, Newborn and Child Health Program through implementation in 10 districts of Sindh province. The RNE has partnered with One UN Program mechanism (with UNDP playing an administrative role and UNICEF, WHO and UNFPA playing implementation roles). A main program element is contracting out of MNCH services through private sector providers and facilities.									
PI duration	Start date 2008/09 (MOU signed 30/01/09)	End date 2013								
PI purpose	<ul style="list-style-type: none"> • The Initiative aims at reducing maternal and child mortality rates with about 40% during the five year period for project implementation, from 2008/2009 till 2013, through a variety of carefully selected interventions from government and non-government agencies (MOU). • The goal of NPPI is to reduce maternal, newborn and under-five mortality in 10 selected districts in Sindh Province in Pakistan. • The purpose is to increase provision of and access to MNCH interventions for the poor and socially excluded in Sindh Province, as well as to raise demand and utilization for those services 									
Stated objective	The expected outcomes to attain the above goals by: increased coverage of quality MNCH/FP service; and improved MNCH/FP self care and care-seeking behaviour among families and communities.									
Funding	<p>Total: NOK 250 million (Norwegian Kroner fifty million approximately USD 50 million) for the five year period from 2008/2009 till 2013. Pakistan will support the NPPI through its MNCH program. Three UN agencies working together within the One UN program mechanisms – UNICEF, WHO and UNFPA.</p> <table border="1"> <thead> <tr> <th>By partner (based on 2010 workplan):</th> <th>By activity area (based on 2010 workplan):</th> </tr> </thead> <tbody> <tr> <td>UNICEF: 40%</td> <td>1. Output 1.1 Integrated MNCH/FP care made available through contracting (incl. public private partnerships) 2010- \$ 5.99m</td> </tr> <tr> <td>WHO: 42%</td> <td>2. Output 1.2 Improved governance and results based management 2010- \$1.24m</td> </tr> <tr> <td>UNFPA: 18%</td> <td>3. Output 1.3 Operational Research conducted to produce knowledge and improve future decision making related to increasing MNCH/FP coverage and self care 2010 – \$2.07m</td> </tr> </tbody> </table>		By partner (based on 2010 workplan):	By activity area (based on 2010 workplan):	UNICEF: 40%	1. Output 1.1 Integrated MNCH/FP care made available through contracting (incl. public private partnerships) 2010- \$ 5.99m	WHO: 42%	2. Output 1.2 Improved governance and results based management 2010- \$1.24m	UNFPA: 18%	3. Output 1.3 Operational Research conducted to produce knowledge and improve future decision making related to increasing MNCH/FP coverage and self care 2010 – \$2.07m
By partner (based on 2010 workplan):	By activity area (based on 2010 workplan):									
UNICEF: 40%	1. Output 1.1 Integrated MNCH/FP care made available through contracting (incl. public private partnerships) 2010- \$ 5.99m									
WHO: 42%	2. Output 1.2 Improved governance and results based management 2010- \$1.24m									
UNFPA: 18%	3. Output 1.3 Operational Research conducted to produce knowledge and improve future decision making related to increasing MNCH/FP coverage and self care 2010 – \$2.07m									

	4. Output 2.1 Strengthened community based & Outreach MNCH/FP Care services 2010 - \$. 606m
	5. Output 2.2 Voucher/incentive schemes implemented to increase demand and service utilization 2010 - \$.309m
	6. Output 2.3 Community networks for advocacy/ Social mobilization/BCC and awareness 2010 - \$.261m
PI within landscape for MDGs 4&5 / development partners	<p>At national level, the national MNCH programme started at the end of 2006 with a budgetary allocation of US\$ 333.33 million over a six year project life from 2006-2012. Sixty percent (60%) of the project cost is to be borne by the GoP, while DFID is contributing 40% of the overall budget¹⁷. In addition, other development partners such as UNICEF, USAID, UNFPA, WHO and GAVI have formally committed resources to the National MNCH Programme. UNICEF supports 34 districts nationwide with support for MNCH implementation with a budget of US \$ 59 million (regular and other sources) over the period 2009-10.</p> <p>Within the 10 selected districts, current estimated allocation by the National Program for FP and PHC is US\$ 4.9 million. The total allocation by Department of Sindh is estimated to be around US\$ 15.4 million for recurrent & development cost. The 10 selected districts will have the following estimated incremental increase during 2006-2012:</p> <ul style="list-style-type: none"> • US\$16.7 million from the National MNCH Programme • US\$ 4.1 million from the GAVI Health Systems Strengthening (2008-2009)¹⁸ • US\$50.0 million from NPPI (2008-2012) • UNICEF, UNFPA and WHO are also working in Sindh province in the area of MNCH, immunization and family planning but at present mostly in other districts. <p>NPPI documentation reviewed by the Team makes no mention of an existing public partnership program between the Sindh Rural Support program/People's Primary Healthcare Initiative (SRSP/PPHI) and the Government of Sindh. In this program, the government has agreed to transfer management of BHUs, dispensaries, and MCHCs along with existing budgetary share to SRSO/PPHI. The program was initially started in five districts and later extended to seventeen districts of Sindh. These districts include Larkana, Badin, Umerkot, Kashmore (four of the ten NPPI districts). Overall, the PPHI covers 883 health facilities in Sindh.</p>
Geographic areas	Ten rural districts in Sindh province which represent 29% of the total province population: Jamshoro, Badin, Tharparkar, Umarkot, Nawabshah, Larkana, Kambar, Shikarpur, Ghotki and Kashmore. The population is mainly scattered into small villages making physical access to health care services difficult.

17 DFID funding to the MNCH-program has been suspended based on issues including the results of an audit report (Punjab) and shortfalls in the GOP contributions to the MNCH-program.

18 GAVI Health Systems Strengthening grant will provide an US\$ 23 million for all of Pakistan.

Criteria for selection	According to the PD, the 10 districts were selected on the basis of their prevailing high MMR, NMR and under five child mortality rates (based 2004 MICS survey data) as well as poorly functioning health systems. As the MICS 2004 was not available for review, it is difficult for the team to independently assess the basis for selection. It seems improbable that the MICS would have provided district-level mortality estimates and that these estimates would differ significantly from other districts in the province. Other undocumented factors are assumed to have played a role in district selection.	
Partners description	Partners	Responsibilities
	Gov't. of Pakistan	<ul style="list-style-type: none"> • supports NPPI through its MNCH program • all sectors of government cooperate with NPPI when such cooperation is called for
	UNDP	Administers the Norwegian contribution to the Pakistan One Fund; collates reports of implementing partners into a single Progress Report
	UNICEF	Implementing partner under the One UN in Pakistan agreement which carries out activities in line with the budget contained in the programmatic documents
	WHO	Implementing partner under the One UN in Pakistan agreement which carries out activities in line with the budget contained in the programmatic documents
UNFPA	Implementing partner under the One UN in Pakistan agreement which carries out activities in line with the budget contained in the programmatic documents	
Institutional arrangements	RNE and UNDP have agreement for RNE funding for NPPI to be administered through the Pakistan One Fund and implemented through three UN agencies: UNICEF, UNFPA and WHO. Implementing agencies have Sindh offices and staff. Activities managed through annual work plans, progress reports and monitoring visits. In principle, UNDP is to provide consolidated progress reporting to RNE.	
Implementation timeline	Planned	Actual
	No detailed NPPI implementation plan beyond PD and annual work plans of implementing agencies.	First tranche of funding disbursed in early 2009 (month?). Delay ensued while implementing arrangements were worked out between UNICEF, UNFPA and WHO. Implementing agencies working at differing paces. Major flooding in summer 2010 lead to requests for re-programming of available NPPI funds for relief operations.

Elements of Innovation (as described in PD)

According to the PD, public-private policy work was to be supported as an element of the provincial reform processes and was to include setting rigorous performance benchmarks and evaluations for PPP efforts. Other than the Feasibility Study, available documentation provided little information as to how this PPP was being made operational. The OR identified below (on contracting out) is presumably the activity associated with this innovation. Team to acquire these materials in their current form.

A fund for innovations equal to US\$20,000 was to be allocated annually to try out different new ways of tackling problems and using operations research to document impact and lessons learnt (see section immediately below).

Operations Research activities	Activity	Status (2010 workplan)
	Reduction/prevention of low birth weight and maternal anaemia in NPPI districts (UNICEF)	Continuation of TA initiated in 2009
	<i>Support to carry out OR on Developing & Testing Models of Public Private Partnerships” for contracting out MNCH services (UNICEF)</i>	<i>Develop ToR for consultant; identify and hire consultant; award contract; initiate OR</i>
	<i>Support to carry out OR on Implementing incentive/voucher schemes for increasing demand and uptake of key MNCH services” (UNICEF)</i>	<i>Develop ToR for consultant; identify and hire consultant; award contract; initiate OR</i>
	Operation Research on an intervention package for the management of maternal Anaemia and LBW in the target districts (WHO)	
	Initiate operations research on reduction of Maternal Mortality and TFR (UNFPA)	Develop ToR for consultancy, develop bidding documents, hiring of firm/institute for research
	Initiate operations research on incentives/ CCTs (UNFPA)	Develop ToR for consultancy, develop bidding documents, hiring of firm/institute for research

Appendix F: Norway - Pakistan Partnership Initiative Monitoring & Evaluation Plan

This table shows a compilation of indicators drawn from available program documentation – it does *not* represent the Team’s recommendations on appropriate indicators for evaluating PI performance or impact.

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	Data sources	Arrangements / Responsibility for data collection
Impact						
1. MMR per 100,000 live births	1	340/350	200			
2. U5MR per 1,000 live births	1	103/120	65	Pakistan Family Planning and Reproductive Health Survey 2001 – 02		
3. IMR per 1,000 live births	1	77/81	55	Pakistan Family Planning and Reproductive Health Survey 2001 – 02		
4. Neonatal mortality rate per 1,000 live births	1	54	36			
5. % newborns who weigh less than 2,500 grams	1	25%	15%			
6a. % under-5s below minus 2 standard deviations from median weight-for-age of WHO reference population*	1	49%	25%	MICS 2004		
6b. % under-5s below minus 3 standard deviations from median weight-for-age of WHO reference population*	1	17%	9%	MICS 2004		
7. % under-5s with diarrhoea during past 2 weeks*	1	33% 28.1%	5%	MICS 2004 Baseline HH survey		

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	Data sources	Arrangements / Responsibility for data collection
8. % under-5s with ARI within past 2 weeks*	1	48% 27.5%	5%	MICS 2004 Baseline HH survey		
9. % under 5s with iron deficiency anaemia	1	40%	20%			
10. % households using iodized salt	1	27.5% 9.7%	90%	Baseline HH survey		
Outcome 1: Coverage of MNCH services: Maternal, Neonatal, Under 5.						
11. % under-5s with diarrhoea in the last two weeks and were treated with oral rehydration salts or an appropriate household solution (ORT)*	1	59% 84.1%	90%	MICS 2004 Baseline HH survey		
12. % under-5s with suspected pneumonia receiving antibiotics (note that alternative definition used in baseline household survey)	1	16% 59.4%	90%	Baseline HH survey		
13. % under-5s who have received Vitamin A twice a year	1			MICS 2004		
14. % children under 6 months who are exclusively breastfed*	1	24%	90%	MICS 2004		
15. % children 12-23 months immunized against six immunisable childhood diseases*	1	37% 45.9%	90%	MICS 2004 Baseline HH survey		
16. % children less than 1 immunised against measles (Household survey children 12-13 months)	1	57.5%		Baseline HH survey		

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	Data sources	Arrangements / Responsibility for data collection
17. % newborns experiencing danger signs who receive adequate care and are referred	1	15%	90%***			
18. % women aged 15-49 who are using (or whose partner is using) a contraceptive method. (married women in HH survey)	1	30% 20.8%	60%	Baseline HH survey		
19. % pregnant women with at least 1 ANC visit to a skilled birth attendant	1	53% 64%	95%	Baseline HH survey		
20. % births attended by a skilled health attendant (SBA)	1	29% 43.7%	60%	Pakistan Family Planning and Reproductive Health Survey 2001 – 02 Baseline HH survey		
21. % births attended by a skilled health attendant at a health facility (simply at a facility in HH survey)	1	25% 42.5%	50%	Baseline HH survey		
22. % new mothers receiving postpartum care	1	15% 43.9%	60%	Baseline HH survey		
23. % women experiencing obstetric complications receive EmOC	1	30%	90%**			
24. % caretakers of children 0-59 months who know at least two of the following signs of seeking care immediately: child not able to drink and breastfeed, child becomes sicker, child develops a fever, child has fast breathing, child has difficulty breathing, blood in stools* (Question in HH survey but indicator not in report)	1	40%	90%	MICS 2004		

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	Data sources	Arrangements / Responsibility for data collection
25. Every maternal, newborn and child complication has access to Basic and Comprehensive EmONC services within two hours	1	5%	90%			
Output 1.1: Integrated MNCH Care available through contracting out services						
# facilities providing contracted out programs	1	None	112 private & public			
# facilities providing BEmONC per 500,000 pop.	1	?	4			
# facilities providing CEmONC per 500,000 pop.	1	?	1			
# facilities implementing IMNCI strategy	1	?	370			
# facilities providing ANC services	1	?	370			
# facilities providing SBA services	1	?	112			
# facilities providing family planning services	1	?	370			
Output 1.2: Improved governance and results based management.						
# districts with performance based agreements	1	0	10			
# staff completed planned management training	1	0	40			
# districts with a PPP coordinator	1	0	10			
# district management structures ISO certified	1	0	10			

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	Data sources	Arrangements / Responsibility for data collection
Output 1.3: Operational and evaluation research conducted to produce knowledge and improved future decision making related to increased coverage and self care for MNCH						
% planned studies completed	1	0	4 specified studies+			
Outcome 2: Improve MNCH health self-care and care-seeking behaviour						
Knowledge level among mothers for recognition of danger signs	1		30% increase			
Output 2.1: Strengthened community based & outreach MNCH Care services						
% Communities (villages) with strong community based MNCH programs	1	None	100%			
% LHW/CHW trained in newborn care and counselling infants and young children	1	None	100%			
Mother and Child weeks	1	None	10 completed			
% Children under five registered	1	None	100%			
% Children (0-23 months) receiving required catch-up immunization	1		100%			
Coverage of ANC & FP outreach services	1		80% of areas not covered by static facilities			
% Children 2-5 years de-wormed	1	?	100%			

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	Data sources	Arrangements / Responsibility for data collection
% Children 6 months to 5 years receiving Vitamin A supplementation	1	?	100%			
% Pregnant mothers in malaria endemic districts with ITN	1	?	100%			
% Pregnant mothers provided with Clean Delivery Kits	1	?	100%			
# Facility staff in outreach interventions trained	1	None	All staff trained			
# Health education sessions provided by LHW & CHW	1		60 sessions			
Output 2.2: Voucher/incentive schemes implemented to increase demand and service utilization						
% of UCs in 10 districts with voucher scheme in place	1	None	15%			
% of pregnant women using VS/incentive schemes in targeted UCs	1	None	35%			
Output 2.3: Community networks for MNCH/FP advocacy and mobilization established & behaviour change communication and awareness raising programme implemented						
% Villages where functional community networks established	1	None	At least 80%			
# Quality acceptable and effective mass media and community campaigns conducted per year	1	None	5			

Table 2: Norway–Pakistan PI: Proposed streamlined Indicator set

Indicators (as included in PI documents)	Keep	Revise	Remove	MNCH aligned?	Notes
Impact					
1. MMR per 100,000 live births			✓	✓	Long-term impact not measurable over the life of the project; not attributable to NPPI; not measured in NPPI baseline
2. U5MR per 1,000 live births			✓	✓	Not measurable over the life of the project; not attributable to NPPI; not measured in NPPI baseline
3. IMR per 1,000 live births			✓	✓	Not measurable over the life of the project; not attributable to NPPI; not measured in NPPI baseline
4. Neonatal mortality rate per 1,000 live births			✓	✓	Not measurable over the life of the project; not attributable to NPPI; not measured in NPPI baseline
5. % newborns who weigh less than 2,500 grams			✓	✓	In baseline survey, 55% of women say no weight taken at birth. Of those whose newborns were weighed, 55% do not know the weight. Therefore, valid data on birth weight is only available for approximately 25% of births. Not clearly related to NPPI interventions
6a. % under-5s below minus 2 standard deviations from median weight-for-age of WHO reference population*			✓		Long-term impact not measurable over the life of the project; improving nutritional status not clearly related to NPPI interventions; not measured in NPPI baseline
6b. % under-5s below minus 3 standard deviations from median weight-for-age of WHO reference population*			✓		Not clearly related to NPPI interventions; not measured in NPPI baseline
7. % under-5s with diarrhoea during past 2 weeks*			✓		Prevention of diarrhea not clearly linked to NPPI interventions

Indicators (as included in PI documents)	Keep	Revise	Remove	MNCH aligned?	Notes
8. % under-5s with ARI within past 2 weeks*			✓		Prevention of ARI not clearly related to NPPI interventions
9. % under 5s with iron deficiency anaemia			✓		Reduction of iron deficiency anaemia not clearly related to NPPI interventions
10. % households using iodized salt			✓		Not clearly related to NPPI interventions
11. % under-5s with diarrhoea in the last two weeks and were treated with oral rehydration salts or an appropriate household solution (ORT)*	✓				
12. % under-5s with suspected pneumonia receiving antibiotics (cough and fever in HH survey)	✓				
13. % under-5s who have received Vitamin A twice a year *	✓				
14. % children under 6 months who are exclusively breastfed*	✓			✓	
15. % children 12-23 months immunized against six immunisable childhood diseases*	✓			✓	
16. % children less than 1 immunised against measles*	✓			✓	Clarify that baseline HH survey calculated by 12 months of age; interpretation of indicator clouded by campaign approaches to measles eradication – not necessarily strengthening of routine systems
17. % newborns experiencing danger signs who receive adequate care and are referred			✓		Not a standard newborn care indicator; depends on proper recognition of newborn danger sign among women;

Indicators (as included in PI documents)	Keep	Revise	Remove	MNCH aligned?	Notes
18. % women aged 15-49 who are using (or whose partner is using) a contraceptive method. (married women in HH survey – note indicator in summary (22.3) differs from source table 8.5)	✓				MNCH uses unmet need
19. % pregnant women with at least 1 ANC visit to a skilled birth attendant	✓				
20. % births attended by a skilled health attendant (SBA)	✓			✓	
21. % births attended by a skilled health attendant at a health facility (simply at a facility in HH survey)			✓		Not a recognized intl. standard; baseline does not measure (Note: clarify whether baseline includes this)
22. % new mothers receiving postpartum care		✓			Emerging intl. consensus on indicator as
23. % women experiencing obstetric complications receive EmOC			✓		Unclear whether denominator is actual complication or estimated number of complications (as per met need indicator); preference for unmet need indicator
24. % caretakers of children 0-59 months who know at least two of the following signs of seeking care immediately. (Question in HH survey - indicator not in report)	✓				
25. Every maternal, newborn and child complication has access to Basic and Comprehensive EmONC services within two hours		✓			This indicator is not operational in its current form.

Indicators (as included in PI documents)	Keep	Revise	Remove	MNCH aligned?	Notes
# facilities providing contracted out programs		✓			Needs further specification and should include both volume and quality of services delivered;
# facilities providing BEmONC per 500,000 pop.	✓			✓	
# facilities providing CEmONC per 500,000 pop.	✓			✓	
# facilities implementing IMNCI strategy					Needs further specification and should include both volume and quality of services delivered;
# facilities providing ANC services					Needs further specification and should include both volume and quality of services delivered;
# facilities providing SBA services					Needs further specification and should include both volume and quality of services delivered;
# facilities providing family planning services				✓	Needs further specification and should include both volume and quality of services delivered;
# districts with performance based agreements					Indicator should include some form of performance
# staff completed planned management training				✓	Indicator would be improved with the inclusion of a denominator (number of staff targeted for mgt. training)
# districts with a PPP coordinator					
# district management structures ISO certified				✓	
% planned studies completed					
Knowledge level among mothers for recognition of danger signs					Redundant with an indicator above.

Indicators (as included in PI documents)	Keep	Revise	Remove	MNCH aligned?	Notes
% Communities (villages) with strong community based MNCH programs		✓			“strong” must be defined and made operational as an indicator
% LHW/CHW trained in newborn care and counselling infants and young children		✓			
Mother and Child weeks		✓			
% Children under five registered			✓		
% Children (0-23 months) receiving required catch-up immunization			✓		
Coverage of ANC & FP outreach services			✓		
% Children 2-5 years de-wormed			✓		
% Children 6 months to 5 years receiving Vitamin A supplementation		✓			Redundant with measure above; consolidate
% Pregnant mothers in malaria endemic districts with ITN			✓		Not clearly associated with NPPI; may be confounded by role of GF grants;
% Pregnant mothers provided with Clean Delivery Kits	✓				
# Facility staff in outreach interventions trained			✓		Use denominator of number for staff targeted
# Health education sessions provided by LHW & CHW			✓		
% of UCs in 10 districts with voucher scheme in place	✓				
% of pregnant women using VS/ Incentive schemes in targeted UCs	✓				

Indicators (as included in PI documents)	Keep	Revise	Remove	MNCH aligned?	Notes
% Villages where functional community networks established		✓			
# Quality acceptable and effective mass media and community campaigns conducted per year			✓		

Appendix G: Norway - Pakistan Partnership Initiative Evaluation Options Appraisal

Appraisal Area 1: Map the basis for Norway's support to the PI using project documents¹⁹ and assess the adequacy and quality of such documentation in terms of potential "impact evaluability" of the PI towards the end of the initial implementation period (typically 5-6 years).

Team's Appraisal:

To date, the Team has reviewed the NPPI Project Document; Memorandum of Understanding and Agreements; annual work plans (2009, 2010) and progress report (2009); National MNCH Programme document; NPPI baseline reports and data sets; PDHS 2006-07 and information from other development partners. A complete list of materials is provided below.

The NPPI Project Document (PD) provides an inadequate foundation for evaluating the partnership. Although the PD outlines overall objectives, outcomes and outputs, there is no clear logical model articulating the operational strategies for priority processes that will lead to those outputs and outcomes. There is clear accounting provided for the national and provincial government actions as well as those of other development partners. The PD (or a subsequent operational document) should have provided a more thorough consideration of the implementation context in Sindh province and the selected districts. This should have included a) a description of available health system resources, both public and private sector and b) identification of other development partners and interventions.

The listing of indicators in the PD is extensive, but little description is provided on how monitoring will be carried out. As a initial step in preparation for evaluation, the overall documentation for NPPI needs to be strengthened with a detailed implementation document. Such a document should:

- provide an understanding of the operational strategies that will be used in the context of the PI to attain the ambitious output and outcome level achievements & indicators included in PD;
- articulate the implementation arrangements between UNICEF, UNFPA and WHO, specifying which partner will be responsible for each program element and with a particular focus on implementation-level monitoring;
- provide an understanding of the relative magnitude and priority of individual components – for example, Output 1.1. (contracting out MNCH services to the private sector) appears to account for 50% of NPPI resources (based on this funding, can one assume that this is a priority activity and therefore worthy of close monitoring?)

¹⁹ Including: project design, results chain, indicators and success criteria, implementation arrangements, methods, approaches, types of activities, monitoring and evaluation systems, stakeholders, intended target groups and geographic coverage, and so on

- provide important benchmarks and the expected pace of NPPI implementation – for example, if districts are to be sequenced for the implementation of Output 1.1, how many districts should initiate the activity each year? What are the expectations in term of the number of years of implementation required to produce measureable changes in population-based measures (i.e. outcomes)?
- describe further how the NPPI activities were expected to complement and/or build on the existing programs of those partners;
- provide revised M&E targets using the baseline data that is now available.

Annual progress reporting could be strengthened by making more explicit linkages between individual activities from year-to-year – particularly in the case of delayed priority activities. For example, the set of activities around contracting out of MNCH services is clearly delayed (with actual implementation expected to start 3rd quarter 2010) but there is no explanation as to the obstacles encountered or lessons learned from this early phase of the start-up. In reviewing the 2010 work plan, it is not clear exactly how many districts have/will initiate contracting out through to the point of actual service delivery. In addition, UNICEF and UNFPA use different indicators for what appears to be the same set of activities (unless they are contracting for different services?)

Baseline surveys are described below. However, in regards to documentation, it is not clear how (or whether) the baseline data have been used to guide program strategies/materials and/or modify the monitoring and evaluation plans. For example, the facility assessments provide detailed information about the status of EmONC capacity within district facilities, but this information does not appear to have been used to create relevant NPPI performance measures²⁰.

Appraisal Area 2: Identify contextual factors likely to influence intended and unintended PI impacts. Then assess the availability, adequacy, and quality of contextual factor data/information currently available, going back at least 5 years in time where possible.

Team's Appraisal:

NPPI is being implemented against a backdrop of stagnating infant and under-five mortality in Pakistan. The PDHS (2006-07) found that trends in mortality reduction over the 1990s came to a halt in the early part of this decade (around 2003). The same is true for neonatal mortality. The factors underlying this unfortunate trend (e.g. unchanged levels of poverty, lack of progress in educating girls) may affect NPPI performance. Data on these trends and the underlying factors are widely available through PDHS surveys as well as the upcoming MICS4 to be conducted in Sindh province in 2010.

Resource allocation and spending in the health sector, particularly for the MNCH Programme, will also influence the performance of the NPPI. Issues have already

²⁰ In communication with the Team, the contractor for the baseline survey implied that the data had been destroyed; "We usually destroy all records related to the Project after receiving the final payment will check if the data of NPPI is still available and if we have then they will be immediately mailed to you. NPPI should take steps to ensure that all baseline data is secured to allow analyses with end line data. In the experience of the Team members, it would be highly unusual to delete baseline survey data in such a short timeframe.

emerged with funding for the national MNCH program with support from DFID discontinued following defaults in funding by GOP and concerns about fiduciary risk. In addition, the relationship of the NPPI efforts to contract out MNCH services with those already underway by the Government of Sindh and implemented through the SPRP/PPHI need to be articulated and mapped out.

Some of the activities included in NPPI may depend on central-level policy change, approvals and/or other forms of action. For example, NPPI activities to establish district level databases and strengthen M&E are on hold until related central level MNCH actions are completed. Such delays and sequencing of activities should be documented by the program for the purposes of any mid-term review and/or final evaluation. Additional contextual factors include effects of national-level policy modifications including Constitutional Amendments devolving responsibility for service delivery and national programmes to provinces and the outcome of the National Finance Commission awards that have increased allocations for health to provinces and reduced these at the federal level.

Finally, Pakistan has seen three major natural disasters (an earthquake in 2005, cyclone in 2007, and devastating flooding in 2010) which have damaged and stressed service delivery systems in many areas, including Sindh province (2010 flooding). As the implementing partners (UNICEF, UNFPA and WHO) have been closely involved in relief and recovery efforts, documentation on the impact of these events and progress in rebuilding is assumed to be available with them.

Appraisal Area 3: Review currently-available baseline data/studies and their appropriateness in terms of impact evaluation reliability and validity, keeping such questions as the following in mind:

- *What are the “right” impacts needing to be measured/verified?*
- *Are these “right” impacts actually verifiable to acceptable standards given the existing baseline information available?*
- *Where information is lacking or incomplete, can data be (re)constructed?*
- *What counterfactuals need to be considered?*

Team’s Appraisal:

Baseline data (household surveys and facility assessments for district headquarter hospitals, tehsil headquarter hospitals, basic health units and rural health units) appear to be comprehensive, and reasonably complete.

Two baseline facility assessments were undertaken in the 10 intervention districts with the purpose of providing baseline estimates of key indicators, as described below. The first covered Basic Health Units (BHUs) and Rural Health Centres (RHCs). A random sample of at least 20 BHUs and 5 RHCs were selected from the official list of functional facilities in each district. The second surveyed District Headquarters Hospitals (DHQHs) and Tehsil Headquarters Hospitals (THQHs). The report indicates that the official number of such facilities was much less than originally assumed, though no further explanation is provided. Because of the limited numbers, all such facilities were assessed, amounting to 5 DHQHs and 25 THQHs. The

documentation relating to methodology and fieldwork was limited, making assessment of survey quality problematic.

A baseline household survey covering the 10 intervention districts was conducted over the period October 2008 through January 2009. The methodology closely followed that adopted for the 2006-2007 Demographic and Health Survey (DHS), with the majority of survey questions taken verbatim from the DHS questionnaire. Female interviewers targeted ever-married women who had given birth over the previous three years. Across the 10 districts, 117 urban and 250 rural primary sampling units (PSU) were randomly sampled. The households in each PSU were listed and 12-16 selected using systematic sampling. This resulted in a total sample size of 5,400 women. The survey is reasonably well documented and appears to have been undertaken with appropriate levels of training and supervision. There was a considerable delay in gaining access to the data and we have therefore only been able to undertake a limited assessment of data quality. As will be seen in the indicator list below, in a few cases the existing analysis has not estimated the precise indicator required, even though this would have been possible.

One concern is that the survey questionnaire for women in the NPPI baseline includes some 600 items. As discussed in the case of India, attempting to collect data on such a large number of variables can prove a serious burden both for respondents and enumerators. In particular it allows very limited time to address issues arising when a respondent does not fully understand a question or is doubtful as to the appropriate response. We would suggest that, as with many surveys of this degree of complexity, this may have somewhat reduced the overall quality of the data, tending to increase problems of interpretation. To illustrate some potential issues, the frequency tables below focus on just two areas of particular relevance in terms of impact indicators – ANC/delivery and child vaccination.

The questions on ANC provision follow good practice in allowing multiple providers and requesting Yes/No responses on each. As can be seen, this allows us to note that there were very few non-responses to this question.

ANC providers accessed

	Yes	No	Missing Value
No one	1,885	3,515	0
Doctor	3,213	2,185	2
Nurse/Midwife/LHV	242	5,143	15
TBA	119	5,265	16
LHW	17	5,368	15
Homeopath	1	5,382	17
Hakim	0	5,385	15
Dispenser	24	5,357	19
Other	3	5,372	25

A surprising number of women reported making a large number of ANC visits, with more than 10% report at least 8 such visits. This raises questions as to their understanding the question and the extent to which enumerators tried to ensure that their understanding was correct.

Number of ANC visits

Number of visits	Frequency	%
1	575	16.5
2	869	25.0
3	658	18.9
4	431	12.4
5	241	6.9
6	206	5.9
7	146	4.2
8 or more	356	10.2

The TT vaccination table, with very few DK or missing values suggests that careful explanation of questions can result in at least a clear Yes/No response.

Tetanus Toxoid Vaccination

Received	Frequency	%
Yes	3,400	63.0
No	1,991	36.9
Don't Know	4	0.1
Missing	5	0.1
Total	5400	

The number of TT vaccinations also seems reasonable, with few women reporting receiving more than 3 and only 1 more than 5.

Number of Tetanus Toxoid Vaccinations

Number of Vaccinations	Frequency	%
1	256	7.6
2	2044	60.5
3	914	27.0
4	73	2.2
5	90	2.7
6	1	0.0

The delivery assistance question is multiple choice, allowing the respondent to identify more than one category of person. A probe is included to ensure that all adults present at the birth are included. The only issue with this format is that it would be helpful to clarify the roles played by different providers and duration of their involvement. It may be that such issues would be best determined at the time of enumeration, with the enumerator/supervisor identifying the key person involved.

Provider assisting delivery

Provider	Yes	No	Missing Value
Doctor	1,998	3,358	44
Nurse/Midwife/LHV	930	4,405	65
DAI/TBA	2,907	2,442	51
LHW	52	5,270	78
Relative/friend	1,777	3,557	66
Other	30	5,272	98
No one	22	5,378	0

Of the 5,400 mothers included in the survey, only some 10% were able to produce a vaccination card. In general the data from these cards seems reasonably complete, though a substantial proportion of entries for DPT3 and all HBV vaccinations are reported as undated or unrecorded.

Vaccination Card Entries

	Dated Entry	Missing Value	Undated	Not Recorded
BCG	467	13	1	3
Polio 0	363	13	11	26
Polio 1	405	13	3	17
Polio 2	352	13	5	29
Polio 3	283	13	10	47
DPT 1	367	13	10	44
DPT 2	312	13	10	50
DPT 3	252	13	15	68
HBV 1	297	13	24	76
HBV 2	249	13	25	84
HBV 3	210	13	26	93
Measles	204	13	6	42

One particular issue with reported vaccinations is that the questionnaire specifically requests the enumerator to check that a child receiving polio vaccine via a national campaign should also be recorded as receiving polio vaccine from any source. This

was clearly not pursued as row three in the table (polio campaigns) indicates a substantially higher number than row 3 (all sources). This failure by enumerators to follow a specific instruction intended to maintain quality is of concern.

Vaccinations reported by mother

Vaccination	Yes	No	DK	Missing
BCG	4,122	702	24	552
Polio	4,825	10	5	560
Polio Campaigns	5,269	52		79
DPT	3,577	1,244	50	529
HBV	3,392	1,354	97	557
Measles	2,800	1,892	154	554

Whereas only a very small proportion of children are reported as having received more than 3 DPT or HBV vaccinations, the great majority are reported to have received 7 or more polio vaccinations. One possible explanation is that children are receiving polio drop repeatedly during national campaigns but it would be useful to clarify this point.

Number of Vaccinations Reported by Mother

Polio	Frequency	%	DPT	Frequency	%	HBV	Frequency	%
1	49	1.0	1	450	12.6	1	399	11.8
2	57	1.2	2	524	14.7	2	535	15.8
3	91	1.9	3	2505	70.2	3	2368	69.9
4	83	1.7	4	13	0.4	4	9	0.3
5	96	2.0	5	0	0.0	5	2	0.1
6	67	1.4	6	2	0.1	6	0	0.0
7 or more	4,380	90.8	7 or more	72	2.0	7 or more	76	2.2

In all cases, it would be preferable to include English-language versions of the survey tools as appendix materials to the reports. In addition, the baseline household survey report would benefit from the use of simple data quality assessment also as appendix material²¹.

The PD includes an extensive set of impact indicators (10 in all). The Team raises three primary concerns with these indicators:

a) Impact measures of this type are best considered to be achieved through joint action and it is typically not possible to reliably assess the contribution of any single donor. The national MNCH Programme and Sindh province have an extensive set of development partners including multilaterals, bilaterals and NGOs. While the NPPI

21 UNICEF MICS surveys provide a complete set of materials for use in creating such data quality assessments including syntax files and model tables. See: http://www.childinfo.org/mics3_modelreports.html

programme could plausibly claim to have made a contribution to any improvement in impact level measures for Sindh, there would be little point in attempting to quantify this contribution. The PD should have clarified this at the outset. Requisite documentation for any impact evaluation would include an institutional mapping to depict the multiple actors, funding sources and parallel, similar initiatives underway in Sindh province as well as the 10 selected districts.

b) NPPI contributions to impact would be most directly observable in the 10 selected districts in which it works, covering some 29% of the provincial population. However, many of the proposed impact measures are not measurable at the level of the districts (all mortality indicators) and were not included in the baseline HH survey.

c) The potential contribution of NPPI towards several of the impact-level indicators is unclear. For example, from the available materials it is not possible to draw a plausible association between NPPI inputs and changes in the prevalence of diarrhea, ARI, iron deficiency anaemia prevalence or changes in nutritional status.

An NPPI evaluation will benefit from the fact that there has been a robust set of household surveys conducted in Pakistan over the past decade. These include the Pakistan Family Planning and Reproductive Health Survey (2001-02), a MICS-type survey (2004)²², a PDHS (2006-07), as well as an anticipated MICS4 survey in Sindh province (2010) and PDHS (2011-2012). These surveys coupled with the HH baseline survey will allow for considerable secondary analysis to examine trends in key impact indicators and explore the contribution of the NPPI.

NPPI's overall monitoring appears to fall seriously short in the area of implementation-level monitoring required to link resource allocation to outputs and outcomes. In the PD, implementation-level measures (which come under output areas) are often ill-defined (numerators and denominators) and insufficiently linked to actual program content. The PD contains a description of available data sets and proposed M&E activities (pages 27-34) but it is not clear how these have been made operational. The utility of annual SAVVY exercises to estimate cause of death and burden of disease is questionable.

Appraisal Area 4: Summarize important findings and conclusions related to each of the above tasks.

Team's Appraisal

Recommendations:

1. As drawn from the PD, NPPI's impact indicators are overly-ambitious in their scope and targets. NPPI is advised to clarify that it is a contributor to the achievement of a smaller set of impact indicators which are measurable at the province level. NPPI should not seek to claim direct attribution of impact -- which is achievable only through joint efforts of the Government of Pakistan and development partners.

²² Reports and data sets from the Pakistan FP and RH Survey (2001-02) and MICS-type survey (2004) are not readily available and not reviewed by the Team.

2. For monitoring and evaluation of NPPI, number of developments/advances have occurred since the PD was finalized in August 2008, including:

- extensive baseline survey data is now available from both households and facilities;
- more clarity (via ToRs, etc.) is available on the package of MNCH services to be contracted out;
- working with development partners, the national MNCH programme is working to review and strengthen its overall M&E;
- other PPP initiatives such as PHHI have come online in several of the districts;
- assumptions about overall support for the national MNCH program have not held.

Given these developments, the Team recommends that NPPI fully elaborates a Monitoring, Evaluation and Analysis Plan. This new Plan should encompass:

- a detailed logical model depicted the intended and actual pathways of NPPI support (through inputs, process, outputs and outcomes) along with an analysis plan demonstrating how available and planned data can be used to make the strongest possible associations between activities, outputs and outcomes;
- reduction in the number of indicators to a core set closely and clearly associated with priority NPPI interventions and program content (Table 2 below provides a suggested reduction in the PD indicators);
- greater clarity on the nature of the performance indicators including specification of numerators and denominators and level of measurement (e.g. are indicators being assessed by individual districts or together as a weighted average?);
- revision of targets taking into account baseline values;
- identification of programmatic benchmarks and expected pace of implementation (e.g. at what point can a district be considered to have a fully operational contracting-out MNCH service?; what volume/reach of contracted out service would be required to make a change in population-based outcome measures?; what is the acceptable standard of contracted-out MNCH services? what percentage of contracted-out facilities meet quality standards?);
- alignment of NPPI indicators with those of the national MNCH program;
- identification of “exposure” variables for use in the end line survey – to allow analyses by associating outcomes with use of NPPI-supported services – particularly the contracted out MNCH services.

3) The Team has identified three main options for moving forward with impact evaluation of the NPPI as follows:

The gold standard approach would involve a pre-post/case-control evaluation design articulated at the inception of the program. This approach would entail assigning project and control districts matched on key socio-economic characteristics and health system variables. In advance of implementation, the evaluators, with the support of province and district program managers, would attempt to limit and control other inputs (i.e. other development partner inputs above and beyond the routine MNCH service delivery provided through GoP) in both case and control districts. By limiting the non-NPPI inputs, the evaluation design would, supposedly,

allow for greater isolation of any NPPI “effect”. The sampling design for baseline and end line data collection would seek to measure statistically significant changes both over time and between project and control districts. The design would include analyses of “exposure” variables which would seek to determine the magnitude of the effect (outcomes) based on the intensity of NPPI-supported service use either at the individual or district level. Finally, the gold standard approach would include careful and systematic monitoring of implementation using indicators that had been developed, validated and used consistently throughout the project. In practice, the Team is doubtful that this gold standard approach could work in the NPPI because of the numerous confounding variables which render the isolation of any NPPI “effect” infeasible.

A global health –normative option that would satisfy internationally-accepted standards might entail evaluation only in the 10 project districts without the use of control districts. In those 10 project districts, evaluators would fully account for the inputs of other development partners -- but not try to limit or control those inputs. Sample parameters would be designed to either a) determine with a known confidence whether pre-determined targets had been achieved, or b) measure statistically significant change over time (pre/post) in key variables (the preferred option). As with the “gold standard”, this design would also analyze the “exposure to” NPPI-supported MNCH providers/services and associate those variables with the desired outcomes. Moreover, this design would maximize NPPI’s sequential introduction across districts (as activities are not initiated in all ten districts simultaneously) as a form of “natural experiment” and incorporate measures of program duration and intensity of implementation as variables in the analyses. Finally, the bare minimum approach should also include careful and systematic monitoring of implementation using valid indicators used consistently throughout the project. The Team proposes that efforts to a) contract out MNCH services to the private sector; and b) generate demand through the use of vouchers/incentives are subject to a bare-minimum option of evaluation. These types of interventions are currently subject to much international support and attention for evaluation. NPPI’s OR approach for these two components should be shared with other partners working in this arena as a means of strengthening and validating the methods employed.

The minimal option would seek to triangulate available data sources (PDHS 2006-07; NPPI baseline and end line surveys; MICS4 2010 in Sindh province; PDHS planned for 2011-12) to either a) determine with known certainty whether NPPI targets were achieved (y/n) or b) measure statistically significant change in outcome measures over time. Triangulation of data sources – including objective, systematic implementation monitoring at multiple levels (district, tensil, facility, provider) -- could help to identify associations using timelines (of implementation) and trends in desired outputs and outcomes²³. The Team believes that this option only be used if, through the development M&E plan, it is apparent that the global health-normative option is not possible.

23 For an example of this approach see: System-level determinants of immunization coverage disparities among health districts in Burkina Faso: a multiple case study Slim Haddad, Abel Bicaba, Marta Feletto, Elie Taminy, Moussa Kabore, Boubacar Ouédraogo, Gisèle Contreras, Renée Larocque, Pierre Fournier. BMC International Health and Human Rights 2009, 9(Suppl 1). Published: 14 October 2009.

Documents reviewed:

1. NPPI Project Document. Reducing maternal, neonatal & child mortality in Sindh. August 25 2008. Delivering as One. United Nations in Pakistan. Ministry of Health/Government of Pakistan. Norway-Pakistan Partnership Initiative.
2. Memorandum of Understanding between The Government of the Kingdom of Norway and The Islamic Republic of Pakistan regarding The Norway-Pakistan Partnership Initiative (NPPI).
3. Result Based Financial Mechanisms for Improving Maternal, Newborn and Child Health Outputs: A Feasibility Study for 10 Selected Districts of Rural Sindh. Research Report Commissioned by MNCH Program, Sindh and UNICEF. Funded by Norwegian Pakistan Partnership Initiative [NPPI]. Prepared by Aga Khan University, Karachi March 2009.
4. Standard Administrative Arrangement between the Kingdom of Norway, the United Nations Resident Coordinator and the United Nations Development Programme.
5. NPPI Baseline Survey in Sindh. Report 1: Baseline Household Survey. Submitted to UNICEF Sindh. Prepared by: Tauseef Ahmed PhD. 30 March 2009.
6. GAVI Alliance Annual Progress Report 2009. Submitted by the Government of Pakistan. 15 May 2010.
7. <http://www.mnch.gov.pk/fl.php>. Accessed 29 November 2010.
8. <http://www.sindhhealth.gov.pk/history.htm>. Accessed 29 November 2010.
9. National Institute of Population Studies (NIPS) [Pakistan], and Macro International Inc. 2008. Pakistan Demographic and Health Survey 2006-07. Islamabad, Pakistan: National Institute of Population Studies and Macro International Inc.
10. GAVI Alliance. GAVI Alliance Health System Strengthening (HSS) Applications. Pakistan. May 2007.
11. National Maternal Newborn and Child Health (MNCH) Program. 2006 – 2012. Government of Pakistan. Ministry of Health. November 2006.
12. Norway Pakistan Partnership Initiative (NPPI). Core Action Plan April-December 2009.
13. Norway Pakistan Partnership Initiative (NPPI). Progress Update (May 2010) WHO-UNICEF-UNFPA Activities.
14. Norway Pakistan Partnership Initiative (NPPI). Core Action Plan January -December 2010.
15. <http://www.thardeep.org>. Accessed 29 November 2010.
16. United Nations Children's Fund. Executive Board. Annual session 2008. 3-5 June 2008. Short-duration country programme document. Islamic Republic of Pakistan. E/ICEF/2008/P/L.11
17. Executive Board of UNFPA and UNDP. Annual session 2008. 16 to 27 June 2008, Geneva. Item 4 of the provisional agenda. UNFPA - Country programmes and related matters. United Nations Population Fund. Extensions of country programmes in the Asia and the Pacific region. Note by the Executive Director.
18. Sindh Development Review 2008-09. People's Primary HealthCare Initiative (PHHI). Planning and Development Department. Government of Sindh. www.sindhpn.gov.pk.
19. Ministry of Health (MoH) and Technical Resource Facility (TRF). Draft Aide Memoire. Second Mini Review of the National Maternal, Newborn and Child Health (MNCH) Programme 12-14 April, 2010.

Appendix H: Norway - Malawi Partnership Initiative Profile

PI duration:	Start date December 2010/January 2011	End date 2013
PI purpose		
Stated objective	The objective is to contribute to reduction in maternal and neonatal mortality (MDG 4 & 5) using the instrument of RBF through the introduction of sustainable demand and supply side mechanisms. Detailed sub-objectives for each component are to be worked out during the initial design phase.	
Funding:	Total: Germany and Norway have pledged a sum of USD 10 million over a three year pilot period, where we will contribute USD 5 mill equivalent each over the programme period. Norway has set aside NOK 30 million for the three year period 2011 – 2013.	
	By partner:	By activity area (funding by area not yet available):
	KfW will be “lead donor” with a contract /agreement between MFA and KfW (being finalised). KfW will be the main agreement signatory together with the Ministry of Finance, Malawi. It is not yet decided if MFA/The Embassy in Lilongwe will have a separate agreement with the Government of Malawi	The Joint KfW/Norway pilot RBF programme will consist of the components: a) Detailed program design will be worked out during the first phase of the program, including the selection of the participating facilities b) Infrastructure improvements - prior to the introduction of RBF interventions to ensure a minimum standard of eligible facilities with likely focus on improvements of existing waiting homes, guardian shelters and maternities; other infrastructure; equipment. c) Component 1 facilitates women’s access to facilities for delivery by reimbursing the cost of transport for her and her guardian as well as incentives for not delaying arrival at the facility, and staying for the required time. d) Component 2 provides quality and performance contracts with public and the Christian Health Association of Malawi (CHAM) health facilities to improve and sustain quality and volume of maternity services

PI duration:	Start date December 2010/January 2011	End date 2013
PI within landscape for MDGs 4&5 / development	According to National Health Accounts data, in 2004/05, total spending for maternal health programs was estimated at US \$ 15.3 million and child health spending reached US\$ 35.85. Half of these funds were provided by donor agencies and the remainder were provided by public and private sectors.	
Geographic Areas and other partners working there	Area 1: The pilot area encompasses five districts in Central Eastern Zone of Malawi (Kasungu, Nkhotakota, Ntchisi, Dowa and Salima). A control area is to be determined during the design phase -- likely to be in a different zone. There will be no NMPI interventions in this zone other than collection of baseline data.	
Criteria for selection	Discussions during the Feasibility Study helped to define the pilot area with the decision to select districts within a single zone to facilitate monitoring. It was agreed that the programme would be implemented in Central Eastern Zone. This was not because this area had higher need but because it was nearer to the Capital from which the project would be administered to act as a learning site. Further, that the health centres in these areas are better staffed and equipped (than many other rural areas)so the programme would be able to finance as little as possible in terms of ORT and also only minor infrastructures were needed (upgrading of staff housing, waiting shelters, clinic etc).	
Partners description	Partners	Responsibilities
	Kfw	Lead donor
	RHU/MOH	day to day management of the program including recruitment of consultants through a tender process
	Consulting firm (TBN)	to support program management and provide technical support to RHU/MoH, support management of funds as well as reporting
Institutional arrangements:	Norway will channel funds through KfW to the programme. The programme will be managed by staff in the MOH. Agreements drafted include: bilateral agreement with Malawi; bilateral contract with KfW, bilateral agreement between KfW and Malawi on delegated partnership. Both Norway and Germany are signatories for the health SWAp – the program will be implemented as a discrete funding item in the SWAp.	

PI duration:	Start date December 2010/January 2011	End date 2013
Implementation timeline	Planned Program start towards the end of 2010 with tender process for consulting firm during July/August 2010. The pilot to take 2-3 years with phased implementation of different components.	Actual
Elements of Innovation (as described in PD)	The pilot, focused on results-based financing, represents an innovative financing approach.	
Operations Research activities	Activity	Status
	No details yet provided.	

Appendix I: Norway - Malawi Partnership Initiative Monitoring and Evaluation Plan

Table 1: Norway- Malawi PI Monitoring and Evaluation Plan

This table shows a compilation of indicators drawn from available program documentation – it does *not* represent the Team's recommendations on appropriate indicators for evaluating PI performance or impact.

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
Reduced direct maternal mortality rate	1					
Reduced direct maternal mortality rate	1					
<i>Visit indicators</i>						
# 1st Antenatal visits	1					
# Women with 2 Antenatal visits	1					
# 1st time family planning visit	1					
# Infants with post natal check up	1					
# Infants given vaccine X	1					
<i>Contents of care indicators</i>						
% Women that benefit from transport+** during last delivery	1					
# Women who use shelters before delivery						
# Facility based deliveries	1					

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
% Women given 2 doses of IPTp	1					
# Caesarean sections	1					
# Obstetric referrals	1					
% Facilities that meet set Infection control standards.	1					
Delivery to skilled nurse / midwife ratio per facility (?).	1					
Health worker satisfaction	1					
Health worker retention rate	1					
<i>Facility level indicators (linked to RBF)</i>						
% Deliveries from 1st ANC till Deliveries (Formula of declining ratio)*	2					*
% C-Sections	2					*
% Complicated and emergency cases referred to other places	2					*
% Pregnant woman fully immunised against tetanus	2					*
# Facility based maternal deaths	2					*
% Cases treated for Septicaemia	2					*
% Neo-natal deaths	2					*

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
% Post natal follow-up	2					*
% Complete HMIS Reporting	2					*
% Fully immunised infants	2					*
Frequency of drug stock-outs	2					*
% Actual deliveries in relation to the agreed number in SLA	2					*
Aggregated proportion of improvements made on the above indicators from the previous year	2					*

Appendix J:

Norway - Malawi Partnership Initiative Evaluation Options Appraisal

Appraisal Area 1: Map the basis for Norway's support to the PI using project documents²⁴ and assess the adequacy and quality of such documentation in terms of potential "impact evaluability" of the PI towards the end of the initial implementation period (typically 5-6 years).

Team's Appraisal:

The available documentation provides an outline of the types of activities to be undertaken and implementation arrangements. In its current form, the available documentation would not suffice for impact evaluation purposes. More detailed description of the actual interventions and implementation seems slated for the first phase of the program. That phase should also develop a more detailed logical model for the pilot intervention which should depict how the elements of the program are envisioned to work together towards the intended results. For example, the existing material does not make clear the duration and/or geographic targeting of planned infrastructure improvements. Presumably, this phase must come to completion before components 1 and/or 2 are put into effect.

Additional documentation should be provided on the selection of pilot areas. While mention was made of discussion during the Feasibility Study, there are no clear criteria or process described for the selection of these sites. In addition, given the sharp increase in per capita development assistance for health in Malawi, it would be extremely difficult to identify comparison districts which are not benefiting from substantial external assistance. These factors and decisions should all be carefully described in the basic documentation developed during early implementation.

As the PI seeks to provide a model for scaling-up, consistent documentation of decision-making throughout the pilot will be a critical element of the evidence base. As a potential model for wider application, the pilot should carefully record the costs associated with the intervention at all stages.

²⁴ Including: project design, results chain, indicators and success criteria, implementation arrangements, methods, approaches, types of activities, monitoring and evaluation systems, stakeholders, intended target groups and geographic coverage, and so on

Appraisal Area 2: Identify contextual factors likely to influence intended and unintended PI impacts. Then assess the availability, adequacy, and quality of contextual factor data/information currently available, going back at least 5 years in time where possible.

Team's Appraisal:

As mentioned above, Malawi is a country which has received enormous support from development partners for the health sector – complicating the job of evaluators to find comparison areas. Particularly in a small country, the experiences and lessons of other actors/districts in improving maternal health and obstetric care may permeate PI intervention districts. Moreover, other innovative schemes may be implemented in so-called control areas.

Appraisal Area 3: Review currently-available baseline data/studies and their appropriateness in terms of impact evaluation reliability and validity, keeping such questions as the following in mind:

- What are the “right” impacts needing to be measured/verified?
- Are these “right” impacts actually verifiable to acceptable standards given the existing baseline information available?
- Where information is lacking or incomplete, can data be (re)constructed?
- What counterfactuals need to be considered?

Team's Appraisal:

The material reviewed is still in a preliminary state making it difficult to reach conclusions on the above questions. As an observation, it seems that much of the M&E outlined seems, in principle, to be targeted on the correct variables and is at the correct level. For example, inputs and processes are described as including infrastructure such as waiting/guardian shelters, equipment, communication, health work force and supply chain of essential drugs while outputs include facility readiness and capacities, intervention utilization, quality and safety. Outcomes are identified as increased intervention coverage and reduced prevalence of risk behaviour (e.g. delays in care seeking).

These reasonably-stated process, outputs and outcomes state in contrast to other portions of the reviewed documentation where, for example, the primary indicators of interest include improved health outcomes (e.g. reduced maternal and neonatal mortalities). Measurement of reduced maternal and neonatal mortality is simply not feasible in a pilot of this duration and scale. The program will also not be able to measure meaningful change in indicators of relatively rare events such as the number of facility based maternal deaths or percent of cases treated for septicaemia. The PI is better positioned to focus on changes in increased use, improved quality and acceptability of the priority interventions.

Designers of the M&E component might want to consider using comparison areas from within the same districts rather than attempting comparison districts. It seems reasonable to assume that these pilot activities will be introduced in a phased manner. In addition, the pilot is largely focused on facilities providing delivery services and their catchment areas. It may well be possible to design an evaluation focused on experience of facilities and catchment areas as they are phased into

the pilot program. The design is similar to that proposed (pre-/post-intervention with intervention and comparison districts) although carried out on a smaller, more “do-able” scale.

Appraisal Area 4: Summarize important findings and conclusions related to each of the above tasks.

Team’s Appraisal

Recommendations:

According to the PI documents, a monitoring and evaluation framework will be defined during initial phase of RBF program design. The NMPI is well-positioned to conduct a **global health – normative** option with important design elements in place at its inception. According to the PI documents, a monitoring and evaluation framework will be defined during initial phase of the results-based financing (RBF) program design (starting January 2011). The Team would encourage the designers to take careful note of existing international experiences and approaches for evaluation of results-based financing – notably the required timeframe and required resources. As part of the initial monitoring and evaluation design, a strategy and plan for documentation should be developed and contracted. Documentation becomes particularly important as the study areas were selected based on their accessibility to the capitol and relatively better infrastructure. These circumstances could be seen as compromising the programmes’ replicability. As a learning site with better than average conditions, the PI partners should assiduously document the conditions required for successful performance (e.g. service accessibility and quality) with an eye towards how those conditions will be made available in other areas and program effect replicated.

The available documents cite the main goal of the M&E component as “improving the availability, quality and use of the data needed to inform results-based financing program reviews and planning processes, monitor health outcomes progress, health system performance and ultimately demonstrate impact of the project on selected outcome indicators with reasonable degree of certainty”. This goal is far too ambitious for a pilot project expected to be evaluated after 2 to 3 years -- likewise the expectation that a baseline, mid-term and end of project evaluation will be carried out.

Program designers are encouraged to moderate their expectation for the M&E component to a more achievable set of aims. While the concept note drafted for the evaluation is a good start (even the basic depiction below of indicator domains and data collection is notably absent in other PIs reviewed), its next version should reflect the short timeframe and add greater specificity on what existing information will be utilized in the evaluation (potential data sources mentioned project and health facility documents, clinical reporting systems and HMIS). The Team encourages the PI to rely on existing sources such as the HMIS and the instrument jointly developed by the MoH and JHPIEGO which will be utilized throughout the country (independent of the PI pilot) and will thus serve as an important measurement of maternal and newborn health services.

Finally, the program designers are strongly encouraged to work closely with the RBF for Health impact evaluation network (part of a Norwegian funded initiative based at the World Bank to support results-based financing innovation in eight countries). There is a readily available set of tools and materials to assist with the further development and refinement of the evaluation approach.

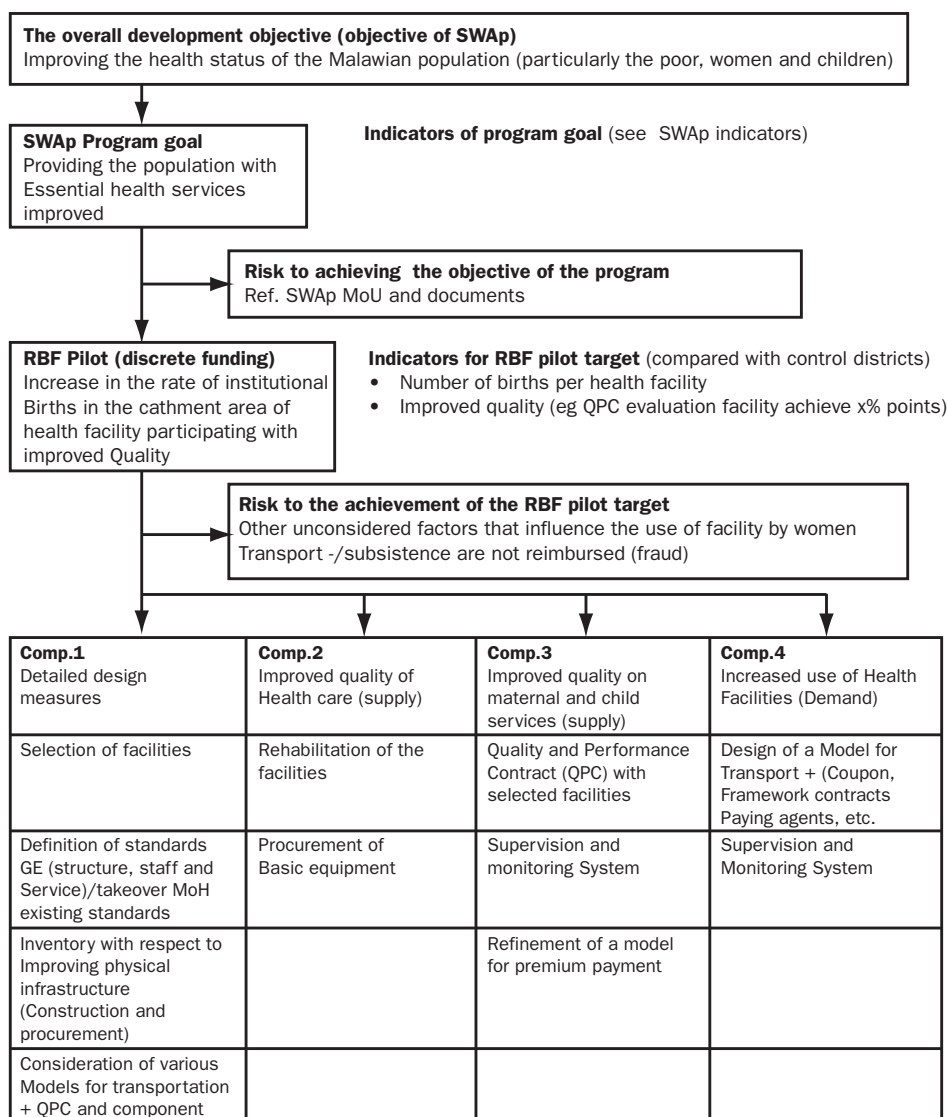
Documents reviewed:

1. Malawi RBF. Pilot on Results-based Financing (RBF) for Maternal and Newborn Health in Malawi – September 2010. Norad.
2. Potential Result-based Financing for reduced Maternal and Child Mortality in Malawi - Background document. Norad.
3. Norwegian-German Initiative to support MDG’s 4 and 5. Feasibility Study on an MCH programme in Malawi using the instrument of RBF. Phase 1. FINAL Draft Report. December 2009. David Griffith, Brigitte Jordan-Harder, Alice Maida.
4. Results-based Financing (RBF) for Maternal and Newborn Health in Malawi. Minutes of Mission. Lilongwe, 21 June 2010
5. A concept note for the evaluation of result based funding mechanism of maternal and child health services in Malawi. Prepared by Jobiba Chinkhumba MBBS MSc., Department of Community Health College of Medicine.
6. Ministry of Health. March 2007 Malawi National Health Accounts (NHA). 2002-2004 with Subaccounts for HIV and AIDS, Reproductive and Child Health, Department of Health Planning & Policy Development, Lilongwe, Malawi. Abt
7. Results-based financing for health. Health results-based financing impact Evaluation network. World Bank.
8. Decision Document – Project and Programme Support. Royal Norwegian Embassy Lilongwe.

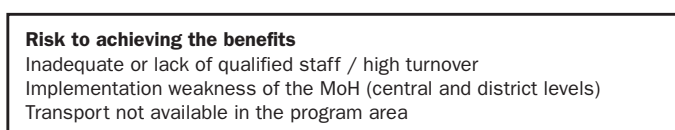
Figure is drawn from: *A concept note for the evaluation of result based funding mechanism of maternal and child health services in Malawi.* Prepared by Jobiba Chinkhumba MBBS MSc., Department of Community Health College of Medicine.

	Inputs & processes			Outputs	Outcomes	Impact
<i>Indicator domains</i>	Transport +	Quality and Performance contract.	Equipment and other supplies	Intervention accessibility Intervention quality Intervention safety Facility readiness	Increased intervention coverage Increased intervention utilization Reduced prevalence of risk behaviour e.g. delayed care seeking. Improved staff working conditions.	<i>Improved health outcomes</i> Reduced maternal and neonatal mortality
<i>Data collection</i>	Project documents Facility records		Facility Assessments	Community based surveys		
<i>Analysis</i>	Health Information Management Systems					
	Project progress and performance & systems capacities					

Log Frame: Pilot on Results-based Financing (RBF) for Maternal and Newborn Health in Malawi



Activities that involve multiple components:



Appendix K: Norway - India Partnership Initiative Profile

Overview	The RNE is investing NOK 500 million in support of the National Rural Health Mission through implementation multiple partners including UNOPS, UNICEF and WHO.	
PI duration:	Start date August 1st 2006	End date Original end date: 31 March 2012 Extension end date: 2013
PI purpose:	<p>The aim of the NIPI partnership is to facilitate rapid scale-up of quality child related health services that are equitable and sustainable with focus on:</p> <p>a) Strengthening Gols National Rural Health Mission initiative by supporting an independently managed enabling network facilitating delivery of MDG 4 related services.</p> <p>b) Testing and introducing new ways of scaling up services by community health workers - Accredited Social Health Activists (ASHAs) at the village level in focus states.</p> <p>c) Engaging the private sector in the delivery of MDG related services at all levels.</p> <p>d) Exploring new opportunities as they arise and conducting operational research to establish their value.</p> <p>NIPI further aims to provide flexible support to enable implementation and innovation and to resolve bottlenecks.</p>	
Stated objective:	<p>These are the three objectives and expected outcomes according to PD:</p> <ul style="list-style-type: none"> • Save an additional half a million under five children each year from 2009 onwards. • Sustain routine immunisation coverage rate in the country at 80% or more from 2007 onwards. • Contribute to improved performance of the health system as a whole and development of best procedures for large scale roll-out of interventions addressing MDG 4 also in other countries. 	

Funding:	<p>Total: Funding of NOK 500 million (US \$ 81.1 million) over approximately 5 years to identified agencies in line with individual agreements/contracts. No funds received directly by the Government of India.</p>	
	By partner:	By activity area:
	<p>UNICEF: MFA grant of NOK 130 million over the period: 1 August 2006 to 31 March 2012; a fifth state, Uttar Pradesh is also supported by NIPI through UNICEF</p>	<p><i>Focal Area A: Quality Services for Child Health:</i> Catalytic interventions related to: universal immunization (cold chain and vaccine management systems); newborn and child health interventions; related maternal health intervention; the Yashoda/Mamta initiatives. This area accounts for the largest portion of the NIPI budget lead by items including the yashoda/mamta; home-based newborn care and sick newborn care units.</p>
	<p>UNOPS LFA: MFA grant of NOK 240 million over the period: 23 November 2006 to 31 March 2012; of which:</p> <p>Rajasthan – USD 3,620,650 (14-12-07) Bihar - USD 3,264, 098 (27-12-07) MP - USD 3, 563, 614 (20-12-07) Orissa –USD 3, 462, 517 (13-12-07)</p>	<p><i>Focal Area B: Enabling Mechanisms</i> Catalytic interventions relating to techno-managerial support that contributes towards enhancing the overall quality and effectiveness of the programme and strengthening of health systems. These include: strengthening state/district and block management structures for child health; catalytic action to galvanize and motivate teams and support training activities; gap management and problem-solving related to technical solutions, system bottlenecks, planning, budgeting management and financial issues. Innovative solutions, operational research, identification of best practices and refinement of approaches.</p>
	<p>WHO: MFA grant of NOK 65 million over the period: 1 December 2006 to 31 December 2011</p>	<p><i>Focal Area C: Learning and Sharing of Experiences -</i> Research, pilot projects and models are considered inherent in Focal Areas A and B. Private sector involvement, capitalizing on new opportunities, pro-poor focus and gender are of particular importance.</p>
	<p>NIPI Secretariat – NOK 50 million</p>	
	<p>RNE – NOK 15 million</p>	
PI within landscape for MDGs 4&5 / other development partners	<p>The majority of funding for the NRHM comes from the Gol. The four states participating in NIPI are high focus states for the NRHM and allocated (central and state gov't) US\$ 1.4 million (Orissa) and US \$ 2.8 million (Bihar) per fiscal year for the NRHM. Other development partners, such as the World Bank, play a significant role in these states as well.</p>	

Geographic Areas (including but not limited to)	Area 1: Bihar	UNICEF, UNFPA, WHO-NPSP, Dfid
	Area 2: Madhya Pradesh	USAID, Dfid
	Area 3: Rajasthan	USAID, EC
	Area 4: Orissa	USAID, Dfid
Criteria for selection	These four states are described as highly populated, poor and majority contributors to India's mortality statistics. They are among the 10 states identified as "high focus" for the NRHM. No further criteria or ranking was found. UNOPS works in selected districts of these four states (3 districts per state). As per the Baseline Household Survey Report: the districts within each state were selected on the basis of various health indicators – "neither good nor poor in health indicators meaning that districts were selected having moderate health indicators". WHO and UNICEF work in other districts and in other states. There is no sub-set of districts where all three operate with RNE funds.	
Partners description	Partners	Responsibilities
	UNOPS / LFA	<p>Local Fund Agent support to NIPI States (State health Societies being the real implementers, with UNOPS having the responsibility of LFA and of facilitating and ensuring the implementation of these activities in an efficient and effective manner). Techno-managerial support to Program Management Units (PMUs) under NRHM</p> <p>Yashoda/Mamta (newborn aides)</p> <ul style="list-style-type: none"> • Birthing kits • Home-based neonatal care (HBNC) • Sick newborn care units (SNCU) intended to improve health within the district and also to serve as model for the state • National Child Health Resource Centre
	UNICEF	<ul style="list-style-type: none"> • Integrated management of neonatal and childhood illnesses (IMNCI) • Improved facility based newborn and child care across NIPI focus States Establishment including Sick New Born Care Units • Accelerate immunization coverage through strengthened cold chain and vaccine management systems • District and Block planning and management support across NIPI focus States • Facilitate community based newborn and child care interventions across NIPI focus States through planning, capacity building, supervision and monitoring • Develop and promote innovations for child health service delivery, prototyping new models and research activities

	WHO	<ul style="list-style-type: none"> • Immunisation/ control of vaccine-preventable diseases (measles) • Pre-service IMNCl training to medical and health professionals • Emergency obstetric care (EmOC) and skilled birth attendance • Curriculum development and assessment of training schools for ANM/ nurse training • Establishment of an accreditation system for facilities in MNCH • Study on malnutrition
Institutional arrangements:	RNE signed MOUs with UNICEF, WHO and UNOPS. A NIPI Secretariat was created to serve as overall program management and coordinator.	
Implementation timeline	Planned	Actual
	Overall agreement between Norway and India signed in mid-2006 with end date of 31 March 2012	Agreements with SHS signed December 2007, program implementation begins, effectively, in 2008. JSC has agreed to no-cost extension to 2013. Data prepared for the MTR shows utilization of allocated funds ranged from 4 – 11% by state.
Elements of Innovation (as described in MTR)	It was originally envisaged that the partnership would work on a flexible basis, providing up-front catalytic support where needed. While such catalytic support has, perhaps, been provided, lack of documentation makes it difficult to pinpoint.	
Operations Research activities	Activity	Status
	WHO: A study to analyze the socio-cultural determinants of childhood malnutrition and its management in the community and health facilities (in collaboration with INCLEN). The data collection is complete and analysis is underway.	Underway
	WHO: A study comparing the effect of commercial energy dense food and iso-calorie locally prepared food on weight gain and body composition in children with Severe Acute Malnutrition (in collaboration with ICMR-National Institute of Nutrition).	Underway

<p>WHO: Capacity building of Medical Officers and Nurses in management of Severe Acute Malnutrition (SAM): The training modules developed and ready for piloting.</p>	<p>Underway</p>
<p>Public Health Foundation of India (PHFI) and Centre for Development and the Environment (SUM), University of Oslo: 'Assessing and supporting NIPI Interventions' in Norway India Partnership Initiative.'</p> <p>Fafo Institute of Applied Studies, Oslo and ANSWERS: 'Assessment of Pivotal Issues related to infant feeding and child nutrition: Inputs for Improving Interventions in Norway India Partnership Initiative.'</p>	<p>Underway</p>

Appendix L: Norway - India Partnership Initiative Monitoring & Evaluation Plan

This table shows a compilation of indicators drawn from available program documentation – it does *not* represent the Team’s recommendations on appropriate indicators for evaluating PI performance or impact.

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
IMR	1			8, 9	8, 9	
U5MR	1			8, 9, 11	8, 9, 11	
% Children fully immunised	1		80%	1, 8, 10	2, 9, 10	
Child Health Resource Network: Appropriate physical infrastructure for District Nodes	1		District Nodes created			
% of cold chain equipment functional at Government Medical Store Depot (MSD), state and regional levels	1		95%			
Vaccine logistics and cold chain MIS	1		Implemented in 4 MSDs			
National/state level IMNCI coordination groups	1		Functioning			
Detailed state and district IMNCI plans	1		Prepared for medium/long roll out and integrated in PIPs of the 5 states			

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
Advanced IMNCI implementation	1		>40% training completed in 100 districts by 2009 All completed by 2011			
% IMNCI trained workers who correctly assess, identify, classify and manage sick under-five children	1		70%			
Number of partners for delivery of essential child services through PPP	1		Increased, particularly in hard-to-reach areas and urban slums			
New evidence and best practices documented for advocacy and disseminated for improved implementation of National Child Health Strategy	1		At least 4 specified studies			
Survival rate of in-station admitted neonates in all public SNCU equipped FRUs	1		>80%			
Zinc treatment guidelines availability	1		All AWCs, SCs & PHCs across country			
NIPi Secretariat	2		Established			

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
Performance of NIPi Secretariat	2		JSC satisfied			
NIPi Director	2		In place with support staff			
Secretariat staff orientation	2		Oriented on NRHM & NIPi			
Secretariat systems	2		Established			
Secretariat reports	2		Submitted to JSC as per schedule			
National Node of the Enabling Mechanism (NNEM) of the Child Health Resource Network	2		Established			
NNEM and supporting technical staff	2		Recruited and oriented on NRHM & NIPi			
NNEM systems	2		Established			
NNEM reports	2		Submitted to JSC as per schedule			
Holding, disbursement & allocation of funds that the JSC and MFA have decided to be channeled through UNOPS as Local Fund Agent	2		Monitored & accounted for in transparent & efficient manner as per agreement			UNOPS LFA

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
Programs and projects for which funds are being entrusted to UNOPS	2		Followed up, monitored & reviewed as per EoL agreements with each implementing agency			UNOPS LFA
NIFI monitoring formats	2		Designed, field tested & used			
State NRHM focal points	2		UNOPS introduced			
Bidding process	2		Guidelines established			
Additional activities identified by JSC	2		Implemented as per established standards			
Feedback from implementing agencies and the field on the role of UNOPS as LFA	2		Satisfactory			
Monitoring of immunization outreach sessions by district and state functionaries	3		50% increase			
Identified laboratories with appropriate facilities for measles outbreak investigation	3		At least one per state			

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
Training of state and district functionaries in new techniques and MIS for surveillance of VPDs in NIPI and EAG states	3		70% of all functionaries			
Improved MIS system for surveillance for VPDs	3		Established in the country & followed in targeted states			
Logistics in place to support measles outbreak investigation	3		In place in targeted states			
Capacity built to support measles outbreak investigation	3		Built in targeted states			
Adoption and implementation of the new vaccine management software	3		MSDs and 65 regional stores			
Training of state and district officers in effective vaccine store management	3		All trained			
Status of national and 5 state nodal pre-service training centres for IMNCI	3		Operational			
Pre-service training in IMNCI in the NIPI states by all the teaching institutions (medical, nursing and ANM schools) in NIPI states	3		All conducting training			

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
MTR by the Gol Maternal Health Division in all the NIPI states	3		Conducted			
Quality Assurance Cells for specialised training in maternal health programmes at national and state level	3		Established and functional			
Guidelines for certification of medical colleges to function as training centres for EOC	3		Established, accepted and used by state governments, including NIPI states			
EOC nodal training centres	3		At least one functional in each state			
Number of Medical Officers trained in life saving anaesthetic skills in EOC in the NIPI states	3		50% increase			
Deliveries in the NIPI states conducted by SBAs	3		50%			
Assessment survey completed in all the ANM training centres in the 5 NIPI states and actions initiated with the state government for filling gaps based on findings	3		NIPI state governments implementing recommendations in all ANM training centres			

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
System for accreditation of facilities offering maternal and child health services in the country with special focus on the NIPi states	3		NIPi states implementing & monitoring accreditation process for government & private facilities			
Outcomes						
% Mothers who had at least 3 Ante- Natal Care visits during last pregnancy	4			1, 3	2, 3, 4, 5	1
% Institutional births	4			1, 3	2, 3, 4, 5	1
Average retention period (hours) for institutional delivery	4			1, 3	2, 3, 4, 5	1
% Home deliveries assisted by a Doctor/ Nurse/LHV/ANM	4			1, 3	2, 3, 4, 5	1
% Children had check up within 24 hours after delivery (based on last live birth)	4			1, 3	2, 3, 4, 5	1
% Children had check up within 10 days after delivery (based on last live birth)	4			1, 3	2, 3, 4, 5	1

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
% Mothers who received post natal care within 48 hours of delivery of their last child	4			1, 3	2, 3, 4, 5	1
(%#?) Referrals for mothers with illness and complications during pregnancy	4			1, 3	2, 3, 4, 5	1
% Children (age 6 months and under) who were exclusively breastfed	4			1, 3	2, 3, 4, 5	1
% Children (12-23 months) fully immunized (BCG, 3 doses of DPT, Polio and Measles)	4			1, 3	2, 3, 4, 5	1
% New born babies immunized zero dose polio and BCG	4			1, 3	2, 3, 4, 5, 6, 7	1
% New born babies breastfed within 1 hour of birth	4			1, 3	2, 3, 4, 5, 6, 7	1
% New born babies with birth weight taken after delivery at home	4			1, 3	2, 3, 4, 5, 6	1
# Women facilitated / motivated by ASHAs for ANC	4			1, 3	2, 3, 4, 5, 6	1
# Women facilitated / motivated by ASHAs for delivery at health facility	4			1, 3	2, 3, 4, 5, 6	1
UNOPS - Local Fund Agent						

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
Techno managerial support						
Number of positions sanctioned as NIPi supported State staff (for each NIPi focus State)	4				LFA format; Quarterly, State level	National Programme in charge, M&E Associate
Number of positions filled vis-à-vis sanctioned NIPi supported State staff (for each NIPi focus State)	4				LFA format; Quarterly, State level	National Programme in charge, M&E Associate
Number of Child health Managers/ Deputy Managers/ Maternal & Child Health Managers supported by NIPi (for each NIPi focus State)	4				LFA format; Quarterly, District level	State Programme in charge, Programme Associate
Number of Child Health Supervisors at facility level (for each NIPi focus State) supported by NIPi	4				LFA format; Quarterly, District level facility	State Programme in charge, Programme Associate
Number of Deputy Child Health Supervisors at facility level (for each NIPi focus State) supported by NIPi	4				LFA format; Quarterly, District level facility	State Programme in charge, Programme Associate
Number of Block Child Health Managers (for each NIPi focus State) supported by NIPi	4				LFA format; Quarterly, Block level	Programme Associate, District Child Health Manager
Number of Yashodas supervised by Child Health Supervisors/ Deputy Child Health Supervisors	4				LFA format; Quarterly, District level facility	Programme Associate, District Child Health Manager

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
Funds spent on remuneration of NIPi supported staff	4				Financial expenditure format-LFA; Quarterly, State level	State Programme Officer
Funds spent as mobility support, organizing meetings, office expenses etc	4				Financial expenditure format-LFA; Quarterly, State level	State Programme in charge, Programme Associate
ASHA payments made for HBPNC	4				Total Number of PNC cards filled, validated; Monthly, Quarterly, Annual, Block, District, State levels	State Programme in charge, Programme Associate
Formulation of District Programme Implementation Plans (PIPs) for child and maternal health under NRHM	4				District PIPs, District level	Programme Associate, District Child Health Manager
Micro plans for Routine Immunization	4				State, District, Block and Community Levels	State Programme in charge, Programme Associate
Yashodas						
Number of active Yashodas	4					Programme Associate, State NIPi Office; M&E Associate, LFA; Programme in charge, LFA
Number / %age of births including still births by gender	4					Programme Associate, State NIPi Office; M&E Associate, LFA; Programme in charge, LFA

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
Number / %age of stillborns by gender	4					Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Number / %age of neonates identified with illnesses by gender	4					Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Number / %age of neonates dead after birth by gender	4					Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Number / %age of neonates discharged within 6 hours by gender	4					Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Number / %age of neonates discharged between 6-12 hours by gender	4					Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Number of neonates discharged between 12-24 hours by gender	4					Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Number / %age of neonates discharged between 24-48 hours by gender	4					Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Number / %age of neonates discharged after 48 hours by gender	4					Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
Average retention time of neonates by gender	4					Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Number / %age of neonates given 0 dose Polio by gender	4					Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Number / %age of neonates given BCG by gender	4					Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Number / %age of neonates breastfed within 1 hour by gender	4					Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Number / %age of neonates weighed by gender	4					Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Number / %age of neonates with weight more than 2.5 kgs by gender	4					Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Number / %age of neonates with weight between 2-2.5 kgs by gender	4					Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Number / %age of neonates with weight less than 2 kgs by gender	4					Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
Number of kits distributed to mothers	4					Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Home Based Post Natal Care						
Number of newborns and mothers visited up to 6 times during 1st 6 weeks after delivery	4				ASHA PHC card	Supervisor at facility - Block Office for ASHAs
Number of Mothers counselled at home in basic newborn care, hygiene, breastfeeding and danger signs	4				ASHA PHC card	Supervisor at facility - Block Office for ASHAs
Number of mothers and children referred to hospital with danger signs	4				ASHA PHC card	Supervisor at facility - Block Office for ASHAs
Percentage of home newborns breastfeed within 1st hour of life	4				ASHA PHC card	Supervisor at facility - Block Office for ASHAs
Sick Newborn Care Units						
Admissions (Inborn/ Outborn) by gender (%age)	4				SNCU Monitoring form	Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Birth weight (Inborn/Outborn)- more than or equal to 2.5 kgs/ less than 2.5 kgs by gender (%age)	4				SNCU Monitoring form	Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
Gestation (Inborn/Outborn)- more than 37 weeks; less than or equal to 37 weeks by gender (%age)	4				SNCU Monitoring form	Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Causes for admission by gender (%age)	4				SNCU Monitoring form	Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Deaths (Inborn/ Outborn) by gender (%age)	4				SNCU Monitoring form	Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Primary cause of Death by gender	4				SNCU Monitoring form	Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Information on deliveries in the hospital- normal, assisted, C section by gender	4				SNCU Monitoring form	Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Live births by gender (%age)	4				SNCU Monitoring form	Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Still births by gender (%age)	4				SNCU Monitoring form	Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Neonatal deaths by gender (%age)	4				SNCU Monitoring form	Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
Maternal deaths by gender (%age)	4				SNCU Monitoring form	Programme Associate, State NIPI Office; M&E Associate, LFA; Programme in charge, LFA
Routine Immunization						
% Session held per planned during month	4				Universal Immunisation Program (UIP)/HMIS	
% sessions with all vaccines and syringes available	4				Gol monitoring form (MF)	
% sessions with ASHA / mobilizer	4				UIP/HMIS	
% BCG & OPV doses administered at birth for institutional deliveries*	4				Yashoda data	
DPT1 to DPT3 drop-out rates	4				UIP/HMIS	
% full immunization coverage at 1 yr*	4				UIP/HMIS	
% unimmunized children*	4				Gol MF	
% partially immunized children*	4				Gol MF	
% fully immunized children*	4				Gol MF	
Financial Monitoring						
UNICEF & WHO						

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
JSC Approved activities, outputs and expenditures	4					Intervention data shared with NIPi Secretariat for analysis. Semi-annual and annual reports to Royal Norwegian Embassy (RNE) and JSC.
Research Institutes						
Operations research activities	4					Operations Research Committee coordinated by NIPi Secretariat. Reports to RNE and JSC.
Percentage of newborns exclusively breastfed throughout first 6 weeks of life	4				ASHA PHC card	Supervisor at facility - Block Office for ASHAs

PI Documents:

1. UNICEF/ELA Appendix 1: Project Summary (13 December 2006)
2. UNOPS ELA Appendix 1: ELA Appendix 1: Project Summary (23 November 2006)
3. WHO ELA Appendix 1: Program Summary (13 December 2006)
4. Monitoring and Evaluation Strategy Document Version 1: August 2010

Responsibilities for M&E

1. The M&E Advisor, NIPi Secretariat will have overall responsibility for compilation and coordination.

Specified data sources

1. DLHS3; 2. DLHS4; 3. NIPi Surveys – Baseline, Midline, Endline; 4. HMIS; 5. Other independent studies; 6. ASHA data; 7. Yashoda data;
8. NFHS; 9. Sample Registration System; 10. Coverage Evaluation Survey; 11. UN Global Summit Reports; SNCU Monitoring form.

Appendix M:

Norway - India Partnership Initiative Evaluation Options Appraisal

Appraisal Area 1: Map the basis for Norway's support to the PI using project documents²⁵ and assess the adequacy and quality of such documentation in terms of potential "impact evaluability" of the PI towards the end of the initial implementation period (typically 5-6 years).

Team's Appraisal:

It can be argued that the defining characteristic of the NIPI lies in its relation to the National Rural Health Mission (NRHM). It originated as a high level political decision to allocate Norwegian aid funding to provide flexible, 'catalytic support' to the NRHM. The Mid-Term Review argues that *"NIPI should be owned and seen as an integral part of NRHM: supported externally by Norway, UNICEF, WHO and UNOPs"*. In line with the Paris Declaration, *"NIPI would operate in a national system of rigorous monitoring and evaluation"*. While there has been much discussion of the balance between implementation (support to scaling-up interventions of assumed value) and innovation, there appears to have been no strategic decision as to the relative importance of these program themes.

Given this background, a key decision is the extent to which any impact evaluation will be required to address attribution. It should be noted that under the standard DAC definition, an impact evaluation is simply one that focuses on *"long-term effects produced by a development intervention"* and thus attribution is not a necessary component. If attribution is a requirement, it may be useful to remember that one way to frame this question is *"what contribution did the allocated resources make to the observed changes resulting from the overall intervention?"* In the present case, it may be reasonable, if difficult, to attempt to estimate the contribution made by NIPI resources. This was recognized early on by the Joint Steering Committee (Minutes of 6th JSC) in which it was noted that it be *"not possible to measure the contribution of NIPI in empirical terms against all of the interventions since some of the interventions have multiple inputs"*.

One potentially simplifying factor is the apparent focus on a limited number of key interventions. Again, this approach was introduced early on NIPI when JSC members argued for a focus on a limited number of interventions in order to better allow empirical measure of impact. Substantial NIPI resources, for example have been allocated to the ASHA and Yahoda/Mamta interventions. One plausible line of analysis might be to assess the contribution that NIPI funds have made to the training

²⁵ Including: project design, results chain, indicators and success criteria, implementation arrangements, methods, approaches, types of activities, monitoring and evaluation systems, stakeholders, intended target groups and geographic coverage, and so on

and support for these two groups and then to estimate the extent to which they have been responsible for increases in specific, measures outcomes (e.g. skilled attendance at birth and post-partum follow-up visits for both mother and newborn). Note that for some outcomes indicators, for example, increased immunisation rates attributed to the NRHM, such a line of questioning may lead to a disappointing result at state or national levels, as the resources allocated under the NIPI are often dwarfed by those from the Gol.

The underlying premise of the innovation work under NIPI was to seize opportunities to test and demonstrate new and potentially effective means of addressing maternal and child health. The value of these innovations is then not in their isolated application, but in having them leverage the far greater resources available through NRHM. Programs aimed at created such leverage typically work with strategies and action plans to help create the environment for such uptake and then promote/ advocate for their application. The JSC and PMG return to this topic repeatedly with (e.g. *“scaling should be taken up only after impact assessment of the interventions to avoid potential HR issues in the state”* 7th JSC minutes). It appears that the NIPI has not created such a strategy or a systematic, documented approach to leverage their innovations in this manner.

In all, the available documents do not describe a well developed program where goals and purposes are linked to a set of expected outcomes, outputs and activities. If there was a hierarchy of clearly articulated, agreed, realistic and well explained targets, it might be possible to argue how such a mix of interventions could lead to the expected results, either alone or in combination with other efforts. An initial attempt at such a framework appears in an appendix to the 2008 NIPI Strategy clearly labeled as an example. However, that framework was apparently never fully elaborated among partners and finalized. In the absence of such guidance, the preferred course may be to focus on evaluating a key few individual interventions (such as the ASHA and/or Yahoda/Mamta) to create an adequate mapping of the results chain.

Throughout the materials, it is difficult to discern the roles to be played by each partner and how they work in concert to achieve results. It would appear that each works independently of the others and reports to the RNE and Joint Steering Committee. While performance measures were annexed to the Letters of Agreement between the RNE and each partner, those measures have not been incorporated into any PI-wide monitoring system or routinely reported. Again, while it is possible to specify the geographic areas in which UNOPS works with RNE resources, it is not apparent in which geographic areas UNICEF or WHO utilize the RNE resources or whether they see those resource as distinct from their on-going programs.

From NIPI inception through mid-2010, it seems that NIPI operated without a functioning M&E plan or structure. Based on recommendation from the MTR, a M&E Strategy was drafted and shared with the Team in early September 2010 and, with revisions, again in December. The Team is reviewing that document and preparing comments. Given the late stage of the program, it is strongly suggested

that the M&E strategy and plan have a limited, manageable focus on activity areas which are priority for the Norwegian support to the NRHM.

Appraisal Area 2: Identify contextual factors likely to influence intended and unintended PI impacts. Then assess the availability, adequacy, and quality of contextual factor data/information currently available, going back at least 5 years in time where possible.

Team's Appraisal:

The most significant contextual factor is the scale of other available resources – notably Gol resources and Gol policy direction and NRHM program implementation could have the greatest potential influence on NIPI. In any final assessment, NIPI interventions will have to be viewed within the context of the NRHM policy and programmatic direction and resources flows. For example, evidence from several data sets shows that facility-based deliveries are increasing rapidly related to the financial incentives provided through the Janani Suraksha Yojana (JSY) program. This increased demand for services could clearly effect the performance of the ASHA Yashoda/Mamta program (e.g. if resource planning does not account for increasing numbers of women seeking facility-based delivery services, these services may be inadequate to the demand). Likewise, programmatic emphasis on polio eradication efforts will impact on the routine immunization system.

In any future evaluation, careful consideration should be given to the sustainability of NIPI activities particularly those related to techno-managerial support. In the four focus states, a substantial cadre of MNCH staff has been hired and is being paid (either directly or indirectly) by NIPI. If these positions are to be sustained, they will have to be absorbed into either existing or newly created government positions. Similar concerns were expressed in JSC/PMG meetings in regards to the supervisory support needed for the Yashoda activity. More detailed consideration of contextual factors and underlying data should be guided with a more focused set of program outputs and outcomes (e.g. home-based newborn care, SNCU, Yashoda).

Appraisal Area 3: Review currently-available baseline data/studies and their appropriateness in terms of impact evaluation reliability and validity, keeping such questions as the following in mind:

- What are the “right” impacts needing to be measured/verified?
- Are these “right” impacts actually verifiable to acceptable standards given the existing baseline information available?
- Where information is lacking or incomplete, can data be (re)constructed?
- What counterfactuals need to be considered?

Team's Appraisal:

Facility assessments were conducted in 3 districts selected for NIPI interventions in each of 4 states. These typically covered the District Hospital, one or two Community Health Centers, 6 Primary Health Centers and around 30 Sub-Centers. These assessments compiled data existing at each facility and conducted a basic audit of staff, buildings, equipment and materials. A baseline household survey was under-

taken in each of these districts. The target population was currently married women who had given birth or been pregnant over the two years preceding the survey. A two-stage stratified cluster sample design was adopted in each district. A total of 50 primary sampling units (PSUs), either rural villages or urban wards, were selected in each district, using probability proportional to population size sampling from a list of all such PSUs based on the 2001 census. A list of eligible respondents was then constructed for each PSU and the final sample selected using systematic sampling. A design factor of two was assumed for key indicators, leading to an overall targeted sample size of 1200 respondents per district. However, this was later revised to ensure the inclusion of a larger sample of neonates and infants. The final sample included all neonates (0-28 days), 600 infants and 1200 children aged 12-23 months with oversampling of 10% to allow for non-response. Across the four states the final sample sizes were 1,395 neonates, 6,485 infants and 7,323 children 12-23.

Two separate agencies conducted the baseline studies, one covering Madhya Pradesh and Orissa, and the other Bihar and Rajasthan. The Team were provided with primary datasets originating only from the first agency and thus most of the comments here relate to that data. Our general impression is that the surveys were generally well managed and implemented, with the reservation that it has proved very difficult to access survey instruments and related materials. Good practice in such surveys is to finalise and maintain data and meta-data files such that any external analyst or reviewer can be assured of easy access to all necessary data and information. Our impression is that limited attention was paid to this requirement once the survey reports had been accepted.

The data appear to be reasonably complete and frequency distributions plausible for most variables. However, one general concern is that the household survey questionnaire contained over 700 items. Such large surveys are known to test both the patience of respondents and the persistence of enumerators. In particular, attempting to collect data on such a large number of variables almost inevitably implies that there is insufficient time for clarification and follow-up by enumerators to ensure that questions are either fully understood by respondents.

Our impression is that this is reflected to some extent in the overall quality of the data and would recommend that in future surveys a tighter focus on the key indicators would be preferable. To illustrate some of these quality issues, the frequency tables below focus on just two areas of particular relevance in terms of impact indicators – ANC/delivery and child vaccination.

The question relating to access to ANC allows for multiple responses – i.e. different types of provider may be seen on different visits. One minor point here, as with many of the other questions, is that non-response (blank) is treated as a negative response and there is no 'Don't Know' (DK) option. As in all such self-reported contact with the health system we are reliant on an assumption that the respondent can correctly identify the type of provider. This may lead to some doubts, for example as to the qualifications of the 'private doctors'.

ANC providers accessed

Provider	Frequency	%
Government Doctor	2676	37.7
ANM/Nurse/Midwife/LHV	2539	35.8
ASHA	48	0.7
Private Doctor	1279	18.0
DAI	10	0.1
Anganwadi/ICDS Worker	617	8.7
Other	50	0.7
No one	8	0.1
Total women in sample	7,094	

The proportion of women making an apparently large number of ANC visits raises questions as the extent to which terms such as 'ANC visit' were clearly explained to respondents or responses assessed by enumerators. As another point of comparison, the Team compared the NIPI baseline (2008) to the DLHS (2007-08) in these same districts and found that significantly larger proportions of women reported seeking any ANC care in the NIPI surveys compared to the DLHS. Based on available information, the Team cannot conclude that one of these surveys is more reliable or valid than another but intend simply to point to the differences²⁶.

Number of ANC visits

Number of visits	Frequency	%
1	561	8.4
2	1,775	26.7
3	1,564	23.5
4	853	12.8
5	672	10.1
6	437	6.6
7	357	5.4
8 or more	425	6.4

Some questions, such as that relating to Tetanus Toxoid vaccination do contain the preferred response options (Yes, No, DK). However, the observation that 108 women are coded in the No category and just 7 in the DK category, as compared to the 373 missing values, may imply that there was insufficient focus on these responses.

²⁶ It should be noted that each state-level NIPI Baseline Report overestimated the number of times women received ANC care by incorrectly using only those women who received any ANC as the denominator as opposed to all women surveyed. Key indicators such as percent of women who received 3+ ANC visits are, in some cases, substantially overestimated.

Tetanus Toxoid Vaccination

Received	Frequency	%
Yes	6606	93.1
No	108	1.5
Don't Know	7	0.1
Missing	373	5.3
Total	7094	

In contrast to the number of ANC visits, the number of TT vaccinations seems reasonable, with very few outliers. Careful review of the guidance given to enumerators may indicate why this should be the case but the full documentation was not available at the time of this review. Such evidence should be examined in any impact evaluation.

Number of Tetanus Toxoid Vaccinations

Number of Vaccinations	Frequency	%
1	360	5.5
2	5490	83.3
3	730	11.1
4	8	0.1
5	1	0.0
7	1	0.0

The survey questionnaire distinguishes between facility and non-facility births. For facility births the respondent is asked to identify the type of provider assisting. For non-facility births, there seem to be a series of yes/no questions as to whether a particular type of provider was involved. As can be seen, this seems to result in a small degree of multiple responses. It is not clear how this was resolved in terms of assessing if the birth was attended by a skilled provider. It would have seemed preferable for the enumerator/supervisor to determine the primary birth attendant in the field.

Provider assisting delivery

Provider	Frequency	%
Facility births	5,312	74.9
Government Doctor	2516	35.5
Private Doctor	850	12.0
ANM / Nurse	1905	26.9
Other	41	.6
Non-facility births	1,926	27.1
Midwife/LHV	20	0.3
Trained DAI	380	5.4
Untrained DAI	542	7.6
ASHA	37	0.5
ANM	53	0.7
Family member	553	7.8
Relative/friends	215	3.0
Other	126	1.8
Total	7,238	102.0

Some 68% of women were able to produce vaccination cards (a further 21% said that they had cards but could not show them to the enumerator), which should greatly enhance the quality of the immunisation data. Very few entries appear to have been uncompleted or undated. It is not clear if HBV data is not entered on these cards or was not transferred.

Vaccination Card Entries

	Dated Entry	Missing Value	Undated
BCG	4,523	1	5
Polio 0	2,485	1	4
Polio 1	4,078	1	14
Polio 2	3,566	1	12
Polio 3	3,019		14
DPT 1	4,013	1	16
DPT 2	3,499		13
DPT 3	2,966	1	16
Measles	1,810		15

Where a card was not available, women were asked to report types vaccinations on the basis of a description provided by the enumerator. There is some confusion as

to the definition and labelling of the variables relating to routine and ‘pulse polio’ vaccinations. The table below is therefore limited to reported polio vaccinations of all types. Note again that the limited number of No and DK response categories raises questions as to their interpretation.

Vaccinations reported by mother

Vaccination	Yes	No	DK
BCG	1,857	186	
Polio	1,997	186	
DPT	1,553	295	39
Measles	1,044	780	63

The proportion of children receiving more than 4 polio vaccination (13.2%) would suggest that it would have been useful to use follow-up questions to check respondents’ understanding of this question.

Number of Vaccinations Reported by Mother

Polio	Frequency	%	DPT	Frequency	%
1	306	15.3	1	206	2.9
2	402	20.1	2	459	6.5
3	656	32.8	3	795	11.2
4	174	8.7	4		
5	135	6.8	5		
6	38	1.9	6		
7 or more	89	4.5	7 or more		
DK	197	9.9	DK	93	1.3

The Indian District Level Household Surveys is an additional data resource which has already been demonstrated as applicable for impact assessment²⁷ and includes many variables of relevance to the NIPI. District –level surveys are also being conducted by development partners including UNICEF and UNFPA on similar topics and sometimes in the same districts. Altogether there is an adequate baseline on population-based outcome variables for the NIPI evaluation.

At this stage in the program (final two years 2011-2013), the NIPI may chose to focus their M&E efforts on a small number of key outcome variables which are most closely associated with priority program activities. This would allow the program to concentrate on program areas where implementation-level data is available (for example, the ASHA PNC card or Yashoda records). While data is being collected and compiled on these two activities, it is difficult to discern the quality

27 India’s Janani Suraksha Yojana, a conditional cash transfer programme to increase births in health facilities: an impact evaluation. Stephen S Lim, Lalit Dandona, Joseph A Hoisington, Spencer L James, Margaret C Hogan, Emmanuela Gakidou. Lancet 2010; 375: 2009–23.

(completeness, timeliness, accuracy) of that data. NIPI should invest in strengthening all aspects of that data collection – within existing government structures and systems. It seems feasible that any form of analyses would focus primarily on UNOPS-supported activities in the 12 selected districts in the four states. Data from partners (i.e. UNICEF) may be relevant for analyses of specific program areas (e.g. SNCU).

Overall, the documentation does not yet sufficiently detail what indicators *are actually being* collected, nor the quality and completeness of that data for impact evaluation purposes. The M&E Strategy and Plan (December 2010) implies that plans are only now being in putting in place to collect information. With nearly three full years of NIPI implementation and multiple sources of population-based data, NIPI should be in a position to produce regular, targeted summaries and analyses of performance on a small number of relevant indicators. However, there is still a lack of clarity as to the precise nature of the key output and outcome measures, targets and performance benchmarks, which partner is collecting those measures, and the quality of that data.

Appraisal Area 4: Summarize important findings and conclusions related to each of the above tasks.

Team's Appraisal

Recommendations:

1) With an approved no cost extension, it is recommended that any follow-up household survey is postponed until a detailed analyses plan is prepared that would take into account: the NIPI baseline data, the data available from the District-Level Household Surveys, implementation-level measures (including ASHA PNC cads and Yashoda records) and relevant findings from the qualitative research on ASHA and Yashoda/mamta. Any form of follow-up survey should be tailored specifically to address a few well-conceived questions about a limited number of priority NIPI support and associated analyses which triangulate these data sources. The best case scenario might entail time-series analyses from targeted districts, assessing associations between outcome measures with intensity of program implementation indicators based on available implementation-level data.

2) The overall value of the NIPI experience might be as a lesson learned for the Partnership Initiatives and future Norwegian support for MDGs 4 and 5. Indeed, each PI program designed after NIPI represents an improvement in terms of program clarity and the results chain – albeit with room for improvement. In addition, the emphasis on innovation needs to be matched with a strategy to document those innovations and strategically engage policy makers on their uptake.

3) The Team has identified three main options for moving forward with impact evaluation of the NIPI as follows:

The gold standard approach would have involved a pre-post case-control evaluation design in a selected number of intervention districts. Control districts would have been matched to intervention districts on a specific set of key socio-economic

characteristics and health system variables. In advance of implementation, the evaluators, with the support of province and district program managers, would have attempted to limit and control other inputs (i.e. other development partner inputs above and beyond the NRHM program) in both case and control districts for greater isolation of any NIPI “effect”. The sampling design for baseline and end line data collection would have sought to measure statistically significant changes both over time and between project and control districts. The design would include analyses of “exposure” variables to estimate the magnitude of the effect (outcomes) based on the intensity of NIPI-supported service use either at the individual or district level (e.g. the number of times that a ASHA visited). Finally, the gold standard approach would have included careful and systematic monitoring of implementation using indicators that had been developed, validated and used consistently throughout the project. *In practice, this gold standard approach is infeasible in the present context. Among the most significant factors are the large-scale changes underway due to the implementation of NRHM with resources dwarfing those of NIPI, the lack of a geographic focus for the three partners (UNOPS, UNICEF and WHO) and the associated diffusion of resources across different program activities would not allow for such impact analyses.*

It is, however, entirely feasible that elements of this approach could be used to evaluate some specific, targeted interventions for their effect on tangible outcomes. For example, examining whether (and how) ASHA activities increase the utilization of home-based newborn care practices or whether the presence of Yashodas improves specific aspects of facility-based delivery services. Such analyses could be relatively small, well-designed studies aimed at answering specific questions related to NRHM priority program elements and estimating the contribution made by NIPI resources.

A global health – normative option that would satisfy internationally-accepted standards might entail evaluation only in the 12 NIPI focus districts without the use of control districts (the de facto situation as control districts were not included at baseline). In those 12 project districts, evaluators would carefully account for the inputs of other development partners -- but not try to limit or control those inputs. Sample parameters would be designed to either a) determine with a known level of certainty (statistical power) whether pre-determined targets had been achieved, or b) measure statistically significant change over time (pre/post) in key variables (the preferred option). As with the “gold standard”, this design would also analyze the “exposure to” NIPI-supported NRHM providers/services, including program duration and intensity, and associate those variables with the desired outcomes (NRHM routine data and health information system data would be important sources). Finally, this approach should also include careful and systematic monitoring of implementation using valid indicators used consistently throughout the project. *At this point in the NIPI program (two-thirds through the extended 2006-2013 timeline) the Team believes that not enough attention has been devoted to the types of implementation-level monitoring needed to support this effort. Retrospectively creating this information, including quantifying the inputs of other partners, may be too time-consuming and costly for the value of the information produced.*

A “minimum option” would seek to triangulate available data sources (DLHS-3, NIPI baseline, UNFPA and UNICEF coverage surveys, NRHM routine data, health information system data as well as the upcoming DLHS-4) to either a) determine with known certainty whether NIPI targets were achieved (y/n) or b) measure statistically significant change in outcome measures over time. This option could include the planned NIPI mid-line and end-line survey data. Triangulation of data sources – including objective, systematic implementation monitoring at multiple levels (district, tensil, facility, provider) -- could help to identify associations using implementation timelines and trends in desired outputs and outcomes²⁸. This approach would focus more on whether change has been achieved and less on identifying and substantiating the NIPI contribution. *The Team believes that this option is entirely feasible for NIPI. If coupled with the recommendation above (i.e. to evaluate specific, targeted interventions for their effect on tangible outcomes), a well-rounded picture of the NIPI experience could be created and useful information generated for NRHM investments and direction.*

Documents reviewed:

36. Norway-India Partnership Initiative. Final Report of the Mid-Term Review. February 2010. Dr Ashok Dutta, Dr. Rani Gera, Dr Antoinette Pirie and Mr Stein-Erik Kruse.
37. Norway India Partnership Initiative. Annual Report 2008. UNOPS. UNICEF. WHO.
38. Minutes of the meeting of the 7th PMG held on 17-04-09.
39. Short summary of NIPI interventions in Bihar. Presented by Ravi Parmar. Executive Director, SHSB. 13-11-09.
40. NIPI Strategy Document. 2008.
41. Norway-India Partnership to achieve MDG4 (NIPI Program Document).
42. UNICEF-NIPI Implementation Status 2009 and Plan for 2010. 8th PMG mtg. 13-11-09.
43. Norway-India Partnership Initiative. Update on WHO Activities. 8th NIPI Programme Management Group Meeting. 13 November 2009
44. Minutes of the 6th JSC held on 28-08-2008.
45. Draft minutes of the 7th JSC held on 04-06-2009.
46. Minutes of the Meeting of the PMG held on 21-08-2008
47. Short summary of NIPI interventions in Madhya Pradesh. Dr. Manohar Agnani. MD NRHM. Madhya Pradesh. 13-11-09.
48. Short summary of NIPI interventions in Orissa, Shri G. Mathivathanan, IAS. Mission Director. NRHM, Orissa. 13-11-09.
49. Short summary of NIPI interventions in Rajasthan. Dr M L Jain, Director (RCH), Rajasthan. PMG mtg. 13-11-09.
50. Summary agenda notes for 8th Program Management Group meeting. Norway India Partnership Initiative on November 13th 2009. New Delhi.
51. Summary Report of Assignment for Norad AHHA re: NIPI, October-December 2008. Cliff Wang, HeSo.
52. Visit to Bihar State and Nalanda District, November 2008. Travel Report. By Cliff Wang and Olav Hernar.

²⁸ For an example of this approach see: System-level determinants of immunization coverage disparities among health districts in Burkina Faso: a multiple case study Slim Haddad, Abel Bicaba, Marta Feletto, Elie Taminy, Moussa Kabore, Boubacar Ouédraogo, Gisèle Contreras, Renée Larocque, Pierre Fournier. BMC International Health and Human Rights 2009, 9(Suppl 1). Published: 14 October 2009.

53. NIPI issues in advance of the mid-term review. Ambassaden i New Delhi, India. Kopi : UD- GIL UD- REG. 19.08.2009.
54. NIPI. Monitoring and Evaluation Strategy and Plan, 2010-2013. DRAFT version 4.
55. International Institute for Population Sciences (IIPS), 2010. District Level Household and Facility Survey (DLHS-3), 2007-08: India. Bihar: Mumbai: IIPS. Accessed at www.rchiips.org.
56. International Institute for Population Sciences (IIPS), 2010. District Level Household and Facility Survey (DLHS-3), 2007-08: India. Madhya Pradesh: Mumbai: IIPS. Accessed at www.rchiips.org.
57. International Institute for Population Sciences (IIPS), 2010. District Level Household and Facility Survey (DLHS-3), 2007-08: India. Orissa: Mumbai: IIPS. Accessed at www.rchiips.org.
58. International Institute for Population Sciences (IIPS), 2010. District Level Household and Facility Survey (DLHS-3), 2007-08: India. Rajasthan: Mumbai: IIPS. Accessed at www.rchiips.org.
59. Baseline Survey On Child And Related Maternal Health Care. Rajasthan. Revised Report. Prepared For: Norway India Partnership Initiative, New Delhi. DRS Pvt. Ltd. GVL Narsimha Rao Development & Research Services Private. August 2009.
60. Baseline Survey On Child And Related Maternal Health Care. State Report: Bihar. Prepared For: Norway India Partnership Initiative, New Delhi. DRS Pvt. Ltd. GVL Narsimha Rao Development & Research Services Private. August 2009.
61. Baseline Survey On Child And Related Maternal Health Care. Madhya Pradesh. Revised Report. Prepared For: Norway India Partnership Initiative, New Delhi. DRS Pvt. Ltd. GVL Narsimha Rao Development & Research Services Private. August 2009.
62. Baseline Survey On Child And Related Maternal Health Care. Orissa. Revised Report. Prepared For: Norway India Partnership Initiative, New Delhi. DRS Pvt. Ltd. GVL Narsimha Rao Development & Research Services Private. August 2009.
63. Activities by State Health Societies in Orissa, Bihar, Madhya Pradesh and Rajasthan under Norway – India Partnership Initiative. Program Update, January 2009.
64. Concurrent Assessment of Janani Suraksha Yojana (JSY) in Selected States. Bihar, Madhya Pradesh, Orissa, Rajasthan, Uttar Pradesh. UNFPA. 2009.
65. Coverage Evaluation Survey 2009 (CES 2009). UNICEF.
66. Activities by State Health Societies in Orissa, Bihar, Madhya Pradesh and Rajasthan under NIPI. Program Update. January 2009.
67. NIPI Expenditure Report. (Excel spreadsheet prepared for MTR committee).
68. UNICEF ELA Appendix 1: Project Summary (13 December 2006)
69. UNOPS ELA Appendix 1: Project Summary (23 November 2006)
70. WHO ELA Appendix 1: Program Summary (13 December 2006)

Appendix N: Norway - Nigeria Partnership Initiative Profile

Overview	Norway is investing 250 million NOK to improve maternal, newborn and child health in Northern Nigeria. In a highly efficient move, Norway partnered with an existing Dfid-funded program focused on routine immunization bringing additional resources and expanding the project scope to include maternal, newborn and child health services improvement. The MNCH component of the program focused on expanding access and demand for life-saving obstetrical services among other areas.	
PI duration:	<p>Start date Mid-2008 (Agreement signed 13 June 2008)</p>	<p>End date 2012</p>
PI purpose	The goal of the initiative is to improve maternal, newborn and child health in Northern Nigeria. The purpose is to improve quality and access to MNCH services in four States – Jigawa, Katsina, Yobe and Zamfara.	
Stated objective	<p>Improved quality & access to MNCH services in Jigawa, Katsina, Yobe and Zamfara. The initiative involves two main components.</p> <ul style="list-style-type: none"> • An established routine immunisation programme (PRRINN) working in four States (Jigawa, Katsina, Yobe and Zamfara) • A new MNCH programme focussing on three of these States (Katsina, Yobe, Zamfara), although cross-cutting issues may apply to Jigawa also. <p>Each component of the initiative has its own outputs and activities.</p>	

Funding:	Total: Norway's contribution in Nigeria of 250 million Kroner will be managed through a delegated cooperation agreement with the UK as outlined in a Arrangement on Delegated Cooperation (Appendix 3). This is consistent with the Paris Declaration and Nordic Plus agreements. The Norwegian contribution more than doubles the existing DFIDinvestment of 19 million GBP.	
	By partner:	By outputs (Programme Memorandum)
	DFID	1. Strengthened State and LGA governance of PHC systems geared to MNCH. Indicative budget: NOK 32,077,238 (13%)
		2. Improved human resource policies and practices for PHC. Indicative budget: NOK 35,409,938 (14%)
		3. Improved delivery of MNCH services via PHC system. Indicative budget: NOK 73,319,400 (29%)
		4. Operational research (OR) providing evidence for PHC stewardship, MNCH policy, service delivery and effective demand. Indicative budget: NOK 31,244,063 (13%)
		5. Improved information generation with knowledge being used in policy and practice. Indicative budget: NOK 32,077,238 (13%)
		6. Increased demand for MNCH services. Indicative budget: NOK 40,825,575 (16%)
PI within landscape for MDGs 4&5 / development partners	Other development partners instrumental in the geographic area and in the MDGs 4 and 5 include: EC, USAID, UNICEF, WHO, UNFPA., World Bank and GAVI HSS.	
Geographic Areas	Area 1: Jigawa	NSLPs, PRRINN and PATHS; PATHS2 state
	Area 2: Katsina	PRRINN, UK Norway MNCH and GEP; EC operating with PRIME and SRIP;
	Area 3: Yobe	PRRINN and the UK Norway MNCH initiative., EC operating with PRIME and SRIP;
	Area 4: Zamfara	PRRINN and the UK Norway MNCH initiative ;
Criteria for selection	<p>Sites include 3 out of 4 states – Katsina, Yobe and Zamfara are already supported with a significant DFIDinvestment and where DFID-supported PRRINN is already running. The fourth PRRINN state Jigawa where DFIDPATHS is operating is to be included in relevant cross-cutting issues such as OR. The DFIDPRRINN states were chosen on the basis of health statistics, logistics for programme management – i.e. two pairs of contiguous states, Katsina, Jigawa, Yobe and Zamfara.</p> <p>Dfid's health systems strengthening program states (PATHS2-- including Jigawa) were chosen through a benchmarking criteria delineated by the World Bank/DFIDCountry partnership strategy.</p>	

Partners description	Partners	Responsibilities
	DFID	Since early 2007, DFID has supported a programme focused on routine immunisation in 4 northern States (PRRINN) with a value of £19 million. DFID further contributes through an ongoing health systems strengthening program, (PATHS and PATHS 2 starting from June 2008 which continues support for Jigawa state)
Institutional arrangements:	<p>Norway's contribution is managed by DFID through a delegated cooperation agreement with the UK as outlined in a Arrangement on Delegated Cooperation. The MNCH initiative is implemented by the same consortium (Health Partners International, Save the Children UK, and GRID Consulting) implementing the Dfid-supported PRRINN and was chosen through an intl. tender with a September 2008 award date.</p> <p>DFID reports to Norway as outlined in the Arrangement and the partners conduct an annual programme review followed by an annual meeting; there is Joint monitoring and evaluation of the two components. A more comprehensive review is planned for the third year to decide on possible expansion to more States or adjustment to the initiative.</p>	
Implementation timeline	Planned To be completed	Actual To be completed
Elements of Innovation (as described in Desk Review)	Some innovative and catalytic initiatives were identified in the project proposal, including: performance-based funding, financial incentives, and Community Based Planning and Services.	
Operations Research activities	Activity	Status
	Approximately 5% of the total budget is to be earmarked for operations research	

Appendix O: Norway - Nigeria Partnership Initiative Monitoring and Evaluation Plan

This table shows a compilation of indicators drawn from available program documentation – it does not represent the Team’s recommendations on appropriate indicators for evaluating PI performance or impact.

Indicators (as included in PI documents)	PI Document	Baseline	Target	Baseline data sources	M&E Data sources	Arrangements / Responsibility for data collection
Goal						
U5MR	1	153	140	1		
% of births attended by a skilled birth attendant (SBA) in targeted CEOC clusters	1	38.9%	50%	1		
Purpose						
% of infants fully immunised by first birthday	1	16%	65%	1, 2, 3	1, 2, 3, 4	
% of women aged 15-49 have appropriate TT doses	1	15%	75%			
Caesarean section rates (in targeted CEOC clusters)	1	0.5%	>1%			
% of women receiving ANC	1	21%	45%			
Measles incidence	1	22,250	2,225	6	6	
Polio incidence	1	237	0	6, 7	6, 7	

Indicators (as included in PI documents)	PI Document	Baseline	Target	Baseline data sources	M&E Data sources	Arrangements / Responsibility for data collection
Output 1: Strengthened state and LGA governance of PHC systems geared to RI and MNCH						
Leadership of annual review and health planning process in all states	1	Facilitated and supported by programme	Led by state teams with no support from programme			
Successful access to new Federally managed health funds by states	1	One year in all states	Four years in all states			
Availability of PHC budget and expenditure reports for LGAs /Gundumas	1	Limited data available in 3 states	Available in 80% LGAs/ Gundumas			
Number of states with their State Health Plan incorporated into their State Development Plan	1	Some linkage to SEEDs in 2 states	4 States			
State health plans reflect project data from 2010	1	Few examples of evidence based planning	At least 7 examples of evidence based planning in each state plan			
Number of donor PHC programmes reflected in State and LGA annual health plans	1	1 per state	At least 3 per state			
State inter-agency coordinating committees (SIACCs) support for RI through PHC system in all states	1	Little	RI fully integrated			
Number of donor field missions and reviews done jointly	1	2	8 in total over project			

Indicators (as included in PI documents)	PI Document	Baseline	Target	Baseline data sources	M&E Data sources	Arrangements / Responsibility for data collection
Output 2: Improved human resource policies and practices for PHC						
% of targeted facilities with at least one health worker trained in LSS	1	5%	100%			
HR policies and plans developed, operationalised, and implemented in each state	1	Some	Implemented			
Number of health professionals trained annually	1	904				
% of professional staff given in-service training in MNCH in targeted PHC facilities	1	0%	100%			
Output 3: Improved delivery of MNCH services (including RI) via the PHC system						
% LGAs reaching performance ranking tool (PPRHAA) scores over 75%	1	18%	60%			
Number of PHC facilities providing basic emergency obstetric care	1	1	36			
Systems for effective supervision in each State	1	Designed	Visits planned, financed and implemented by each state			
Number of 1-year-old children immunised against measles	1	1,26,439	485,624			
% of health facilities providing RI experiencing vaccine stock-outs of TT	1	38%	15%			

Indicators (as included in PI documents)	PI Document	Baseline	Target	Baseline data sources	M&E Data sources	Arrangements / Responsibility for data collection
% of PHC facilities with tracer drugs available	1	50%	66%			
Output 4: Operational research providing evidence for PHC stewardship, RI and MNCH policy and planning, service delivery, and effective demand creation						
Number of pieces of OR into supply & demand aspects of MNCH feed into programme	1	None	6 in total			
State plans reflect OR results	1		1 example per state per year (12 in total)			
Output 5: Improved information generation with knowledge being used in policy and practice						
Demonstrated level of understanding in use of information by trained HMIS officers in each state	1	Some	Substantial understanding			
State plans increasingly built on evidence from HMIS	1	None	Substantial			
% of LGAs with HMIS MNCH data collated at State level on a monthly basis	1	<10%	85%			
Output 6: Increased demand for MNCH (including RI) services						
Increased political support for MNCH (including RI) evidenced by high level public events	1	1 at state level; 1 at LGA level	1 at state level; 2 at LGA level each year			

Indicators (as included in PI documents)	PI Document	Baseline	Target	Baseline data sources	M&E Data sources	Arrangements / Responsibility for data collection
% wards with a development committee and/or health partnership implementing a community action plan	1	4%	40%			
% women in targeted areas who have standing permission to take their child to a health facility	1	55%	83%			
% women who know at least four of the maternal danger signs in targeted areas	1	1.4%	40%			
% facility health committees for intervention facilities in targeted areas actively monitoring drugs	1	0%	75%			
% mothers of children <2 in targeted areas who know the childhood vaccination schedule	1	10%	60%			
% never immunised children <2 in targeted areas	1	25%	11%			
Output 7: Improved capacity of Federal Ministry level to enable States' MNCH (including RI) activities	1					
Formal systems for leveraging, accessing and utilising additional PHC funding	1	None	Systems functioning			

Indicators (as included in PI documents)	PI Document	Baseline	Target	Baseline data sources	M&E Data sources	Arrangements / Responsibility for data collection
% vaccines and supplies to states delivered on time by Federal level	1	TBD	100%			
Agreed strategies to improve efficiency of RI	1	Some strategies developed	Strategies implemented			

PI Documents:

5. PRRINN-MNCH Annual Review Draft Narrative Report, Carol Bradford and Maisha Strozier, 10 March 2010 Appendix 11

Specified data sources:

1. DHS 2008
2. NICS 2006, 2009, 2012
3. MNCH Household Surveys 09, 10, 12
4. MNCH Mini Household Survey 2011
5. HMIS/PMS
6. Integrated Disease Surveillance and Response (IDSR) previously DSN
7. WHO surveillance system
8. State health plans and budgets

Appendix P: Norway - Nigeria Evaluation Options Appraisal

Appraisal Area 1: Map the basis for Norway's support to the PI using project documents²⁹ and assess the adequacy and quality of such documentation in terms of potential "impact evaluability" of the PI towards the end of the initial implementation period (typically 5-6 years).

Team's Appraisal:

The strategy of the PRRINN-MNCH intervention, as indicated by the MNCH Inception Review of February 2009, is to target selected programme activities on a number of catchment areas or 'clusters' (typically 2 or 3 LGAs) in three states, Katsina, Yobe and Zamfara. The intention was to focus on one cluster in each state in the initial phase, with the aim of covering up to 3 clusters over the course of the programme. A CEOC hospital was to be established in each of these clusters linked to 4 BEOC facilities, which would in turn have a referral chain from PHC facilities. It was intended that a substantial component of overall M&E activities would target these clusters and facilities. The revised monitoring plan envisages systematic monitoring of all facilities in each cluster where activities are underway. To underline this strategy, the designation 'Programme Monitoring Sites' (PMS) has been adopted by the M&E framework in referring to these clusters. The intention was that work on the PMS system would be undertaken alongside that on the development of CEOC and BEOC services within each cluster.

On the assumption that the basic facility implementation plans are successful, i.e. that each cluster acquires a well equipped and staffed CEOC hospital linked to 4 such BEOC facilities, direct comparisons with non-cluster areas (based on the planned household surveys) in terms of service utilisation indicators are obviously likely to demonstrate a substantial advantage for populations living within clusters, even though there are likely to be considerable spill-over effects, which will have to be taken into account in any analysis. It should also be noted that the original motivation for the intervention was the very poor quality of services in the program states, again suggesting that substantial improvements from a very low base could reasonably be expected. Given this background, it may be most interesting for any impact evaluation to explore the process of service improvement. For example, analysing the time line linking facility construction/repair, provision of equipment, staffing, community engagement activities, etc. with service utilisation. The plan to focus M&E resources on the cluster facilities should make this a serious possibility.

²⁹ Including: project design, results chain, indicators and success criteria, implementation arrangements, methods, approaches, types of activities, monitoring and evaluation systems, stakeholders, intended target groups and geographic coverage, and so on

Appraisal Area 2: Identify contextual factors likely to influence intended and unintended PI impacts. Then assess the availability, adequacy, and quality of contextual factor data/information currently available, going back at least 5 years in time where possible.

Team's Appraisal:

As with the other PIs, the crowded environment in which development partners operate is an important contextual variable. The decentralised nature of the programme is a central issue in determining an appropriate impact evaluation strategy. National partners in each state play a lead role in determining both the range of interventions and the detailed intervention process. Gaining political and community support for specific interventions has been a major objective. This would suggest that for some outcomes there will be limited value in aggregation of findings to programme level, though this will clearly be required for key purpose level indicators.

A further complication in terms of overall impact assessment arises from the focus on specific clusters within each state, particularly as the number of such areas will be determined as the programme progresses. If the impact evaluation were to be limited to cluster level – i.e. what improvements have occurred in cluster as compared to non-cluster areas – the rational course of action for managers would be to limit expansion of the programme, allocating their resources to improving services as much as possible in as few clusters as possible. On the other hand, insisting on too rapid an expansion of the program to new areas could greatly increase the risk of relative failure which has been the fate of many previous health sector interventions in Nigeria. The project appraisal document seems to leave this question open, recognising that the available resources are far too limited to transform services across the program states and indicating that the primary aim is *“helping to achieve a rapid and tangible improvement in the MNCH outcomes of poor women and children and to generate greater political commitment to leverage more resources for the wider reform agenda through a functioning PHC”*. On balance, it may well be worthwhile to err on the side of caution and accept that the beneficiary population may be somewhat more limited than perhaps originally intended. A clearly successful intervention, even on a small scale, could provide a valuable demonstration model for use in advancing the above agenda.

Appraisal Area 3: Review currently-available baseline data/studies and their appropriateness in terms of impact evaluation reliability and validity, keeping such questions as the following in mind:

- *What are the “right” impacts needing to be measured/verified?*
- *Are these “right” impacts actually verifiable to acceptable standards given the existing baseline information available?*
- *Where information is lacking or incomplete, can data be (re)constructed?*
- *What counterfactuals need to be considered?*

Team's Appraisal:

Considerable efforts were made over 2009 by both donors and project managers to refine the indicators in the log frame for the combined programme. There is

probably little to be gained by re-visiting these indicators at this time. Most of the purpose indicators target increased service utilization, for example in terms of ANC, attended births and immunisation coverage. The lack of indicators of quality of service has been a cause for concern by some reviewers but previous experience in Nigeria would strongly suggest a preference for simplicity of measurement when selecting key indicators, and measurements of service quality are generally regarded as problematic. Baseline estimates for all these indicators have been compiled and there is a well documented strategy in place to deliver mid-term and end-line estimates.

A series of 'MNCH Surveys' is planned, covering both households and facilities. The first of these took place early in 2009 and covered Katsina, Yobe and Zamfara. The intention is to repeat the household surveys in 2010 and 2012, with a possible 'light' version undertaken in 2011. In addition to these programme-based exercises, it is intended that data from a number of planned national surveys, including a DHS in 2012 or 2013 and a National Immunisation Coverage Survey (NICs) in 2012, will be accessed to provide state-level estimates of selected coverage indicators.

The baseline household survey used a two stage cluster sampling design in which enumeration areas (EAs) were first selected with probability proportional to population size and then households were selected using systematic sampling. The sample was stratified, with EAs being classified as 'core', areas which were targeted by MNCH activities, or 'non-core', all other areas. In each state, 30 EA were sampled from the first group and 15 EAs from the second. A systematic sample of 47 households was then selected in each EA and all women aged 15-49 approached for interview. The final sample size was 7,442. The survey is well documented and appears to have been implemented with considerable attention to survey quality. However, the history of household surveys in Nigeria encourages caution. It is of interest that one programme document indicates widely differing estimates of immunisation coverage for the MNCH states across DHS, MICS and NICS surveys for similar time periods.

Appraisal Area 4: Summarize important findings and conclusions related to each of the above tasks.

Team's Appraisal

Recommendations:

The Partnership Initiative in Nigeria is already following a **global health – normative** approach to evaluation. The aim of that evaluation is to measure change over time in specified outcome indicators in both cluster and non-cluster areas within each state. The Team's concern, is that 'success' at this level, given the dysfunctional nature of services prior to the intervention, will need to be carefully interpreted in terms of scaling up to state level. We would strongly recommend that this is a case where the use of implementation data to link inputs to intermediate outputs/outcomes along the project timeline to analyse the implementation process could be of greater value than a simple assessment of project success.

The set of indicators seems well-tailored to the priority interventions although several of the goals- and purpose-level measures seem beyond the attributable “reach” of the project (under-five mortality, measles and polio incidence).

Documents reviewed:

1. UK-Norway MNCH Initiative in Northern Nigeria. Outline of Joint Programme. May 2008.
2. The UK-Norway MNCH Initiative in the Northern States Of Nigeria. Desk Appraisal April 2008. Mercy Bannerman. Nordic Consulting Group
3. PRRINN-MNCH Annual Review Draft Narrative Report, Carol Bradford and Maisha Strozier, 10 March 2010 Appendix 11.
4. Project Memorandum. Nigeria Northern States. Maternal, Newborn And Child Health Programme. April 2008.
5. PRRINN-MNCH Baseline Studies 2009. Summary Report.
6. Arrangement On Delegated Cooperation Between UK Department For International Development (Dfid) and the Norwegian Ministry Of Foreign Affairs (MFA) regarding support to Maternal, Newborn And Child Health Initiative In Northern Nigeria. Signed 13 June 2008.
7. Northern States Maternal, Newborn and Child Health Programme. CNTR 200808350. Part B – General and Technical Proposal. Submitted by Health Partners International in Joint Venture with Save the Children UK, and GRID Consulting Ltd.
8. 2009 Annual Report. PRRINN-MNCH.
9. PRRINN-MNCH Annual Review. Narrative Report. Carol Bradford and Maisha Strozier 24 March 2010. DFID Human Resource Development Center.
10. Project concept note. Reviving routine immunization, Nigeria. Mis code: 048-555-xxx. 06-09-04.
11. Joint logframe and Final M&E Framework. PRRINN-MNCH

Appendix Q: Norway - Tanzania Partnership Initiative Profile

Overview	Norway is investing 225 million NOK to improve maternal and child health in Tanzania. Through an agreement with the Government of Tanzania, the majority of the NTPI funding (80%) support the implementation of Tanzania's national health sector strategy through a pooled funding mechanism ("basket financing"). The additional funds will be used to support activities including a pilot project on pay for performance mechanisms, strengthening of the health management information system and support for NGO-provided MNCH services.				
PI duration:	<table border="1"> <thead> <tr> <th data-bbox="719 880 879 909">Start date</th> <th data-bbox="1114 880 1222 909">End date</th> </tr> </thead> <tbody> <tr> <td data-bbox="683 913 879 1070">Est. October 2007 Joint Statement between Got and Norway in February 2007</td> <td data-bbox="1102 913 1233 943">FY 2011/12</td> </tr> </tbody> </table>	Start date	End date	Est. October 2007 Joint Statement between Got and Norway in February 2007	FY 2011/12
Start date	End date				
Est. October 2007 Joint Statement between Got and Norway in February 2007	FY 2011/12				
PI purpose	<p><i>The goal of NTPI is to contribute to the implementation of the National Roadmap Strategic Plan to Accelerate the Reduction of Maternal and Newborn Mortality and Morbidity (2007), and the attainment for the MDGs related to maternal, newborn and child health in Tanzania.</i></p> <p>The <i>purpose</i> of NTPI is to provide additional flexible funding to district health services to support the implementation of interventions guided by the Roadmap, and to contribute to innovation and strengthened result focused through performance based financing approaches for reaching MDG-4 and MDG-5 in Tanzania.</p>				
Stated objective	<ul style="list-style-type: none"> • The objective of the basket fund contribution is to increase the amount of flexible funding available to contribute to the implementation of the Roadmap and other relevant child survival strategies at district level • The objectives of the P4P pilot are to: a) undertake a detailed examination of the current P4P design elements to identify strengths and weaknesses; b) Test variations of P4P design to explore good practices c) Document good practices for adoption in the National P4P program. • The objectives of the HMIS component are to improve the quality of data and completeness of reporting to support performance monitoring of MNC health service outputs and outcomes; improve timeliness of data transfer from lower levels to central offices; and improve mechanism for software and hardware trouble shooting 				

Funding:	Total: 225 million NOK										
	<table border="1"> <thead> <tr> <th>By partner:</th> <th>By activity area:</th> </tr> </thead> <tbody> <tr> <td>GoT</td> <td>1. “Basket health fund” channeled into three forms of support: per-capita payment to Districts; priority areas (contraceptives, drugs, vaccines etc.); and the Joint Rehabilitation Fund which targets district dispensaries/centres and provides short-term emergency funding for rehabilitation and maintenance (78%).</td> </tr> <tr> <td>Ifakara Health Research Institute</td> <td>2. A pay for performance (P4P) pilot project expected to start in January 2011 and run for 18 months. (6%)</td> </tr> <tr> <td>A consortium -- University of Dar es Salaam and Muhimbili College of Health Sciences (MUCHS) and Ifakara Health Research</td> <td>3. Support to strength the Health Management Information System (HMIS), in collaboration with other DPs (through MOHSW) (10%).</td> </tr> <tr> <td>Haydom Lutheran Hospital</td> <td>4. Support to Haydom Lutheran Hospital earmarked for maternal and newborn health (10%) Funds may be channeled directly by the Norwegian Embassy.</td> </tr> </tbody> </table>	By partner:	By activity area:	GoT	1. “Basket health fund” channeled into three forms of support: per-capita payment to Districts; priority areas (contraceptives, drugs, vaccines etc.); and the Joint Rehabilitation Fund which targets district dispensaries/centres and provides short-term emergency funding for rehabilitation and maintenance (78%).	Ifakara Health Research Institute	2. A pay for performance (P4P) pilot project expected to start in January 2011 and run for 18 months. (6%)	A consortium -- University of Dar es Salaam and Muhimbili College of Health Sciences (MUCHS) and Ifakara Health Research	3. Support to strength the Health Management Information System (HMIS), in collaboration with other DPs (through MOHSW) (10%).	Haydom Lutheran Hospital	4. Support to Haydom Lutheran Hospital earmarked for maternal and newborn health (10%) Funds may be channeled directly by the Norwegian Embassy.
By partner:	By activity area:										
GoT	1. “Basket health fund” channeled into three forms of support: per-capita payment to Districts; priority areas (contraceptives, drugs, vaccines etc.); and the Joint Rehabilitation Fund which targets district dispensaries/centres and provides short-term emergency funding for rehabilitation and maintenance (78%).										
Ifakara Health Research Institute	2. A pay for performance (P4P) pilot project expected to start in January 2011 and run for 18 months. (6%)										
A consortium -- University of Dar es Salaam and Muhimbili College of Health Sciences (MUCHS) and Ifakara Health Research	3. Support to strength the Health Management Information System (HMIS), in collaboration with other DPs (through MOHSW) (10%).										
Haydom Lutheran Hospital	4. Support to Haydom Lutheran Hospital earmarked for maternal and newborn health (10%) Funds may be channeled directly by the Norwegian Embassy.										
PI within landscape for MDGs 4&5 / development	<p>26 development partners contribute to the health SWAp with all funds aligned with MoHSW policy. A agreed sector plan is determined by a SWAp Technical Advisory Committee. A common set of indicators is used to monitor HSSPII, and six-monthly annual review meetings are conducted. Within the SWAp, eight development partners (World Bank, Irish Aid, SDC, GtZ/KfW, CIDA Canada, Netherlands, Danida and UNFPA – with UNICEF is shortly become the ninth partner) support the pooled “basket health fund” aimed at district-level support. The NTPI contribution raises the basket fund from \$0.75 per capita per year to approximately \$ 0.90 per capita.</p> <p>Other programs and partners work in support of MDG 4 and 5, including (but not limited to):</p> <ul style="list-style-type: none"> • GAVI support for health system includes funding for recruitment of about 1447 first level MCH staff, basic equipment for clinics, increasing the number of community mobilizers , CORPs and community health workers cadres, vehicles and supervision; • African Development Bank to finance a large health infrastructure development program and capacity building in 3 regions to support maternal and newborn health interventions. • EC through WHO to support scaling up Maternal and Newborn strategies in 4 regions • Health Metrics Network - HMIS Assessment and Strategic Framework 										

Geographic Areas and other partners working there:	The P4P pilot is slated for the Coast Region and will encompass 7 districts -- 4 intervention districts and 3 control districts.	
Criteria for selection	The Coast region was selected for the P4P, in part, because of an on-going District Health Information System pilot project.	
Partners description	Partners	Responsibilities
	MOHSW Clinton Health Access Initiative (CHAI) - lead partner within consortium for P4P pilot	MOHSW - overall responsibility to implement the components of the NTPI program and for P4P, will manage and co-ordinate with a consortium of domestic and intl. partners Clinton Health Access Initiative – lead partner and provide secretariat support including project management, administrative/ documentation support, training support and sub-contracting of technical expertise.
	Ifakara Health Research Institute	research elements of the P4P pilot including on-going process monitoring and impact evaluation.
	University of Dar es Salaam, Muhimbili College of Health Sciences (MUCHS) and Ifakara Health Research, University of Oslo (?)	support the MOHSW to implement the Health Management Information System (HMIS) strengthening activity
	Haydom Lutheran Hospital	MNCH service delivery
Institutional arrangements:	The MOHSW has overall responsibility to implement the components of the NTPI program. Other NTPI agreement partners are answerable and should report to the MOHSW and the Norwegian Embassy for the activities under this initiative.	
Implementation timeline	Planned All NPTI activities were have started in late 2007.	Actual While the contributions to the basket funding started on a timely basis, the P4P pilot has been delayed for a variety of reasons. The P4P pilot is now expected to start in January 2011.
Elements of Innovation (as described in PD)	The P4P pilot represents an innovation (and comparator to the national P4P program) and will be subject to on-going implementation-level monitoring as well as impact evaluation.	
Operations Research activities	Activity	Status
	None identified except the pilot P4P pilot.	

Appendix R: Norway - Tanzania Partnership Initiative Monitoring and Evaluation Plan

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
<i>Community indicators:</i>						
Proportion of women booking early for ANC	1					
Proportion of villages with community owned resource for MNCH (eg. CBD, IMCI)	1					
<i>Neonatal indicators:</i>						
Neonatal mortality rates	1					
Number of district hospitals that have a functional newborn resuscitation place in the delivery room	1					
TT2 coverage	1					
<i>Child indicators</i>						
Under five mortality rate	1					
Proportion of health facilities providing IMCI services	1					
Prevalence of underweight in under fives	1					

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
<i>Family Planning indicators:</i>						
New FP acceptors	1					
CPR (DHS)	1			DHS		
<i>Maternal Health indicators:</i>						
Maternal mortality ratio (DHS)	1			DHS		
Proportion of ANC (RCH) clinics offering routine core PMTCT services	1					
Proportion of deliveries taking place in a health facility.	1					
Proportion of births assisted by a skilled attendant.	1					
Proportion of facilities offering Basic EmOC services	1					
Proportion of facilities offering Comprehensive EmOC services.	1					
Postnatal care attendance rate	1					
<i>Increased political will and commitment indicators:</i>						

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
Proportion of funds allocated to MNCH and FP at different levels	1					
<i>Dispensary RBB targets</i>						
IPT 2+ doses	2	*	*		ANC register	
Deliveries	2	*	*		Book 2, Jedwali 41A	
OPVO	2	*	*		Monthly EPI returns	
Infant ITN vouchers	2	*	*		Voucher stubs	
HMIS returns	2	*	*		Councils & Regions will maintain a register of HMIS returns for every facility/council	
<i>Hospitals</i>			*			
As above, plus partographs properly completed	2	*	*		Partograph file & monthly tally	

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
<i>CHMTs</i>						
Aggregate performance of their council on the 5 indicators above plus complete HMIS quarterly reports for all facilities	2		70% of facilities meet targets on first 4 indicators. All hospitals meet partograph targets. All facilities meet HMIS targets			
<i>RHMTs</i>						
Aggregate performance of all facilities in a region; plus complete HMIS quarterly reports for all councils	2		70% of facilities meet targets on first 4 indicators. All hospitals meet partograph targets. All facilities meet HMIS targets			
<i>Facilities</i>						
DPT-H3 by skilled attendant	3					
Deliveries by skilled attendant	3					
MTUHA reporting	3					

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
<i>CHMTs and RHMTs</i>						
# and % Health facilities that qualify for P4P	3					
Overall DPT Hb3 for district/region	3					
Overall deliveries in health facilities in district/region	3					
% Health facilities visited for a supervisory visit each quarter	3					
Clean financial reports	3					
<i>Facilities</i>						
Immunization - DTPHb 3	4		≥ 80%		Vaccination registration Book 7	
Immunization - OPVO	4		≥ 60%		Vaccination registration Book 7	
Deliveries in health facilities	4		≥60%		Book 2, Jedwali 41A and book 12 delivery register	

Indicators (as included in PI documents)	PI Document	Baseline	Target	Sources of baseline data	M&E Data sources	Arrangements / Responsibility for data collection
IPT 2+ doses for pregnant women	4		≥60%		ANC register Book 6 - additional column added by hand	
Quarterly MTUHA report timely, complete and accurate	4		100%		F004 and F005 (Book 2)	
CHMTs						
Aggregate performance of council on facility indicators	4					
RHMTs						
Aggregate performance of region on facility indicators	4					

*Targets for facilities determined as follows:

- If baseline coverage of the intervention / indicator lies between 0% and 50%, target must require at least 10 percentage point improvement
- If baseline coverage lies between 50% and 75%, target must require at least 7 percentage point improvement.
- If baseline coverage is above 75%, target must require at least 5 percentage point improvement.

PI Documents:

1. Norway Tanzania Partnership Initiative Programme Document. MOHSW.
2. Results-Based Bonus: Design, Implementation & Budget (February 2008).
3. Payment for Performance Appraisal. Report to Norad and the Royal Norwegian Embassy Final Report (April 2009). Centre for Health and Social Development
4. Implementation Guidelines: Payment for Performance (December 2008). MOHSW.

Indicators identified for the payment scheme -- national program:

Facility / Team	Indicator and reward targets
Dispensaries, Health Centres and Hospitals	Immunization - DTPHb 3 equal or above 80%
	Immunization - OPV 0 equal or above 60%
	Deliveries in health facilities equal or above 60%
	IPT 2 for pregnant women equal or above 60%
	Quarterly HMIS report timely, complete and accurate 100%
CHMTs and co-opted members	Aggregate performance of council on facility indicators (above)
RHMTs and co-opted members	Aggregate performance of region on facility indicators (above)

Indicators for the pilot have not yet been concluded, but will most likely be the same as for the national program – except for the indicators for the CHMTs and RHMTs, which will be related to management performance and delinked from the performance of the facilities under their charge. There are some concerns about the inclusion of the IPT 2 indicator as this is not captured in the DHIS. Also, there is a discussion on whether the pilot should include additional indicators.

NB: Indicators to be used to trigger P4P payment to facilities are not the same as those which will be need to evaluate the effects of the program.

Appendix S: Norway - Tanzania Partnership Initiative Evaluation Options Appraisal

Appraisal Area 1: Map the basis for Norway's support to the PI using project documents³⁰ and assess the adequacy and quality of such documentation in terms of potential "impact evaluability" of the PI towards the end of the initial implementation period (typically 5-6 years).

Team's Appraisal:

The documentation available to date is very useful in tracking the progress of the P4P program in Tanzania over the past three years. It will be important to ensure that evaluators do not inadvertently use an older (and now obsolete) version of program documentation to use as a foundation of the current program. Clearly, there are some important lessons to be learned from this experience which should be succinctly compiled and shared. Perhaps as the P4P pilot rolls out, the opportunity will emerge for NTPI and MOHSW to reflect together on the experience in light of the pilot. These lessons may be valuable for other low-income countries seeking to embark on similar processes with an initiative that is "country-owned" yet coupled with a heightened international interest and attention.

As the pilot is prepared for launch, it will be important to create a new consolidated program document which leaves behind past machinations and focuses clearly and concisely on the parameters of the pilot. As with the other PIs reviewed, as part of this document the designers should prepare a single logical model which depicts the sequence of expected input, processes, outputs and outcomes. The logical model should be accompanied by an evaluation design document which, in a stepwise manner, describes and maps the types of data which will be required, the schedule and arrangements for data collection. Although reference was made to such a document, it was not included among those received. [Note: Team to confirm whether Ifkara has prepared an evaluation plan for the pilot or just a proposal.]

As with the Malawi P4P pilot, the Team would urge evaluation designers to be realistic and focused in their evaluation design and requirements for data collection. The timeline for the pilot has been compressed to just 18 months and the evaluation design will have to accommodate this truncated implementation schedule. Similarly, evaluation designers are strongly encouraged to work closely with the RBF for Health impact evaluation network (part of a Norwegian funded initiative based at the World Bank to support RBF innovation in eight countries). There is a readily

³⁰ Including: project design, results chain, indicators and success criteria, implementation arrangements, methods, approaches, types of activities, monitoring and evaluation systems, stakeholders, intended target groups and geographic coverage, and so on

available set of tools and materials to assist with the further development and refinement of the evaluation approach.

Some concern is raised with the following statements from the available documents: “process monitoring is a key feature it will take the form of an implementation science research project and employ scientifically rigorous empirical monitoring, documentation and analysis of key P4P+ design features and processes” Again, with an 18 month long pilot, the evaluation design should be streamlined and realistic.

Appraisal Area 2: Identify contextual factors likely to influence intended and unintended PI impacts. Then assess the availability, adequacy, and quality of contextual factor data/information currently available, going back at least 5 years in time where possible.

Team’s Appraisal:

The Coast region has been selected for the P4P pilot due to the parallel implementation of an HMIS intervention. While this is an opportunity for the P4P to benefit from stronger routine data, it is also potential confounder. The evaluation will need to be able to describe minimum requirements for HMIS operation during any proposed scale-up of the P4P pilot – with the cognition that not all districts will have benefitted from the HMIS activity. One possibility would be that facilities in the analyses could be grouped according to the level and quality of their use of the DHIS (District Health Information System). Planners have already considered roll-out to those regions (Mtwara and Lindi) where the newly revamped DHIS2 has been introduced with development partner assistance (CHAI).

As elements of the national P4P programme have been rolled out and implemented in an uneven fashion (with training completed in many regions/districts but little actual implementation), the evaluation design will need to account for this level of background “noise”⁷. Evaluators will also have to work closely with NTPI program managers to ensure a consistent set of messages about how to compare the pilot with the existing national P4P program. Steps should be taken to ensure that the evaluation of the pilot is seen as transparent and highly credible. National decision-makers will need to be brought along and confident in the findings of the evaluation before they can be expected to fully consider scale-up.

Appraisal Area 3: Review currently-available baseline data/studies and their appropriateness in terms of impact evaluation reliability and validity, keeping such questions as the following in mind:

- *What are the “right” impacts needing to be measured/verified?*
- *Are these “right” impacts actually verifiable to acceptable standards given the existing baseline information available?*
- *Where information is lacking or incomplete, can data be (re)constructed?*
- *What counterfactuals need to be considered?*

Team's Appraisal:

Work on evaluation design is expected to begin in parallel to intervention design activities in January 2011. A list of indicators was included in the Programme Document and raises some concerns with the Team. It should be noted that in compiling indicators across documents, the Team may have unintentionally combined two types of indicators: a) those proposed to evaluate the effectiveness of the P4P intervention; and b) those intended to serve as a trigger for performance payments to individual units. For the purposes of a) evaluating the effectiveness of the P4P intention, the Team, would simply urge that a minimum set of indicator be employed as the timeframe for implementation is quite short. Do not include indicators above the level of outcome (defined as population-based measures of use of services and behaviours). Specifically, the Team urges that indicators related to mortality measures (as appear above: neonatal mortality rate, under-five mortality rate and maternal mortality ratio). Even though program documents correctly point to the DHS as the source of such measures, the scale and timeframe for the NTPI simply do not allow for inclusion of such measures. These measures typically cover events that occur over a five or ten year period of time and are not generated for areas as small as the intervention area.

The Team does encourage the fullest possible use of existing data and triangulation methods as have been mentioned in the documentation available. The emphasis on "facilities readiness" is appropriate and should be given sufficient time to develop, test and garner buy-in from national decision-makers (as there are clear implications for the national P4P programme). The HESO appraisal document (#9 below) includes a number of specific and highly-relevant recommendations for monitoring of the pilot and potential pitfalls.

As of June 2010, a joint assessment recommended 6 additional indicators to be monitored by the P4P initiative. These include:

- The proportion of children under one year of age receiving the BCG;
- The proportion of children under one year of age receiving the measles vaccine;
- The percentage of pregnant women receiving antiretroviral drugs (ARVs) for prevention of mother to child transmission (PMTCT);
- The proportion of pregnant women receiving at least two does of tetanus toxoid (TT);
- The number of days Integrated Management of Childhood Illness (IMCI) tracer drugs are out of stock.

Appraisal Area 4: Summarize important findings and conclusions related to each of the above tasks.

Team's Appraisal

Recommendations:

The majority of the NTPI investment is directed through the pooled funding mechanisms and assessed according to jointly agreed procedures. However, as with the PI in Malawi, the NTPI is well-positioned to conduct a **global health – normative**

option around the pay-for-performance (P4P) component with important design elements in place at its inception.

As per above, the Team recommends:

- prepare a review of lessons learned over the period 2007-2010 as a guide to other low-income countries seeking to embark on similar processes with an initiative that is “country-owned” yet coupled with a heightened international interest and attention;
- create a new consolidated programme document -- and as part of this document prepare a single logical model which depicts the sequence of expected inputs, processes, outputs and outcomes, accompanied by an evaluation design document;
- be realistic and focused in evaluation design and requirements for data collection – as pilot’s timeline has been compressed to just 18 months, the evaluation design will have to accommodate this truncated implementation schedule;
- work closely with the RBF for Health impact evaluation network (part of a Norwegian funded initiative based at the World Bank to support RBF innovation in eight countries) and to extent possible, draw on the readily available tools and materials to assist with the further development and refinement of the evaluation approach;
- be able to describe minimum requirements for HMIS operation during any proposed scale-up of the P4P pilot – with the cognition that not all districts will have benefitted from the HMIS activity;
- evaluators to work closely with NTPI program managers to ensure consistent messages comparing the pilot with the existing national P4P program;
- ensure that the pilot’s evaluation is seen as transparent and highly credible and that national decision-makers are brought along;
- make fullest possible use of existing data and triangulation methods;
- “facilities readiness” measures need sufficient time to develop, test and garner buy-in from national decision-makers (as there are clear implications for the national P4P programme).

Documents reviewed:

1. Appraisal of the Norway Tanzania Partnership Initiative Programme. Report to the Norwegian Agency for Development Cooperation. John James, Wanjiku Kamau. August 2007.
2. Performance-Based Financing Report On Feasibility and Implementation Options Final Report. Paul Smithson, Dr Nelly Iteba, Oscar Mukasa, Dr Ali Mzige, Max Mapunda, Gradeline Minja. September 2007
3. Norway Tanzania Partnership Initiative. Programme Document. Submitted by the Ministry of Health and Social Welfare of Tanzania to the Royal Norwegian Embassy.
4. Memorandum of Understanding between the Partners (Government of Tanzania and Donors participating in the pooled funding(“basket financing”) of the health sector concerning the pooled funding for the Government of Tanzania’s Health Sector Programme based on Second Health Sector Strategic Plan and the Health Sector medium term expenditure Framework. July 2003; June 2008.

5. Results-Based Bonus Design, Implementation & Budget “Malipo kwa Ufanisi Bora katika Huduma za Afya” (MUBHA). Final Report. Report of the design team describing detailed model, implementation strategy and funding requirements. Paul Smithson, Dr Rena Eichler, Dr Samson Winani, Dr Aziz Msuya, Ingvar Theo Olsen, Erica Musch. Commissioned by Royal Embassy of Norway, Tanzania. 20th February 2008
6. The United Republic Of Tanzania. Ministry Of Health And Social Welfare. Agenda No. 5.1. Payment For Performance Strategy 2008-2015. December, 2008.
7. The United Republic Of Tanzania. Ministry Of Health And Social Welfare. Agenda No. 5.2 Implementation Guideline. Payment For Performance. December, 2008.
8. Memorandum of Understanding between the Partners (Government of Tanzania and Development Partners participating in the pooled funding (“basket”) of the health sector concerning the pooled funding for the Government of Tanzania’s Health Sector Programme based on Second and Third and the Health Sector medium term expenditure Framework based on the Ministry Of Health And Social Welfare and Prime Minister’s Regional Administration and Local Government and the Comprehensive Council Health Plans. July 1, 2008 to June 30, 2015.
9. Centre For Health And Social Development. Payment For Performance Appraisal. Report To Norad And The Royal Norwegian Embassy, Tanzania. Marilyn Lauglo and Goddfrey B. R. Swai. Final Report. 2 April 2009.
10. Programme document. Tanzania P4P for maternal and newborn health pilot. Undated. No author. Draft Programme document. Pilot to improve P4P for maternal and newborn health in Tanzania. Undated. No author.
11. Pay for Performance in Tanzania. Undated. No author.
12. Informing the Design of a Pay for Performance Initiative Pilot. Report prepared by Gregory Kabadi, Dominic Mosha, Flora Kessy and Josephine Borghi. June 2010.
13. A Review Of The Status Of HMIS Data Sources. By Gregory Kabadi and Dr Dominic Mosha, Ifakara Health Institute. Undated.
14. Paying for performance to improve the delivery of health interventions: an outline of good practices and lessons in relation to the design of the P4P pilot in Tanzania. Flora Kessy (PhD). Draft, 16th May 2010.
15. Learning from current P4P initiatives in Tanzania – An Assessment of possible linkages Report prepared for the Ministry of Health and Social Welfare and Norad. Josephine Borghi. May 2010.
16. Pay for Performance (P4P) in Tanzania – Status September 2010.

EVALUATION REPORTS

- 1.98 "Twinning for Development". Institutional Cooperation between Public Institutions in Norway and the South
- 2.98 Institutional Cooperation between Sokoine and Norwegian Agricultural Universities
- 3.98 Development through Institutions? Institutional Development Promoted by Norwegian Private Companies and Consulting Firms
- 4.98 Development through Institutions? Institutional Development Promoted by Norwegian Non-Governmental Organisations
- 5.98 Development through Institutions? Institutional Development in Norwegian Bilateral Assistance. Synthesis Report
- 6.98 Managing Good Fortune – Macroeconomic Management and the Role of Aid in Botswana
- 7.98 The World Bank and Poverty in Africa
- 8.98 Evaluation of the Norwegian Program for Indigenous Peoples
- 9.98 Evaluering av Informasjons støtten til RORGene
- 10.98 Strategy for Assistance to Children in Norwegian Development Cooperation
- 11.98 Norwegian Assistance to Countries in Conflict
- 12.98 Evaluation of the Development Cooperation between Norway and Nicaragua
- 13.98 UNICEF-komiteen i Norge
- 14.98 Relief Work in Complex Emergencies
- 1.99 WID/Gender Units and the Experience of Gender Mainstreaming in Multilateral Organisations
- 2.99 International Planned Parenthood Federation – Policy and Effectiveness at Country and Regional Levels
- 3.99 Evaluation of Norwegian Support to Psycho-Social Projects in Bosnia-Herzegovina and the Caucasus
- 4.99 Evaluation of the Tanzania-Norway Development Cooperation 1994–1997
- 5.99 Building African Consulting Capacity
- 6.99 Aid and Conditionality
- 7.99 Policies and Strategies for Poverty Reduction in Norwegian Development Aid
- 8.99 Aid Coordination and Aid Effectiveness
- 9.99 Evaluation of the United Nations Capital Development Fund (UNCDF)
- 10.99 Evaluation of AWEPA, The Association of European Parliamentarians for Africa, and AEI, The African European Institute
- 1.00 Review of Norwegian Health-related Development Cooperation 1988–1997
- 2.00 Norwegian Support to the Education Sector. Overview of Policies and Trends 1988–1998
- 3.00 The Project "Training for Peace in Southern Africa"
- 4.00 En kartlegging av erfaringer med norsk bistand gjennom frivillige organisasjoner 1987–1999
- 5.00 Evaluation of the NUFU programme
- 6.00 Making Government Smaller and More Efficient. The Botswana Case
- 7.00 Evaluation of the Norwegian Plan of Action for Nuclear Safety Priorities, Organisation, Implementation
- 8.00 Evaluation of the Norwegian Mixed Credits Programme
- 9.00 "Norwegians? Who needs Norwegians?" Explaining the Oslo Back Channel: Norway's Political Past in the Middle East
- 10.00 Taken for Granted? An Evaluation of Norway's Special Grant for the Environment
- 1.01 Evaluation of the Norwegian Human Rights Fund
- 2.01 Economic Impacts on the Least Developed Countries of the Elimination of Import Tariffs on their Products
- 3.01 Evaluation of the Public Support to the Norwegian NGOs Working in Nicaragua 1994–1999
- 3A.01 Evaluación del Apoyo Público a las ONGs Noruegas que Trabajan en Nicaragua 1994–1999
- 4.01 The International Monetary Fund and the World Bank Cooperation on Poverty Reduction
- 5.01 Evaluation of Development Co-operation between Bangladesh and Norway, 1995–2000
- 6.01 Can democratisation prevent conflicts? Lessons from sub-Saharan Africa
- 7.01 Reconciliation Among Young People in the Balkans An Evaluation of the Post Pessimist Network
- 1.02 Evaluation of the Norwegian Resource Bank for Democracy and Human Rights (NORDEM)
- 2.02 Evaluation of the International Humanitarian Assistance of the Norwegian Red Cross
- 3.02 Evaluation of ACOPAMA ILO program for "Cooperative and Organizational Support to Grassroots Initiatives" in Western Africa 1978 – 1999
- 3A.02 Évaluation du programme ACOPAMA Un programme du BIT sur l'« Appui associatif et coopératif aux initiatives de Développement à la Base » en Afrique de l'Ouest de 1978 à 1999
- 4.02 Legal Aid Against the Odds Evaluation of the Civil Rights Project (CRP) of the Norwegian Refugee Council in former Yugoslavia
- 1.03 Evaluation of the Norwegian Investment Fund for Developing Countries (Norfund)
- 2.03 Evaluation of the Norwegian Education Trust Fund for Africa in the World Bank
- 3.03 Evaluering av Bistandstorgets Evalueringsnettverk
- 1.04 Towards Strategic Framework for Peace-building: Getting Their Act Together. Overview Report of the Joint Utstein Study of the Peace-building.
- 2.04 Norwegian Peace-building policies: Lessons Learnt and Challenges Ahead
- 3.04 Evaluation of CESAR's activities in the Middle East Funded by Norway
- 4.04 Evaluering av ordningen med støtte gjennom paraplyorganisasjoner. Eksemplifisert ved støtte til Norsk Misjons Bistandsnemda og Atlas-alliansen
- 5.04 Study of the impact of the work of FORUT in Sri Lanka: Building Civil Society
- 6.04 Study of the impact of the work of Save the Children Norway in Ethiopia: Building Civil Society
- 1.05 –Study: Study of the impact of the work of FORUT in Sri Lanka and Save the Children Norway in Ethiopia: Building Civil Society
- 1.05 –Evaluation: Evaluation of the Norad Fellowship Programme
- 2.05 –Evaluation: Women Can Do It – an evaluation of the WCDI programme in the Western Balkans
- 3.05 Gender and Development – a review of evaluation report 1997–2004
- 4.05 Evaluation of the Framework Agreement between the Government of Norway and the United Nations Environment Programme (UNEP)
- 5.05 Evaluation of the "Strategy for Women and Gender Equality in Development Cooperation (1997–2005)"
- 1.06 Inter-Ministerial Cooperation. An Effective Model for Capacity Development?
- 2.06 Evaluation of Fredskorpset
- 1.06 – Synthesis Report: Lessons from Evaluations of Women and Gender Equality in Development Cooperation
- 1.07 Evaluation of the Norwegian Petroleum-Related Assistance
- 1.07 – Synteserapport: Humanitær innsats ved naturkatastrofer: En syntese av evalueringsfunn
- 1.07 – Study: The Norwegian International Effort against Female Genital Mutilation
- 2.07 Evaluation of Norwegian Power-related Assistance
- 2.07 – Study Development Cooperation through Norwegian NGOs in South America
- 3.07 Evaluation of the Effects of the using M-621 Cargo Trucks in Humanitarian Transport Operations
- 4.07 Evaluation of Norwegian Development Support to Zambia (1991 - 2005)
- 5.07 Evaluation of the Development Cooperation to Norwegian NGOs in Guatemala
- 1.08 Evaluation: Evaluation of the Norwegian Emergency Preparedness System (NOREPS)
- 1.08 Study: The challenge of Assessing Aid Impact: A review of Norwegian Evaluation Practise
- 1.08 Synthesis Study: On Best Practise and Innovative Approaches to Capacity Development in Low Income African Countries
- 2.08 Evaluation: Joint Evaluation of the Trust Fund for Environmentally and Socially Sustainable Development (TFESSD)
- 2.08 Synthesis Study: Cash Transfers Contributing to Social Protection: A Synthesis of Evaluation Findings
- 2.08 Study: Anti- Corruption Approaches. A Literature Review
- 3.08 Evaluation: Mid-term Evaluation the EEA Grants
- 4.08 Evaluation: Evaluation of Norwegian HIV/AIDS Responses
- 5.08 Evaluation: Evaluation of the Norwegian Research and Development Activities in Conflict Prevention and Peace-building
- 6.08 Evaluation: Evaluation of Norwegian Development Cooperation in the Fisheries Sector
- 1.09 Evaluation: Joint Evaluation of Nepal's Education for All 2004-2009 Sector Programme
- 1.09 Study Report: Global Aid Architecture and the Health Millennium Development Goals
- 2.09 Evaluation: Mid-Term Evaluation of the Joint Donor Team in Juba, Sudan
- 2.09 Study Report: A synthesis of Evaluations of Environment Assistance by Multilateral Organisations
- 3.09 Evaluation: Evaluation of Norwegian Development Cooperation through Norwegian Non-Governmental Organisations in Northern Uganda (2003-2007)
- 3.09 Study Report: Evaluation of Norwegian Business-related Assistance Sri Lanka Case Study
- 4.09 Evaluation: Evaluation of Norwegian Support to the Protection of Cultural Heritage
- 4.09 Study Report: Norwegian Environmental Action Plan
- 5.09 Evaluation: Evaluation of Norwegian Support to Peacebuilding in Haiti 1998-2008
- 6.09 Evaluation: Evaluation of the Humanitarian Mine Action Activities of Norwegian People's Aid
- 7.09 Evaluation: Evaluation of the Norwegian Programme for Development, Research and Education (NUFU) and of Norad's Programme for Master Studies (NOMA)
- 1.10 Evaluation: Evaluation of the Norwegian Centre for Democracy Support 2002–2009
- 2.10 Synthesis Study: Support to Legislatures
- 3.10 Synthesis Main Report: Evaluation of Norwegian Business-related Assistance
- 4.10 Study: Evaluation of Norwegian Business-related Assistance South Africa Case Study
- 5.10 Study: Evaluation of Norwegian Business-related Assistance Bangladesh Case Study
- 6.10 Study: Evaluation of Norwegian Business-related Assistance Uganda Case Study
- 7.10 Evaluation: Evaluation of Norwegian Development Cooperation with the Western Balkans
- 8.10 Evaluation: Evaluation of Transparency International

Norad

Norwegian Agency for
Development Cooperation

Postal address

P.O. Box 8034 Dep. NO-0030 OSLO

Visiting address

Ruseløkkveien 26, Oslo, Norway

Tel: +47 22 24 20 30

Fax: +47 22 24 20 31

No. of Copies: 200

postmottak@norad.no

www.norad.no

