END OF PROGRAM EVALUATION OF THE PERFORMANCE BASED FINANCING FOR HEALTH PROGRAM IN OROMIA AND AMHARA REGIONS, ETHIOPIA

(Final Report)

Submited to

Cordaid

Submited by

ICOS CONSULTING PLC.

Website: www.icosconsult.com
Email: info@icosconsult.com/ icos.consult@gmail.com
Tel: 2516394200/+251911208078
P.O.Box. 461058, Addis Ababa, Ethiopia

Addis Ababa, Ethiopia
April **2023**

Executive Summary

Introduction and Methodology

Cordaid, with funding from The Embassy of the Kingdom of the Netherlands (EKN) in Ethiopia, is implementing a 4-year Performance Based Financing (PBF) project in the Jimma and Borena Zones in Oromia Region, Ethiopia, and a 1-year PBF project in the North Achefer woreda in West Gojjam Zone in Amhara Region, Ethiopia. The PBF program aims to improve healthcare service delivery and quality, governance, and the health information system in selected areas. The program rewards healthcare facilities and management based on performance, with subsidies paid quarterly as a reward for facilities based on their performance. As the project implementation is in its final year, end of program evaluation of the PBF program was conducted to generate knowledge to support the Federal Ministry of Health in scaling up PBF in Ethiopia.

The evaluation focused on program management, activities, and partnerships to assess achievements, draw lessons learned, and provide information for future planning. It addressed relevance, effectiveness, efficiency, sustainability, and coherence of the program, assessing progress towards achieving outputs and outcomes, appropriateness of management processes, and efficiency of implementation. The evaluation also analyzed the likelihood of continuation and sustainability of program outcomes and benefits and compatibility with other projects and programs.

The evaluation used a quasi-experimental study design and was conducted based on standard evaluation criteria, and data was collected from both primary and secondary sources. The study employed a consultative and participatory approach and used mixed methods, including both qualitative and quantitative research approaches. The primary data were gathered through client exit and health worker interviews , key informant interviews, focus group discussions, and observations from patients and health workers, and a quality of care assessment was conducted at both intervention and non-intervention health facilities using the routine quality assessment checklist. Secondary data collection processes included charts audit to assess adherence of clinicians to treatment guidelines in addition to national Health Management Information System (NHMIS).

A total of 45 health facilities, including 28 in the intervention areas and 17 in the non-intervention areas, were included in the sample using probability proportional to sample size method. Sample charts were also audited to assess adherence of clinicians to treatment guidelines.

Qualitative data were gathered through interviews and focus group discussions, transcribed, and analyzed using OpenAI's GPT-3-powered writing assistant. Quantitative data were used to assess the impact of the program on various outputs and outcomes and were analyzed using descriptive and inferential statistics. Different models, including difference-in-differences models, were used to evaluate the impact of the program on health service utilization, quality of care, health worker motivation, patient satisfaction, and cost-effectiveness analysis.

Findings

The program has proven to be effective in improving quality of healthcare services. By aligning with the Ethiopian Health Sector Transformation Plan II (HSTP II) and addressing the needs of patients, health workers, and health authorities, the program has contributed to each of the five pillars of HSTP II. Key informants have generally provided positive feedback on the program's inputs and strategies, with some concerns regarding the adequacy of financial support and incentives provided.

The PBF program aligns with the Sustainable Development Goals (SDGs), particularly SDG 3. At the regional level, the African Union also supports PBF as a means to improve health services in African countries. In Ethiopia, the PBF program is consistent with national health policies, such as the Second Ethiopian Health Sector Transformation Plan (HSTP II) and the 1993 Health Policy. The PBF program complements the community-based health insurance (CBHI) scheme by providing financial support needed by health facilities to buy essential medicines, and equipment, ensuring quality care and accessibility for CBHI beneficiaries.

The PBF program has contributed to efforts to address gender disparities in access to and utilization of maternal and child health services, and improved quality of care. The program's focus on maternal health has specifically helped vulnerable groups, including mothers, to access and utilize healthcare services. The program's efforts to address shortages of essential medicines and laboratory equipment have had a positive impact on vulnerable communities, particularly women seeking healthcare services. However, challenges related to distance, transportation, and the quality of care remain, highlighting the need for continued efforts to promote gender-sensitive healthcare delivery.

The evaluation found that the program had a significant positive impact on ANC1 service utilization (188.7; P=0.017). However, the project did not have a significant impact on family planning long-term (504.9; P=0.138) and short-term (-28.62; P=0.339), ANC4 (63.61; P=0.469), PNC1 (P=1.359; 0.321), skilled deliveries (10.35; P=0.251), OPD Under-5 (-1937.5; P=0.270) and all project-supported services combined (-135.3; P=0.974). This may be partly because of reduction in false reports/utilization data from intervention health facilities due to verification, while control facilities have continued with their old way of reporting without being scrutinized for quality of their data. Trend analysis of data quality also found improvements in data reliability in intervention facilities overtime. It could also indicate that other factors, such as inadequate infrastructure or low demand for services, may have played a role in limiting the impact of the PBF program on service utilization. These findings may also suggest the need for additional interventions, such as improving infrastructure and addressing demand-side factors, to increase service utilization in some settings.

The evaluation found that intervention facilities have generally higher quality of care scores compared to control health facilities. Although mean quality of care scores on almost all domains (except nutrition domain for control facilities) grew statistically significantly ($p<0.05$) for both intervention and control health centers, the magnitude of difference is very large for intervention facilities (from 24 at baseline to 81 at endline) compared to control facilities whose score grew from 16 to 50. In the nutrition domain, average quality of care scores of the intervention facilities increased statistically significantly ($p<0.05$) compared to baseline, while this was not the case in the control facilities ($p>0.05$). The results suggest that technical quality of care improved more in the intervention health centers than in the control ones. A difference-in-difference analysis results also shows that the intervention health centers generally experienced greater increase in quality scores compared to the control group in most of the domains, indicating a positive impact of the intervention on technical quality of care. Similar results were also found for hospitals. Mean quality of care scores for the intervention hospitals increased significantly ($p<0.05$) in all domains (except laboratory) compared to control ones. In the intervention hospitals, the mean scores for all quality of care domains increased from 21.7 at baseline to 88.3 at endline. The control hospitals also showed some improvement, but not to the same extent as the intervention hospitals.

Patients from intervention health facilities received better quality antenatal care than the control group in many areas, including obstetric history, blood pressure measurement, weight, pallor/anaemia check, oedema check, diet advice, preparation for delivery, and tetanus injection. The intervention group also had better knowledge of danger signs during pregnancy and received

more counseling on family planning and counseling. The intervention group outperformed the control group in terms of recording important data, physical examination data, and clinical history documentation. Overall, the survey results suggest that the intervention group received better quality healthcare than the control group. Logistic regression results also showed that the intervention had a significant positive effect on patient satisfaction for all service types (ANC (AOR=2.325, p<0.01), PNC (AOR=2.758, p<0.01), OPD ((AOR=3.200, p<0.01), IMNCI Diarrhea ((AOR=1.710, p<0.01), IMNCI Pneumonia (AOR=2.751, p<0.01)), Patients expressed satisfaction with the quality of services they received, citing privacy, confidentiality, being listened to, getting the support/service they needed, and getting necessary medicine as positive experiences.

The mean number of weeks pregnant at the time of the visit was significantly (p<0.05) lower in the intervention group (16.23 weeks) than in the control group (20.96 weeks). The intervention group also had a higher percentage of health workers that received training on most areas, indicating that the intervention may have had a positive impact on the provision of training.

The job satisfaction of health workers in both intervention and control health facilities was evaluated, and the program was found to be successful in improving overall job satisfaction and satisfaction with work load as well as compensation and benefits. However, no statistically significant differences (p>0.05) were found between intervention and control health facilities' workers' satisfaction in most other domains such as working environment, performance, recognition, and others. The results may suggest that compensation and benefits may be powerful instruments for enhancing job satisfaction in the Ethiopian context. The motivation of health workers in intervention health facilities was also statistically significantly (p<0.05) higher than those from control facilities.

The PBF program in Ethiopia aimed to improve governance in health service delivery by increasing the capacity of WorHOs and ZHDsto perform regulatory tasks and provide supportive supervision. It also strengthened the capacity of health facility management, resulting in good leadership and commitment from health care providers.

The institutionalization of PBF in the Ethiopian health system is still in its early stages, but the program has been well-received by government and health facilities, and is considered as one of the strategies to address funding gaps in implementing the Second Health Sector Transformation Plan. The program is aligned with health sector priority indicators, and appears to have contributed to achieving health outcomes such as decreasing maternal and neonatal mortality. While many stakeholders have positive attitudes towards the program, doubts remain about the sustainability of the outcomes achieved once the program phases out.

Qualitative findings suggest that the PBF program significantly impacted the quality of health data, evidence-based decision making, and transparency. The program brought about changes in health workers' attitudes towards importance of data. Improved in data quality has also encouraged use of the data to inform decision making different levels.

The findings show that health centers have a substantially higher average cost per capita of ETB 48.35 compared to hospitals, which have a cost per capita of ETB 13.44, and the difference was statistically significant (p=0.000). The WHO recommendation on cost-effectiveness states that interventions with an Incremental Cost-Effectiveness Ratio (ICER) less than the Gross Domestic Product (GDP) per capita are considered cost-effective[1]. As the ICER of the project is significantly lower than the 2021 national GDP per capita of USD 925.1 (at current prices) estimated by the

[1] World Health Organization (WHO). (2001). Macroeconomics and health: Investing in health for economic development: Report of the Commission on Macroeconomics and Health. Geneva: World Health Organization. Retrieved from https://apps.who.int/iris/handle/10665/42435

World Bank[2] or 40,524[3] ETB reported by the Ethiopian Statistics Service, the interventions may be considered cost-effective.

The interaction variable in the DiD analysis shows that there was no statistically significant variation (35.62; P=0.872) in the cost per DALY averted between the intervention and control groups over time. This suggests that the intervention did not have a significant impact on the cost-effectiveness of the program, as measured by the cost per DALY averted, compared to the control group. However, it is important to interpret this finding in the context of the overall study design, methodology, and other relevant factors. The lack of statistical significance in the interaction variable may be due to a variety of reasons, including small sample size, measurement error, or confounding variables that were not accounted for in the analysis. Therefore, further analysis and interpretation may be required to fully understand the implications of this finding.
Incremental Cost-Effectiveness Ratio (ICER) analysis results indicate that it costs approximately ETB 967 and ETB 832, respectively, to avert one additional DALY in Jimma and Borena. Overall, it costs Birr 943 to avert one additional DALY.

The sustainability of the outcomes and benefits of a health program that used performance-based financing (PBF) is uncertain, with some stakeholders skeptical that the results can continue without the continued support of the PBF project. The continuation and sustainability of program benefits are dependent on commitment from government bodies and health facility leaders, availability of financial support, and capacity of the health facilities.

Recommendations

- While it increases quality of care, it faces challenges related to increasing volume of service utilization, sustainability, efficiency, and institutionalization. By focusing on these areas and building on the positive impacts of the program, it is possible to continue to improve the quality of care in Ethiopia. Additionally, further piloting and evaluation with a more complete data is important before scaling up the PBF program in Ethiopia. Moreover, it's important to use public fund to pay for performances and to integrate PBF into the regular provider payment mechanism (either via MoH/RHB) as well as the routine M&E system to enhance sustainability and efficiency.

- Strengthen health management information system to ensure that quality data is available for measuring effectiveness, impact and efficiency of PBF program, and support evidence-based decision making.

- Strengthen the relationship between the PBF program and CBHI scheme to further enhance the quality and accessibility of care for CBHI beneficiaries.

- Conduct a comprehensive review of the program design and implementation to identify the potential reasons for the lack of impact on health service utilization.

- Encourage and facilitate the sharing of best practices and lessons learned among health facilities and other stakeholders to promote continuous improvement in data quality and transparency.

[2]GDP per capita (current US$) – Ethiopia: Accessed from:
https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=ET

[3] Determined by multiplying the GDP per capita in current US dollars by the average exchange rate of 2021, which was 1 USD equals 44 ETB.

- Ensure that future PBF programs in Ethiopia gives due attention to health worker compensation and benefits to enhance their job satisfaction, while doing further analysis to identify the other domains that affect job satisfaction among health workers in the country.

- Extend the program's duration particularly in areas with shorter implementation periods, and support further pilot and scaling up of the program.

TABLE OF CONTENTS

List of Tables

List of Figures

Acronyms and Abbreviations

AIDS	Acquired Immuno Deficiency Syndrome
ANC	Ante Natal Care
BOFEC	Bureau of Finance and Economic Development
CAPI	Computer-assisted personal interviewing
CBHI	Community Based Health Insurance
CBO	Community Based Organization
CEA	Cost Effectiveness Analysis
COVID-19	Corona Virus Disease-19
CORDAID	Catholic Organization for Relief and Development Aid
DAC	Development Assistance Committee
DALYs	Disability Adjusted Life Years
DHIS2	District Health Information System
DiD	Difference in Difference
E.C.	Ethiopian Calendar
EFY	Ethiopian Fiscal Year
EKN	Embassy of the Kingdom of The Netherlands
EMONC	Emergency Obstetric And Neonatal Care
EPI	Expanded Program on Immunization
ETB	Ethiopian Birr
FGD	Focus Group Discussion
FMoH	Federal Ministry of Health
FP	Family Planning
G.C.	Gregorian (or European) Calendar
GDP	Gross Domestic Product
GM	Growth Monitoring
HC	Health Centre
HCF	Health Care Financing
HEP	Health Extension Program
HEW	Health Extension Worker
HF	Health Facility
HIV	Human Immunodeficiency Virus
HMIS	Health Management Information System
HP	Health Post
HR	Heart Rate
HRH	Human Resource for Health
HRM	Human Resource Management
HSTP	Health Sector Transformation Plan
ICER	Incremental Cost-Effectiveness Ratio
IMNCI	Integrated Management of Newborn and Childhood Illnesses
IT	Information Technology
IUCD	Intra-uterine Contraceptive Device
KII	Key Informant Interviews
MCH	Maternal and Child Health
MNH	Maternal & Newborn Health
KII	Key Informant Interview
MTR	Mid-Term Review
M&E	Monitoring and Evaluation
NGOs	Non-governmental Organizations
MD	Medical Doctor

MRN	Medical Record Number
NCD	Non-Communicable Disease
OECD	Organization for Economic Co-operation and Development
OPD	Outpatient Department
PBF	Performance Based Financing
PH	Primary Hospital
PHC	Primary Health Care
PHCU	Primary Health Care Unit
PICT	Provider Initiated Counselling and Testing
PIM	Project Implementation Manual
PMTCT	Prevention of Mother to Child Transmission
PNC	Postnatal Care
PTB+	Pulmonary Tuberculosis Positive
PPA	Performance Purchasing Agency
RHB	Regional Health Bureau
RMNCAYH	Reproductive, Maternal, Neonatal, Child, Adolescent, and Youth Health
RMNCH	Reproductive, Maternal, Neonatal and Child Health
SAM	Severe Acute Malnutrition
SC	Steering Committee
SDG	Sustainable Development Goal
SOP	Standard Operation Procedure
SPSS	Statistical Package for the Social Sciences
STD	Sexually Transmitted Disease
STI	Sexually Transmitted Infection
TA	Technical Assistance
TB	Tuberculosis
TOR	Terms of Reference
TT	Tetanus Toxoid
UHC	Universal Health Coverage
UN	United Nations
US$	United States Dollar
WFO	Woreda Finance Office
WorHO	Woreda Health Office
WHO	World Health Organization
ZHD	Zonal Health Department
ZOFED	Zonal Finance and Economic Development Department
&	Ampersand
%	Percent

Note:

- *Ethiopian Fiscal Year runs from July 7 to July 6*
- *Ethiopian Calendar runs from September 11 to September 10*
- *1 Euro = 57.28 (buying rate) Ethiopian Birr at end of February 2023*
- *1 USD = Ethiopian Birr (ETB) 54 (buying rate) mid April 2023.*
- *Terms such as end of program evaluation, endline evaluation, and end-term evaluation and impact assessment used alternatively in the report.*
- *Terms such as program and project are used alternatively in the report when referring to the PBF program.*

1. BACKGROUND

1.1. Introduction

Ethiopia is currently facing a triple burden of diseases that affects all age groups, with a disproportionately higher burden among children and women in their reproductive age. In 2019, 58% of disability-adjusted life years were due to maternal and neonatal conditions, communicable diseases, and malnutrition (GBD 2019 Diseases and Injuries Collaborators, 2020[4]). Ethiopia has made significant efforts to make high-quality essential health services available, accessible, acceptable, and affordable to the community. However, the decline in maternal and neonatal mortalities has been modest, and out-of-pocket spending on health remains high at 31% of Total Health Expenditure in 2016/17, with a significant proportion of households facing catastrophic health expenses (Berhanu et al., 2021[5]).

The country has made growing investments in expanding health services, infrastructure, and health workforce, and access to primary health coverage has potentially increased from 50.7% in 2000 to more than 90% in 2019. However, the universal health coverage (UHC) service coverage index remains at 43% (Berhanu et al., 2021)[6].

Health services are provided by a network of health facilities arranged in a three-tier health care delivery model: primary-, secondary-, and tertiary-level health care. The primary health care unit consists of health posts, health centers, and primary hospitals, with each level providing various services and serving as a referral center for the next level. Health posts and health centers are staffed by health extension workers, nurses, midwives, and other health professionals who provide various preventive and curative services, while primary hospitals offer inpatient and ambulatory services to about 100,000 people within their catchment areas, including emergency surgery. Secondary and tertiary healthcare consist of general hospitals and specialized hospitals that serve larger populations and provide referral services for lower-level health facilities.

Despite efforts to make essential health services available, accessible, and affordable, health service utilization remains low, especially among rural and socioeconomically deprived populations. To address this, Ethiopia has implemented a Health Extension Program (HEP) that deploys health extension workers to provide preventive and basic curative services to communities, supported by volunteer community-level workers. The government has also implemented policies to increase access to health services, including the Health Development Army and the Health Sector Transformation Plan, which aim to improve the quality of care and increase health service utilization.

The Reproductive, Maternal, Neonatal, Child, Adolescent, and Youth Health (RMNCAYH) services in Ethiopia have seen improvements in family planning, maternal and child health, and immunization. Family planning services are available in almost all public health facilities, but contraceptive prevalence rates are still below the target of 55% for 2020, and there is a high unmet need for family planning. Maternal, neonatal, and child health services coverage has increased, but quality of care remains a concern, with low coverage of postnatal care and functioning emergency obstetric and neonatal care (EMONC) facilities. Although under-five mortality rates have reduced,

[4] GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet. 2020 Oct 17;396(10258):1204-1222. doi: 10.1016/S0140-6736(20)30925-9. PMID: 33069326.
[5] Berhanu, D., Taddesse, N., & Taddesse, H. (2021). Ethiopia Health System Assessment 2020. Washington, DC: Health Finance & Governance Project, Abt Associates Inc. Retrieved from https://www.hfgproject.org/wp-content/uploads/2021/02/Ethiopia-Health-System-Assessment-2020.pdf
[6] Ibid

preventable childhood diseases still cause 189,000 under-five deaths every year, and neonatal disorders, diarrhoea, and lower respiratory infections are the leading causes of death. Immunization coverage has improved (WHO, 2021[7]} but with high regional variations (coverage of all basic vaccinations was highest in Addis Ababa (83%) and lowest in Afar (20% in 2019), and about 19% of children had no vaccination at all. (Ethiopian Public Health Institute [EPHI] and ICF, 2019[8])

Despite improvements in RMNCAYH services, there are still significant challenges, including low contraceptive prevalence rates, inadequate quality of care, high neonatal mortality rates, and low immunization coverage. The health system still faces difficulties in ensuring equitable access to services, with uneven distribution of health resources and sub-optimal quality of care. More efforts are needed to address these challenges to achieve the Sustainable Development Goal targets related to RMNCAYH.

The prevention and control of major diseases in Ethiopia face a number of challenges. While progress towards the 2020 90-90-90 targets for HIV has been achieved for the second and third 90s, suboptimal HIV case finding, especially in pediatric and adolescent age groups and in key and priority populations, remains a major challenge. Ethiopia is making progress in achieving the End TB Strategy targets, with a decline in TB incidence and improving TB case notification and treatment success rates. Malaria mortality and morbidity have also declined dramatically. Between 2015 and 2019, malaria deaths dropped from 3.6 to 0.3 per 100,000 among populations at risk, malaria case incidence dropped from 5.2 million in 2015 to under 1.6 million in 2019/20[9]. However, high-level resistance of malaria vectors to insecticides and sub-optimal usage of interventions by target communities are among the remaining challenges. Hepatitis prevention and control program in Ethiopia is under-resourced, which accentuates the challenges of access to diagnosis, treatment, and preventive measures, and most patients do not know their hepatitis B or C infection status.

Health Care Financing

There are various health care financing mechanisms available, including tax-based financing, public health insurance, private health insurance, and out-of-pocket payments. While traditional financing mechanisms such as Social Health Insurance, National Health Insurance, Tax-based system, and Out-of-pocket financing are still prevalent, innovative financing mechanisms have emerged as viable alternatives. One innovative financing mechanism that has gained popularity in recent years is performance-based financing (PBF). PBF involves providing financial incentives to health care providers based on their performance in achieving predetermined targets. PBF has been implemented in several low-income countries such as Rwanda, Burundi, and the Democratic Republic of Congo, where it has been successful in improving health service delivery and health outcomes (Basinga et al., 2011)[10]. PBF has also been praised for its efficiency in resource allocation and its potential to strengthen health systems (Renmans et al., 2017)[11]. However, the effectiveness of PBF has also been questioned, with some studies suggesting that it may lead to unintended

[7] World Health Organization. (2021). Ethiopia Immunization Profile. Retrieved from https://www.who.int/immunization/monitoring_surveillance/data/eth.pdf

[8] Ethiopian Public Health Institute (EPHI) [Ethiopia] and ICF. 2019. Ethiopia Mini Demographic and Health Survey 2019: Key Indicators. Rockville, Maryland, USA: EPHI and ICF.

[9] Federal Democratic Republic of Ethiopia Ministry of Health. (2021). Health Sector Transformation Plan (HSTP 2): 2020/21-2024/25. Addis Ababa, Ethiopia.

[10] Basinga, P., Gertler, P. J., Binagwaho, A., Soucat, A. L., Sturdy, J. R., & Vermeersch, C. M. (2011). Paying primary health care centers for performance in Rwanda. Policy Research Working Paper, (5688).

[11] Renmans, D., Holvoet, N., Orach, C. G., Criel, B., & Ssengooba, F. (2017). Opening the 'black box'of performance-based financing in low-and lower-middle-income countries: a review of the literature. Health Policy and Planning, 32(5), 690-708.

consequences such as focusing on incentivized services at the expense of other health care needs (WHO, 2011)[12].

The PBF Program for Health in Ethiopia

Currently, Cordaid with funding from The Embassy of the Kingdom of The Netherlands in Ethiopia is implementing a 4-year PBF project in the Jimma and Borena Zones in Oromia Region, Ethiopia, as well as a 1-year PBF project in the North Achefer woreda in West Gojjam Zone in Amhara Region, Ethiopia.

Before the inception of the Performance Based Financing (PBF) program, health care financing system in the Jimma and Borena Zones – as in the rest of Ethiopia – relied mainly on an input based approach, such as payment for salaries or drugs for primary health care facilities, irrespective of the results which are being achieved.

In 2015, Cordaid introduced PBF in the Ethiopian health sector, in the Borana zone (Oromia Region), while in 2019 a second PBF project started in the Jimma zone (also Oromia region). This expansion was made possible by the financial support of the Embassy of the Kingdom of The Netherlands (EKN) in Addis Ababa, which subsequently also took over the funding of the PBF intervention in Borana zone. From July 2021, the project covers 100% of the Health Facilities in Jimma and Borana. The project was also expanded in early 2022 to North Achefer woreda of West Gojjam zone in Amhara region to demonstrate to the different levels of the Ethiopian Ministry of Health how PBF can contribute to an arid lands context (Borana) as well as in a highland context (Jimma and West Gojjam) to realising the Ministry's objectives formulated in its HSTP-II.

The overarching objective of the PBF program is to improve availability and accessibility of good quality healthcare at primary and secondary level in Jimma and Borena Zones (Oromia Region, Ethiopia) and North Achefer Woreda (West Gojjam Zone, Amhara Region, Ethiopia), and a stronger health system at large, which supports the progressive realization of Universal Health Coverage (SDG target 3.8) in this geographical area". This is intended to be achieved through three interrelated outcome pathways:

1. Improved Health Service Delivery in selected woredas of the West Gojjam, Jimma and Borena Zones, reflected in:
    a. Increased utilization of good quality services
    b. Increased equity in access
2. Improved governance of health service delivery through:
    a. Increased capacity at the level of Woreda Health Offices (WorHO) and Zonal Health Departments (ZHDs) to perform their regulatory tasks and provide supportive supervision
    b. Institutionalisation of PBF in the Ethiopian health system
3. An enhanced health information system that supports:
    a. Data based decision making at Woreda, Zonal and Regional level
    b. Additional financing potential for the health system through enhanced transparency

The scope of the PBF program is as follows:

- Improve the utilisation, equity and quality of selected health care services for 3,644,414 inhabitants in Jimma, 780,966 inhabitants in Borena and 263,000 inhabitants in North Achefer, West Gojjam.
- Contract the West Gojjam, Jimma and Borena ZHDs

[12] Meessen, B., Soucat, A., Sekabaraga, C., & Mushagalusa, P. (2011). Performance-based financing: Just a donor fad or a catalyst towards comprehensive health-care reform? Bulletin of the World Health Organization, 89(2), 153-156. https://www.who.int/bulletin/volumes/89/2/11-088179/en/

- Focus on 21 Woredas in Jimma, 14 Woredas in Borena and 1 Woreda (North Achefer) in West Gojjam; and contract their WorHOs.
- Contract 121 Health Centres (HC) and 8 Hospitals in Jimma, 44 Health Centres (HC) and 5 Hospitals in Borena and 8 Health Centres (HC) and 1 Hospital in North Achefer, West Gojjam
- Contract 116 CBOs in Jimma, 44 CBOs in Borena and 9 CBOs in North Achefer, West Gojjam to perform quarterly community verification surveys for each Health Centre

Under PBF program, healthcare facilities and management are rewarded based on performance. Performance is measured through periodical quantity and quality verifications and patient satisfaction surveys, with a quarterly subsidy paid as a reward for facilities based on their performance. The facilities have the autonomy to spend the earned funds in line with their priorities set in their business plans. The program also aims to strengthen good quality of care to patients, especially those in remote areas through the application of equity bonuses, by stimulating community outreach, and liaising with health posts and health extension workers. The improved quality of care is indirectly also expected to lead to a higher job satisfaction level among health staff. The project also targets improving governance by incentivizing local governing bodies, such as WorHOs and ZHDs, to improve their performance while engaging higher levels of government to integrate PBF mechanisms into policies and strategies.

As the project implementation is in its final year, Cordaid commissioned ICOS Consulting PLC to conduct an independent end of program evaluation of the PBF for health program with the aim to generate knowledge that will support the Federal Ministry of Health in scaling up PBF in Ethiopia. The evaluation is also intended to assess program achievements and capture effectively lessons learned and provide information on the nature, extent and, where possible, the effect of the PBF program to the Government of Ethiopia.

The evaluation looked at program management, program activities, and partnerships with the government and other development partners. The purpose of the evaluation was to assess achievements and processes, draw lessons learned, and provide information to the government and Cordaid for future planning. The evaluation questions were framed around the relevance, effectiveness, efficiency, sustainability, and coherence of the program.

The evaluation assessed the progress made towards achieving the overall outputs and outcomes, as well as the relevance of the program to the needs of beneficiaries and stakeholders such as patients, health workers, and health authorities. It also assessed the management processes and their appropriateness in supporting delivery, including the effectiveness of monitoring and evaluation mechanisms and strategies and tools used in implementation. The evaluation analyzed the efficiency of program implementation, considering program costs, resources utilization, collaboration with the government, management, accountability structures, and financial management processes and procedures. It also examined the likelihood of continuation and sustainability of program outcomes and benefits after completion of the program, and the compatibility of the intervention with other projects and programs.

This report presents the findings of the end of program evaluation that took place from December 2022 to February 2023.

2. EVALUATION METHODOLOGY

2.1. Study Design

The end of program evaluation focused on the time period and health facilities supported through funding from the Embassy of the Kingdom of The Netherlands in Ethiopia (EKN). A quasi-experimental research design was employed to evaluate the impact of PBF program on utilization and quality of care of key health services. The evaluation was also based on the Organization for Economic Co-operation and Development (OECD) – Development Assistance Committee (DAC) evaluation criteria of relevance, coherence, effectiveness, efficiency, impact, and sustainability.

2.2. Sample Size Estimation

Sample woredas and facilities

At least 30% of the intervention woredas were selected for the evaluation. North Achefer was included in the sample as it was the only intervention woreda in West Gojjam zone. For Guji and Bedele control zones, at least 30% control woredas were also selected. For North Achefer woreda, South Achefer woreda was selected as a control woreda.

A total of 45 health facilities, including 28 in the intervention areas (22 PHCUs/HCs and 6 hospitals), representing over 15% of all intervention hospitals and health centers, were selected and visited for the evaluation. Additionally,17 control health facilities (14 PHCUs and 3 hospitals), representing over 22% of total control health facilities in all control zones, were assessed for the evaluation as planned as was also the case at midterm evaluation.

See Table 1.

Table 1: Planned and actual number of health facilities surveyed by phase intervention zone and type of health facility

Intervention Zone	Type of health facility	Planned						Actual					
		# of sample HFs in intervention areas by enrolment phase				# of sample HFs in non-intervention areas	Total sample HFs in intervention and non-intervention areas	# of sample HFs in intervention areas by enrolment phase				# of sample HFs in ccontrol areas	Total sample HFs in both areas
		Phase 1	Phase 2	Phase 3	Total			Phase 1	Phase 2	Phase 3	Total		
Borena	PHCU	1	2	3	6	3	9	1	2	3	6	3	9
	Hospital		1		1	1	2	-	1	-	1	1	2
	Sub-total	1	3	3	7	4	11	1	3	3	7	4	11
Jimma	PHCU	8	6		14	10	24	8	6	-	14	10	24
	Hospital	2	2		4	1	5	2	2	-	4	1	5
	Sub-total	10	8	0	18	11	29	10	8	-	18	11	29
North Achefer	PHCU	2	-	-	2	1	3	2	-	-	2	1	3
	Hospital	1	-	-	1	1	2	1	-	-	1	1	2

	Sub-total	3	0	0	3	2	5	3	-	-	3	2	5
Grand Total	PHCU	11	8	3	22	14	36	11	8	3	22	14	36
	Hospital	3	3	0	6	3	9	3	3	-	6	3	9
	Total	14	11	3	28	17	45	14	11	3	28	17	45

Sample size estimation for patient exit interviews:

Purposive sampling technique was utilized to select patients/clients for exit interviews from the clients available on the day of the interviews. As was the case at mid-term evaluation, 15 clients from intervention facilities and 10 clients from control facilities who consented to be interviewed were surveyed as they come and leave the sample hospitals and HCs. Similar to the mid-term evaluation, ANC, PNC, hypertensive patients, mothers/ guardians with under 5-year-old children suffering from diarrhea or pneumonia participated in the survey.

Sample Size and Sampling for FGDs

Convenience sampling was utilized for FGDs with patients/clients of sample intervention health facilities that visit the facilities for ANC or PNC services or treatment of their under-5 children suffering from pneumonia or diarrhea. Given the low client flow at health facilities and particularly at health centers as reported in the mid-term evaluation report, FGDs were conducted with male and female patients or parents/guardians that utilized the target services together. A total of 12 FGDs were conducted with patients (six FGDs in Jimma zone, four FGDs in Borena zone and 2 FGDs in North Achefer).

Sample size and Sampling for Health Workers

Three health workers were selected from each of the sample intervention and control health facilities to participate in the survey. The participants were composed of medical director/head representing the senior management; one randomly selected clinical department/case team head/coordinator from middle level management; and one patient-facing clinical staff from lower level. The medical director/head was included in the sample by default. To select a sample clinical department head from middle level management in each sample facility, we obtained list of heads of clinical departments stratified by department. Thereafter, a head of one of the clinical departments was selected using systematic sampling technique. Similarly, to select one lower level/patient-facing health worker, we obtained list of patient-facing health workers in all clinical departments and one of them was selected randomly using systematic sampling technique.

Sample Size for Chart Audit

Sample charts were selected from a list of patients/charts of patients that received normal delivery service, treatment for hypertension, admitted under-5 children that received treatment for pneumonia, ambulatory under-5 children treated for pneumonia, admitted under-5 children treated for diarrhea, and ambulatory under-5 children treated for diarrhea during the past one month. One chart was selected from each service category. If all the targeted services were available at a health facility, then a maximum of six charts would be selected in that facility (1 for normal delivery, 1 for hypertension, 1 for admitted under-5 children with pneumonia, 1 for ambulatory under-5 children with pneumonia, 1 for admitted under-5 children with diarrhea, and 1 for ambulatory under-5 children with diarrhea) for audit. To that end, the list of clients in the register were numbered from 1 $to\ n$ using unique identification numbers for random sampling. Based on these numbers, a random sample of clients from each available service category were selected using lottery method.

2.3. Sampling Strategy

In terms of sampling technique, the sample intervention health facilities were drawn from the list of all facilities that received support from the funding of the Embassy of The Kingdom of The Netherlands (EKN). Out of the 187 health centers and hospitals (14 hospitals and 173 health centers) that have been receiving support from the EKN –funded program, we sampled 28 health centers and six hospitals in the three intervention areas/ zones, which represents 15% of the total intervention hospitals and health centers. A total of 17 control health facilities were visited as was the case at midterm evaluation. The number of sample intervention and control health facilities to be visited is largely allocated among the intervention areas considering the number of supported health facilities in each zone.

Selection of woredas in Jimma and Borena started by stratifying the woredas by phase of enrollment into the program. Once woredas are stratified by phase, we then selected sample intervention woredas in each phase using systematic sampling technique. Woredas included in the sample but found to have security issues were replaced by others following the systematic sampling technique method. The sampled health facilities in each sample intervention woreda were selected randomly. North Achefer intervention being a woreda itself, the field team selected the required number of intervention health facilities allocated to the woreda randomly.

For Jimma and Borena zones, Bedele and Guji zones served as control zones. The list of woredas and health facilities under them in these control zones (Bedele and Guji) was collected from Cordaid Country Offices. The list excluded those woredas and facilities that had security issues or were not accessible. The evaluation team selected control woredas randomly using systematic sampling technique from each of the control zones. Once the sample control woredas were selected, we then sampled the required number of control health facilities randomly. For North Achefer intervention, South Achefer served as a control woreda. The required number of sample control health facilities in this control woreda was also selected randomly.

2.4. Study Method

Data for the end of program evaluation were gathered from a sample of intervention woredas and health facilities in all intervention zones namely: Borena, Jimma and West Gojjam. Additionally, data were gathered from a sample of control woredas, and health facilities. For Borena, data was gathered from a sample of Phase 1, 2 and 3 intervention health facilities in the zone, and a sample of control woredas and health facilities in neighboring Guji zone. For Jimma, a sample of phase 1&2 intervention woredas and health facilities in Jimma zone, and a sample of control woredas and health facilities were taken from neighboring Bedele zone. For North Achefer woreda, South Achefer woreda and a sample of health facilities in it served as a control woreda.

The evaluation team employed a consultative/participatory approach to ensure the active involvement of stakeholders at all levels in the end of program evaluation process. We utilized a mixed study method for the end of program evaluation of PBF program, which incorporates both qualitative and quantitative research approaches.

Data for the end of program evaluation was obtained from both primary and secondary sources. Desk review of secondary documents, including program documentations; Second Health Sector Transformation Plan (HSTP II) and progress reports; global, regional and national polices, strategies, instruments and declarations; publications of WHO, the World Bank, and many other organizations in relation to PBF programs and several others was conducted to inform the inception report and support the findings. Health service administrative data was sourced from the old and new District/Woreda Health Information Systems (DHIS2).

Primary data was gathered from a range of sources/stakeholders at all levels through questionnaires/ survey, key informant interviews, focus group discussions, and observations.

i. **Client Exit Interview**: Exit interviews were conducted with 597 patients (427 patients from intervention and 170 patients from control facilities) following their consultations for antenatal care (ANC) or postnatal care (PNC), or mothers/ guardians with under 5-year-old children suffering from diarrhea or pneumonia at each sample health facility (HF) in both intervention and non-intervention areas (see Table 2). The four sets of structured questionnaires that were developed for each type of target service category and utilized for the midterm evaluation were administered to a sample of these patients in order to assess their level of satisfaction with the services provided by the health center (HC) and hospital, travel time and expenditures, problems they encountered, and so on for the final evaluation. These questionnaires were translated into Amharic and Afan Oromo, and back-translated to English to ensure consistency of the translation/improve the quality of the translation. The questionnaires were pilot-tested in a sample of purposely selected intervention health facilities that are not selected for the main survey. The local language versions of these questionnaires (Afan Oromo and Amharic) were administered to sample clients by trained local language speaking enumerators. The surveys were administered using tablets. To that end, the English as well as the Afan Oromo and Amharic versions of the questionnaire were programmed on SurveyCTO.

ii. Health worker survey: Survey was administered to a sample of 133 health workers (84 health workers from intervention and 49 health workers from control health facilities) in sample health facilities in both intervention and control zones. The evaluation team surveyed three health personnel per sample health facility to capture their assessment of staff training, compensation, motivation, satisfaction, and knowledge (see Table 2). Additionally, interviews were conducted with 21 health extension workers (HEWs) (14 in Jimma, 6 in Borena and 1 in North Achefer) over the phone to capture their views about their use of the funds, and the relevance of the selected health services or indicators for health posts were for which they have been incentivized by the PBF program; and obtain their suggestions about how the program can improve and verify health post level indicators.

iii. Quality of care assessment checklist: The evaluation team conducted technical quality of care assessment at both intervention and non-intervention health facilities by utilizing the quality of care assessment tools that were adapted for the mid-term evaluation from the assessment tools that ZHDs and WorHO use to assess quality of care in hospitals and health centers, respectively. The quality of care assessment checklists were completed through review of documents, observations, interviews with relevant staff/officials of health facilities. The checklists measure both the conditions in which health services are produced and the quality of service itself. Technical quality of care of health centers and hospitals was assessed on 16 domains and 13 domains, respectively. The domains for health centers include General Appearance and Safety, Administration, Financial Management, Human Resource Management (HRM), and Planning, Health Management Information System and Supervision, Infection Control and Waste Management, General Out-Patient Department (OPD), Under 5 OPD, Emergency Services, Antenatal Care (ANC), Maternity Services, Expanded Program on Immunization (EPI) and Growth Monitoring (GM), Nutrition Services, Inpatient Services, Referral Services, Outreach and Health Post Supervision, Laboratory Services, and Logistics, Medicines, and Supplies. On the other hand, hospitals have fewer domains, including General Appearance and Safety, Administration, Financial Management, HRM, and Planning, Health Management Information System and Supervision, Infection Control and Waste Management, General Out-Patient Department (OPD), Emergency Services, Maternity Services, Expanded

Program on Immunization (EPI) and Growth Monitoring (GM), Inpatient Services, Referral Services, Outreach and Health Post Supervision, Laboratory Services, Logistics, Medicines, and Supplies, Surgical Services, and Radiological Services. Weight was assigned to each domain based on its importance in the overall evaluation of the facility.

iv. **Key Informant Interviews (KIIs):** Interviews were conducted with 71 key informants from diverse range of stakeholders, including the Federal Ministry of Health, donor, regional health bureaus, ZHDs, finance and economic development departments, WorHOs, woreda finance and economic development offices, hospitals and health centers in intervention woredas, Cordaid global and country offices, as well as Cordaid field office staff and Performance Purchasing Agency (PPA) in Jimma, Borena, and West Gojjam zones. Semi-structured interview guides were designed for each category of key informants, and interviews were recorded with the consent of respondents. The list of key informants and their organizations along with the number of key informants interviewed at all levels is presented in Appendices 3 and 4.

v. **Focus Group Discussions (FGDs):** Twelve Focus Group Discussions (FGDs) were conducted with patients (or caretaker's) (mostly from Reproductive, Maternal, Newborn, Child and Adolescent Health clients) at sample health facilities at time of the facility visit to assess their perception of quality of care, availability of medicines, and satisfaction. The focus group discussants were selected purposively and in consultation with staff of health facilities. The FGD guides that were developed for the midterm evaluation were utilized for the end of program evaluation. FGDs were voice recorded with the consent of participants.

vi. **Chart audit**: Chart audits were conducted for hypertension admissions, ambulatory Diarrhoea, ambulatory pneumonia, normal deliveries, pneumonia admissions, and Diarrhoea admissions. The total number of audits conducted across all regions was also higher in the intervention group, with 127 audits conducted compared to 106 in the control group. The highest number of audits were conducted for hypertension admissions, with a total of 55 chart audits (30 in intervention and 25 in control). The lowest number of audits were conducted for diarrhea and pneumonia admissions.

Table 2: Number of patients and health workers that participated in the survey by zone, intervention status and type of health facility

Type of Survey and Health Facility		**Grand Total**						**Jimma**				**Borena**				**West Gojjam**			
		Intervention		**Control**		**Total**		**Intervention**		**Control**		**Intervention**		**Control**		**Intervention**		**Control**	
		Target	**Actual**	**Target**	**Actual**	**Target**	**Actual**	**Target**	**Actual**	**Target**	**Actual**	**Target**	**Actual**	**Target**	**Actual**	**Target**	**Actual**	**Target**	**Actual**
Patients Survey	Hospital	90	91	30	30	120	121	60	61	10	10	15	15	10	10	15	15	10	10
	Health Centre	330	336	140	140	470	476	210	215	100	100	90	90	30	30	30	31	10	10
	Total	420	427	170	170	590	597	270	276	110	110	105	105	40	40	45	46	20	20
Health Worker Survey	Hospital	18	18	9	8	27	26	12	12	3	2	3	3	3	3	3	3	3	3
	Health Centre	66	66	42	41	108	107	42	42	30	29	18	18	9	9	6	6	3	3
	Total	84	84	51	49	135	133	54	54	33	31	21	21	12	12	9	9	6	6

Table 3: Number of Chart Audits conducted at sample health facilities by zone, service type and intervention status

Type of Service	Total				Jimma				Borena				West Gojjam			
	Intervention		Control		Intervention		Control		Intervention		Control		Intervention		Control	
	Target	Actual	Target	Actual	Target	Actual	Target	Actual	Target	Actual	Target	Actual	Target	Actual	Target	Actual
Ambulatory Diarrhoea	28	29	17	25	18	18	11	17	7	6	4	4	3	5	2	4
Ambulatory Pneumonia	28	30	17	24	18	19	11	16	7	6	4	4	3	5	2	4
Deliveries	28	27	17	25	18	16	11	17	7	6	4	4	3	5	2	4
Diarrhoea Admissions	28	5	17	3	18	3	11	1	7	1	4	1	3	1	2	1
Pneumonia Admissions	28	6	17	4	18	4	11	1	7	1	4	1	3	1	2	2
Hypertension Admissions	28	30	17	25	18	19	11	17	7	6	4	4	3	5	2	4
Total	168	127	102	106	18	79	66	69	42	26	24	18	18	22	12	19

2.5. Data Management and Analysis

Qualitative Data Preparation and Analysis:

Qualitative data gathered through interviews and focus group discussions were transcribed and then translated into English on Microsoft Word. The transcripts were then entered into Excel file by categorizing responses to each question. We then uploaded the categorized transcripts on Google Docs and used OpenAI's GPT-3-powered writing assistant "AI Writer" that is available as a Google Docs add-on to analyze and categorize qualitative data by identifying themes and patterns, and generate summaries. Thematic analysis and summaries were generated for each question, which in turn were thematized under each evaluation criteria.

Quantitative Data Preparation and Analysis

Quantitative data obtained from primary and administrative data sources were used to assess the impact of the program on a range of outputs and outcomes, such as the utilization of health services, quality of healthcare services, patient satisfaction, health worker motivation, and efficiency.

Descriptive statistics such as frequencies and means were calculated for each dataset. Different graphs were also utilized to present some results/ findings. Whenever possible and appropriate, results are presented comparing intervention and control health facilities. Comparisons were also made between baseline and endline when baseline data was available.

Inferential statistics were calculated to determine the impact or effect of the PBF program on health service utilization, quality of care, health worker motivation and patient satisfaction, and cost effectiveness analysis. Statistical significance level or p-value 0.05 (5%) is used in analysis unless other significant levels are specifically mentioned. This means that if the calculated p-value is less than or equal to 0.05, then we conclude that the results are statistically significant. On the other hand, if the calculated p-value is greater than 0.05, then we conclude that the results are not statistically significant.

Details on quantitative data analysis methodologies utilized are presented under each survey below.

i **Quantity/Utilization of health services**: Quantity of service utilization was calculated focusing on some key indicators/ services targeted by the PBF program. The quantity of service utilization data on intervention health facilities was provided by Cordaid. Utilization data from control health facilities was gathered from WorHO, ZHDs or Regional Health Bureaus (RHB). Quantity of service utilization data was annualized for analysis. Based on the available data, the evaluation team annualized the data using the period September 2018 to August 2019 as baseline and September 2021 to August 2022 as endline values. Health facilities, and indicators or services for which the data was unavailable or nearly unavailable were excluded in determining impact of program on quantity of service utilization. These include Newborn PMTCT Option B, management of newborn HEI, SAM Under-5 children, STIs Treated, Hospital Bed days, Referrals, and Post Abortion Care. Others such as Immunization 1 year, growth monitoring, Vitamin A suppl 5 years, Testing for HIVAIDS, Malaria Dx Tx, Diabetic Patients Tx, Hypertensive Patients Tx, OPD Adults, Microscopy TB Dx, TB Treated Cured were excluded from the analysis as Cordaid wanted the analysis to focus on key selected services namely: Family Planning (FP) long-term and short-term, ANC 1, ANC 4, skilled deliveries and OPD Under-5. The evaluation team used disability-adjusted life year (DALYs) to standardize measurement of

quantity of service utilization across services. The service specific total DALYs averted are calculated by multiplying the DALY conversion rate by the quantity of service provided by the facilities. The DALY conversion rate is a measure used to quantify the impact of a particular health intervention or service on disease burden. It represents the number of disability-adjusted life years (DALYs) averted for every unit of health service utilization; we adopted the conversion rate utilized by the WB for similar assignment[13]. We utilized difference-in-difference (DiD) model to assess the net effect of the project on service utilization while controlling for confounding factors, including catchment population, facility type, and pastoral or agrarian context. (see Equation (1)).

$$Total\ DALY = \beta_0 + \beta_1[Health\ facility\ type] + \beta_2[Interaction\ (Time * Intervention)] + +\beta_4[\text{Catchment}] + \varepsilon \ldots\ldots\ldots (Eq\ 1)$$

Where; $Time * Intervention = \begin{cases} 1\ if\ intervention\ at\ the\ endline \\ else\ 0 \end{cases}$

Livelihood = The main livelihood of the community utilizing the facility The confounding factors are selected based on the theoretical literature and the bivariate analysis conducted to see the effect of the confounders on the outcome variable.

ii **Quantity data reliability:** The PBF program utilized verified HMIS/DHIS2 data on health service utilization as a primary source of data to determine payments to health facilities or providers, among others. Payment to health facilities on each performance indicator has been made to health facilities as long as the difference between declared and verified data did not exceed the contractually agreed error margin/data reliability. Data reliability was calculated for quantity indicators for each intervention health facility for each quarter using the formula presented below.

$$\boldsymbol{Data\ reliability = (validated/verified\ data\ /\ Declared\ data) * 100}\ \ \textbf{(Eq 2)}$$

iii **Technical quality of care assessment**: Technical quality of care at health centers was assessed on 107 indicators identified under 16 domains, while quality of care at hospitals was assessed on 116 indicators identified under 13 domains. Each domain was assigned a maximum weight, which is the maximum possible value that a health facility could acquire on each domain. Total technical quality of care score for a facility was determined by dividing total actual scores acquired by the facility on all domains (16 domains for HC and 13 domains for hospital) by the total possible score for all domains. The evaluation team took baseline technical quality of care data captured on facilities that also included those that later became project beneficiaries.

DiD model formulated in Equation 3 below was used to evaluate the impact of the project on quality of care.

$$Y_{qlty} = \beta_0 + \beta_1[Time] + \beta_2[Intervention] + \beta_3[Time * Intervention] + \beta_4[Covariates] + \varepsilon \ldots (Eq3)$$

Where; $Time = \begin{cases} 1\ if\ Endline \\ 0\ if\ Baseline \end{cases}$

$$Intervention = \begin{cases} 1\ if\ intervention \\ 0\ if\ control \end{cases}$$

[13] https://documents1.worldbank.org/curated/en/238291593572868686/text/Cost-Effectiveness-Analysis-of-Performance-Based-Financing-for-the-Delivery-of-a-Health-Benefits-Package-in-The-Republic-of-Congo-HRBF-Impact-Evaluation.txt

Chart audit index/ Adherence index: From the chart audit data for Diarrhoea, pneumonia, normal deliveries and hypertension for both hospital and HCs, charts clinical completeness was checked for all the services to determine if clinicians adhered to standard clinical practices. Charts were reviewed on a case-by-case basis for all applicable procedures based on the condition of the patient on admission. All procedures not documented were taken as not provided to patients. For each facility that received chart audit for each category of service, applicable and acquired points were first calculated. Then adherence score was calculated for each assessed facility on each service by dividing points acquired by the facility by the total applicable points taking into consideration procedures not applicable for each patient whose chart is reviewed. The adherence index for a service was calculated by averaging the adherence scores across all facilities that received chart audit for that particular service. Difference-in-difference analysis was done to assess the impact of the project on adherence of intervention health facilities to standard practices. The DiD model formulated in Equation 4 below was used to evaluate the impact of the project on adherence..

$$Y_{Adherenceindex} = \beta_0 + \beta_1[Time] + \beta_2[Intervention] + \beta_3[Time * Intervention] + \beta_4[Covariates] + \varepsilon \ldots (Eq4)$$

Where;
$$Time = \begin{cases} 1\ if\ Endline \\ 0\ if\ Baseline \end{cases}$$
$$Intervention = \begin{cases} 1\ if\ intervention \\ 0\ if\ control \end{cases}$$

iv **Patients satisfaction survey:** Patients' overall satisfaction with the service, and with the security and trust of health care providers was assessed on a four-point Likert type scale with 17 items, seven items, respectively (1: Strong disagreement, 2: slight disagreement, 3: slight agreement, 4: strong agreement). Patient satisfaction survey data gathered for the PBF program's mid-term evaluation was merged with the endline patient survey dataset to make difference-in-difference analysis. To assess the impact of the program on patients' satisfaction, we first constructed composite quality of care measure by summing the quality measures for the five services, namely ANC, PNC, hypertensive patients, mothers/ guardians with under 5-year-old children suffering from diarrhea or pneumonia. We then determined average and grouped the responses into two categories: "Dissatisfied" and "Satisfied". We used the Equation 5 to assess the impact of the program on patients' satisfaction.

$$Patients's\ satisfaction = \beta_0 + \beta_1[Sex] + \beta_2[Interaction\ (Time * Intervention)] + +\varepsilon \ldots\ldots\ldots\ldots\ldots\ldots (Eq5)$$

Where;
$$Time * Intervention = \begin{cases} 1\ if\ intervention\ at\ the\ endline \\ else\ 0 \end{cases}$$

v **Health worker motivation and satisfaction:** We also assessed the impact of the program on workers' motivation and job satisfaction. To accomplish this, we utilized a 5-point Likert scale to assess participants' satisfaction levels across nine domains proposed for measuring worker satisfaction: nature of job and responsibilities, workload, compensation and benefits, organizational practices and functioning, working environment, career development and job security, performance, recognition, and overall well-being. To prepare the data for analysis, we merged the endline health worker survey data with the same gathered as part of the midterm evaluation of the PBF program as provided by Cordaid. For each aspect of their work, the survey participants were asked to rate their level of agreement or disagreement with a series of statements using a Likert scale that ranged from "Strongly Agree" to "Strongly Disagree." Workers' responses to statements were measured from 1 to 5 where 1 was "strongly agree" and 5 was "strongly disagree". Prior to the analysis, we reversed statements written in negative form. We also reversed the responses to each of the statements to make sure that higher values such as

"5" represent strong agreements while lower values such as "1" represent strong disagreements.

We computed a composite variable to obtain an overall satisfaction score for each of the nine domains and used it for analysis. Additionally, we included a question that asked workers to rate their overall job satisfaction on a scale of 1 to 5. We computed the overall satisfaction score by summing the results of all participants and comparing it to the general average computed for the entire population. Participants with a rating above the general average were classified as "satisfied," while those with a rating below were labeled as "unsatisfied." Equation 6 below was used to analyze the impact of the program on patients' and health workers' satisfaction.

$$Worker'satisfaction = \beta_0 + \beta_1[Sex] + \beta_2[Interaction\ (Time * Intervention)] + \beta_3[\text{Marital status}] + \varepsilon \ldots\ldots\ldots\ldots\ldots\ldots \text{(Eq6)}$$

Where; $Time * Intervention = \begin{cases} 1\ if\ intervention\ at\ the\ endline \\ else\ 0 \end{cases}$

vi **Cost-Effectiveness/Efficiency analysis methods**

The first step in the cost-effectiveness analysis (CEA) was to collect and allocate program financial costs to health facilities considering their time of entry. Financial costs refer to the actual expenditures made, such as payments made to facilities, regulators, overall program administration and management of the program. Additionally, expenditure data of health facilities from their regular budget was included in the dataset. The cost per disability-adjusted life year (DALY) averted was used to evaluate efficiency of the program. Per Unit Delivered (DALYs) compiled by the World Bank[14] and used in Congo was utilized for cost effectiveness analysis as the evaluation team did not find service-disaggregated DALYs on Ethiopia (See Appendix 5). DALYs averted was used to assess the project's effectiveness based on the DALYs averted from an increase in the volume of services ascribed to the PBF program.

Two different difference-in-difference models presented below were used. In the first model, we utilized the DID model controlling for confounding factors, including catchment population, facility type, and pastoral or agrarian context. (See Equation (7)). We also controlled for the quality of care (see Equation (8)) to test if improved quality of services has contributed to the cost per DALY.

$$Cost\ per\ DALY = \beta_0 + \beta_1[Health\ facility\ type] + \beta_2[Interaction\ (Time * Intervention)] + \beta_3[\text{pastoralists}] + \beta_4[\text{Catchment}] + \varepsilon \ldots\ldots\ldots \text{(Eq7)}$$

$$Cost\ per\ DALY = \beta_0 + \beta_1[Health\ facility\ type] + \beta_2[Interaction\ (Time * Intervention)] + \beta_4[\text{Catchment}] + \varepsilon \ldots\ldots\ldots \text{(Eq8)}$$

Where; $Time = \begin{cases} 1\ if\ Endline \\ 0\ if\ Baseline \end{cases}$

$Intervention = \begin{cases} 1\ if\ intervention \\ 0\ if\ control \end{cases}$

Apart from the DID the evaluation team utilized incremental cost-effectiveness ratio (ICER) to assess cost effectiveness of the services. The process involves estimating the incremental costs of

[14] https://documents1.worldbank.org/curated/en/238291593572868686/text/Cost-Effectiveness-Analysis-of-Performance-Based-Financing-for-the-Delivery-of-a-Health-Benefits-Package-in-The-Republic-of-Congo-HRBF-Impact-Evaluation.txt

the PBF program by adding the PBF cost and other expenses from the facilities budget. The effectiveness of the program is estimated by the change in utilization of health services in facilities where the PBF program is implemented. For this specific purpose we have utilized DALYs averted as a measure of utilization. The ICER is then calculated by dividing the incremental cost of the PBF program per DALY averted by the standard of care.

$$ICER = \frac{Difference\ in\ cost}{Difference\ in\ outcomes} \ (Eq\ 9)$$

The final output of the CEA is the ICER, which compares the national GDP per capita in which the facilities are found.

2.6. Ethical Issues

We implemented standard research ethical practices to protect participants' rights, privacy and confidentiality of data. We have obtained approval for data collection from regional, zonal and woreda government authorities. Respondents were requested to provide informed oral consent to participate in the study. The rights of participants, their privacy and confidentiality of individual data were protected during and after the field data collection.

2.7. Limitations of the Evaluation

The end of program evaluation has several limitations. Firstly, the number of sample health facilities and health workers surveyed is small, which can limit the generalizability and reliability of evaluation findings. In other words, the small sample sizes may not accurately represent the larger population from which they are drawn. Small sample sizes can also lead to reduced statistical power, which makes it more difficult to detect significant differences between groups or to find significant relationships between variables. The small sample sizes can also make it difficult to control for confounding variables, which can affect the relationship between the independent and dependent variables. Furthermore, the small sample sizes can lead to imprecise estimates of population parameters, such as means and standard deviations. This can limit the accuracy of evaluation findings and make it difficult to draw valid conclusions. Hence, it is important for readers/users of this evaluation report to carefully consider these limitations when interpreting results.

Secondly, the evaluation employed quasi-experimental program/project evaluation design, which has some limitations that should be considered when interpreting the findings. One of the limitations in this regard arises from lack of randomization as there was no random assignment to allocate participants to intervention and control groups. This can introduce bias into the evaluation, making it difficult to determine whether observed differences between groups are due to the intervention or other factors. The quasi-experimental design may not also be able to establish causality as convincingly because they cannot rule out alternative explanations for the observed outcomes. The results may be less generalizable, and causality cannot be fully established due to the lack of randomization. Quasi-experimental designs can suffer from selection bias if the intervention and control groups are not equivalent at baseline which is also the case for this evaluation study. But it is important to note here that although it is not the gold standard for establishing causality as random control trial (RTC) design, quasi-experimental design is more flexible and realistic implementation in real-world settings than RTC. It also allows the estimation of causal effects compared to simple pre-post comparisons.

Thirdly, inter-rater reliability could be a potential issue as different people assessed quality of care at health facilities at baseline and endline.

Fourthly, the enrollment of health facilities at different times in the program could also affect the generalizability of the evaluation study. As the health facilities were enrolled at different times, it is possible that other factors that affect health outcomes may also change over time. This can make it difficult to determine whether changes in desired project outcomes are due to the PBF program or other factors that are changing over time. To mitigate this, the evaluation team used same baseline year across sample health facilities regardless of their time of enrollment in the program.

Fifthly, there were major differences in health service utilization data obtained from the DHIS2 database and the data we obtained from WorHO/ZHDs/RHBs, which might have seriously affected the quality of analysis done to evaluate the impact of the PBF program on health service utilization and cost effectiveness analysis. Similarly, financial (revenue and expenditure) data gathered from the health facilities involved a lot of missing values. In some situations, facilities provided budget data but not expenditure data. The evaluation team had to include the budget data as expenditure data for the cost-effectiveness analysis. Those facilities that provided incomplete budget and expenditure data were excluded from the cost-effectiveness analysis. Additionally, the analysis was conducted based on declared data assuming that such data would be comparable for both intervention and control groups. However, it is important to note that the declared data reported by intervention facilities may not be comparable to those reported by control facilities due to the verification activities that were carried out to discourage false reporting/improve data quality in intervention facilities.

Finally, security and inaccessibility also posed challenges to the end of program evaluation, as the evaluation team had to replace some sample woredas and health facilities due to security concerns.

3. RESULTS/FINDINGS

In this section, the evaluation findings are presented with a focus on the desired outcome, outputs, and partnership strategy. The findings are presented according to the DAC evaluation criteria of Relevance, Coherence, Effectiveness and Impact, Efficiency, and Sustainability. The section starts with description of background characteristics of survey participants.

3.1. Background Characteristics of Survey Participants

Characteristics of sampled health facilities

Table 4 provides information on the mean number of beds, staffing rates, and catchment population of sampled health facilities.

Looking at the health centers, it can be observed that the intervention health centers have more beds than the control health centers in all three zones. However, the difference in the number of beds is not very large except in Jimma where intervention facilities are more likely to have more beds compared to control health centers. The average staffing rate (compared to norm) is also higher in the intervention health centers compared to the control health centers in all zones except for Borena. However, the difference in staffing rates between intervention and control health centers is not statistically significant ($p>0.05$). The average catchment population that is being served by the health centers is also higher in the intervention health centers in Jimma compared to control facilities, while the reverse is true in Borena, and the difference is statistically significant ($p<0.05$) in all zones except in West Gojjam zones.

In the case of hospitals, the staffing rate is generally higher in the intervention hospitals compared to the control hospitals, except for West Gojjam where the control hospital has a higher staffing rate. The number of beds is also generally higher in the intervention hospitals, except for Jimma where the control hospitals have more beds. The catchment population served by the hospitals is generally higher in the intervention hospitals compared to the control hospitals in all zones.

Table 4: Characteristics of sampled health facilities by zone and intervention status, mean

Facility Type			Jimma		Borena		West Gojjam		Total	
			Inter-vention	Control	Inter-vention	Control	Inter-vention	Control	Inter-vention	Control
Health Centers		Average staffing rate	0.54	0.48	0.6	0.7	0.89	0.81	0.59	0.54
		Number of beds	4	2	6	5	4	3	5	3
		Catchment population	36,008	25,978	17,880	43,590	51,522	15,268	31,840	27,612
Hospital	Primary Hospital	Average staff staffing rate	1.15	.	.	0.48	1.1	1.35	1.14	0.91
		Number of beds	60	.	.	26	36	40	52	33
		Catchment population	360,427	.	.	200,000	267,784	170,000	329,546	185,000
	General Hospital	Average staff staffing rate	2.07	1.57	1	.	.	.	1.71	1.57
		Number of beds	83	91	42	.	.	.	69	91
		Catchment population	2,000,000	80,039	1,200,000	.	.	.	1,733,333	80,039
	Total	Average staffing rate	1.61	1.57	1	0.48	1.1	1.35	1.42	1.13
		Number of beds	71	91	42	26	36	40	61	52
		Catchment population	1,180,214	80,039	1,200,000	200,000	267,784	170,000	1,031,440	150,013

Patient Satisfaction Survey

The results show that a total of 597 patients (427 from intervention and 170 from control facilities) were interviewed when they exited the health facilities after receiving service. The patients were from OPD, ANC, PNC, and caregivers of children Under-5 years suffering from pneumonia and diarrhoea. In intervention areas, the majority of patients (79%) received care at a health centre, followed by primary hospitals (12%) and general hospitals (9%).

In total, 61% of the patients were attended by male health workers and 39% were attended by female health workers. There was a significant difference ($p<0.05$) in the sex of the health worker who provided service between the intervention and control groups. In the intervention group, 64% of patients were attended by male health workers, while in the control group 51% were attended by male health workers. This difference was particularly pronounced in Borena and West Gojjam.

Most of the clients (53%) received care from nurses, followed by midwives (34%) and doctors (10%). There were some differences between the intervention and control groups in terms of the type of health worker who provided service. In the intervention group, more clients received care from nurse midwives and doctors, while in the control group more clients received care from nurses.

The highest percentage of patients in both the intervention and control groups were antenatal care (ANC) service clients, with 42% and 40% respectively. The second-highest percentage of patients were out-patient department (OPD) clients, with 15% in the intervention group and 28% in the control group.

The mean age of patients in the intervention group was 22.72 years, while the mean age of patients in the control group was 24.91 years. Regarding the highest level of education, the majority of respondents had no education, with 54% in the intervention group and 40% in the control group. The second most common level of education was primary education, followed by secondary education.

A chi-square test of independence was conducted to determine whether there were statistically significant differences ($p<0.05$) between the intervention and control groups. The results showed statistically significant differences between the groups for patient category ($p=0.003$), patient age ($p=0.003$), and highest level of education completed ($p<0.001$). However, there were no statistically significant differences ($p>0.05$) between the groups for reasons for choosing the health facility.

In regards to travel and expenditures, the results show that the majority of patients in both intervention and control groups traveled by foot. The patient satisfaction survey data shows that in Jimma and Borena, the percentage of patients who used motorcycles/bajaj as their primary mode of transportation was higher in the intervention group than in the control group, whereas in West Gojjam, more patients in the control group used motorcycles/bajaj. However, there were no statistically significant differences ($p>0.05$) in transportation mode between the intervention and control groups.

The mean distance traveled by the intervention group (5.69 km) was higher than that of the control group (4.54 km). The mean travel time for the intervention group (1.01 hours) was also higher than that of the control group (0.80 hours). These differences were statistically significant ($p<0.05$). In terms of travel cost, the mean cost of travel for the intervention group (48.52 birr) was slightly higher than that of the control group (47.78 birr), but the difference was not statistically significant ($p>0.05$).

Table 5: Background characteristics of patients that participated in the survey

Questions and Response Options		Intervention Status		Jimma		Borena		West Gojjam	
		Intervention	Control	Intervention	Control	Intervention	Control	Intervention	Control
		x̄ / %	x̄ / %	x̄ / %	x̄ / %	x̄ / %	x̄ / %	x̄ / %	x̄ / %
Type of Facility	Primary Hospital	12	11	13	0	0	25	33	45
	General Hospital	10	6	9	9	14	0	0	5
	Health Centre	79	82	78	91	86	75	67	50
Health worker sex that provided you the service you wanted at the health facility?	Female	39	49*	44	52	28	43	37	45
	Male	61*	51	56	48	72	58	63	55
Type of health worker	Doctor	10	6	12	5	7	5	9	15
	Clinical Officer	6	11*	2	0	12	30	15	30
	Nurse Midwife	30	27	25	18	36	40	52	50
	Nurse	53	56	62	77	43	25	24	5
	Nurse aid	0	0	0	0	1	0	0	0
	Other	0	0	0	0	1	0	0	0
Client category	Antenatal Care (ANC) Service Client	42	40	43	41	22	33	76	50
	Child Suffering from Pneumonia (IMNCI Pneumonia Client)	18*	10	18	12	23	10	4	0
	Child Suffering from Diarrhoea	13	13	10	10	24	28	2	0
	Out-Patient Department (OPD) Client	15	28*	15	31	15	10	17	50
	Postnatal Care Client (PNC) Client	13	9	13	6	16	20	0	0
Patient's age (years)		22.72	24.91	22.99	25.86	20.07	21.69	27.13	26.15
Can you (the respondent) read and write?	Yes	45	58*	49	63	27	38	63	70
What is the highest level of education obtained/ completed by you (RESPONDENT)?	Preschool	4	5	3	5	5	5	11	5
	Primary	26	31	34	40	11	13	17	15
	Secondary	10	21*	10	18	6	18	22	40
	Tertiary	4	4	4	4	1	3	9	10
	None	54*	40	49	34	76	63	35	30
	Other	1	0	0	0	1	0	7	0
What was the main reason you chose this health facility today instead of a different source of care?	Location close to home	58	70*	68	85	35	43	48	40
	Low cost	5	2	3	0	11	10	2	0
	Trust in providers/ high quality care	25	18	16	12	41	33	39	20
	Availability of drugs	4	2	3	1	7	8	2	0
	Availability of female provider	1	2	0	0	2	5	0	5
	Recommendation or referral	3	2	4	1	1	0	2	10
	Free services readily available	4*	1	4	0	2	3	4	0
	Other	0	4*	0	1	1	0	2	25
What was the next most important reason you chose this health facility today instead of a different source of care?	No other reason	23	32*	22	32	14	8	43	85
	Location close to home	18*	8	14	5	30	18	17	10
	Low cost	8	14*	7	16	16	13	0	0
	Trust in providers high quality care	33*	24	41	29	25	23	4	0
	Availability of drugs	8	8	8	5	8	20	7	0
	Availability of female provider	1	1	1	1	2	0	0	0
	Recommendation or referral	1	1	1	0	0	0	2	5
	Free Services readily available	4	11*	4	12	3	15	9	0
	Other (specify)	3	2	1	1	3	5	17	0
How far is your household from this health facility? Kilometres		5.69	4.54	4.57	3.76	9.08	5.75	4.61	6.37
Time travelled (in hours) to reach the facility (one way)		1.01	0.8	1.11	0.81	0.84	0.65	0.77	1.04
What was your primary mode of transportation today? (One way)	By foot	75	79	79	80	69	78	70	75
	Private car	0	0	0	0	0	0	0	0
	Motorcycles/bajaj	16	18	13	19	25	15	15	15
	Horse	1	0	1	0	0	0	0	0

	Ambulance	1	1	1	0	3	3	0	0
	Public car/bus	5	2	5	1	4	3	13	10
	Traditional Ambulance	0	1	0	0	0	3	0	0
	Other Specify	0	0	0	0	0	0	2	0
How much did it cost (in birr) for you/the patient to travel to the health facility today? (One way)		48.52	47.78	42.16	57.73	62.27	30.56	42.5	35
Waiting time in hours		0.7	.55*	0.75	.52*	0.52	0.57	0.78	0.62
Do you think this was too long?	Yes	26*	15	31	15	19	13	13	25
Was a registration/administration/ consultation/ doctor fee charged?	Yes	19	18	17	17	15	13	39	30
Was a laboratory test done?	Yes	46	41	49	35	41	68	46	25
Was an x-ray done?	Yes	5	2	6	1	0	8	13	0
Were medicines dispensed to you at the pharmacy in the health centre?	Yes	70	74	70	79	77	88	48	15
How much was paid for registration/administration/ consultation/ doctor fee?		17.94	21.5	18.4	20.26	13.13	32	21	16.67
How much was paid for laboratory test?		10.08	5.64	11.8	4.87	5.81	1.85	7.86	32.00 *
How much was paid for an x-ray done?		34.09	0	40.63[4]	.00[5]	.[4]	.00[4]	16.67[4]	.[4]
How much was paid for medicines dispensed to you at the pharmacy in the health centre?		17.26	11.94	16.98	8	16.77	18.71	21.5	47
Total cost spent without transportation		21.84	14.89	23.15	11.51	17.31	21.63	24.26	20.05
How much was spent in total at the facility for this visit, not including transportation costs?		23.78	16.24	24.2	11.51	22.48	26.63	24.26	21.45
Where did the money come from that was used to pay for health care today?	Savings or regular household budget	21	21	19	16	9	23	59	40
	Medical insurance	66	76*	63	82	90	73	22	55
	Selling household possessions	0	0	1	0	0	0	0	0
	Mortgaging or selling land	0	0	0	0	0	0	0	0
	From a friend or relative	2	1	2	0	2	3	0	0
	No payment / free of charge	100	99	99	100	100	98	100	100
	Other (specify)	0	1	1	0	0	3	0	0
Is your family covered under a medical insurance?	Yes	78	79	77	84	91	73	57	70
What type of medical insurance?	Public	96	98	100	100	100	100	54	79
	Private	4	2	0	0	0	0	46	21
How long (in months) has your family been enrolled in the medical insurance?		26.63	26.48	23.12	30.40 *	33.69	19.24 *	29.38	15.71
What services are covered under the medical insurance?	Routine well baby visits incl vaccination	42	52	24	36	75	90	69	79
	Sick child care	81*	64	79	50	89	93	77	93
	Other outpatient care	53	75*	35	66	82	93	85	93
	Antenatal care for pregnant women	60*	50	49	33	78	90	77	79
	Delivery care for pregnant women	56	50	43	33	75	93	88	79
	Post partum care for women and new borns	47	48	32	29	75	93	69	79
	Medicines	81	96*	81	98	81	97	85	86
	Medical tests	77	95*	79	99	77	90	58	79
	Hospital admission and inpatient care	24	69*	14	60	44	90	42	86
	Other (specify)	1	0	1	0	1	0	0	0
Does your family have to pay the following before using the insurance?	Premium	88	100	81	100	100	100	100	100
	Deductible	7	0	10	0	0	0	12	0
	Co-insurance	6	0	9	0	0	0	0	0

**p<0.05*

Qualitative responses given by patients regarding their access to health centers or hospitals suggest that there are various challenges they face. One of the most common issues is the distance and the cost of transportation. For example, one patient said, *"The hospital is located too far from*

our residence and we come here by transport. Nowadays, the transportation fee is very expensive. This is very challenging for women who frequently come here for maternal services." Another patient also mentioned that the health post and health centers located near their residence were not capable of providing necessary services for maternal and child care, which forced them to travel far to access the needed care.

In addition, some patients mentioned that the lack of transportation options or the high cost of available transportation methods made it difficult to reach the health facility. For instance, one patient stated*, "I walked 3 hours to get here. There are no motorcycles because there is no convenient road. The cost to rent a horse is too high."* Another patient mentioned that *"my child was seriously sick at home. There is not transportation so i carried my child long distance on foot."*

Many patients mentioned that the rising cost of transportation hindered their access to healthcare services. One patient stated, "*"We walked 25 to 30 hours on foot to come here. There is a motorcycle for transportation but it is costly. It costs us 50 birr."* Another patient also said, *"When you travel in such ways, the transportation cost rises significantly and goes beyond the cost of the service."* This highlights the need for affordable transportation options to ensure that patients can access healthcare services without financial barriers.

Patients' responses reveal that there are various challenges that hinder their access to healthcare services, including distance, lack of transportation options, high transportation costs, and poor road infrastructure. It is also important to note that many patients do not face these challenges as their residence is located near the health facilities. A patient, for example, said *"The health center is very accessible since I live nearby."*

Regarding the question about whether a registration/administration/consultation/doctor fee was charged, a slightly higher percentage of patients in the intervention group (19%) reported being charged a fee compared to the control group (18%), but the difference was not statistically significant ($p>0.05$) except in Borena.

In terms of reasons for choosing the health facility, the highest percentage of patients in both groups chose the facility due to its location close to home, with 58% in the intervention group and 70% in the control group. Trust in providers/high quality care was the second most common reason for choosing the facility, with 25% in the intervention group and 18% in the control group. The difference was statistically significant at $p<0.05$.

Health Worker Survey Participants

Appendix 1 presents demographic characteristics of health workers that participated in the endline survey. As shown in Appendix 1, the total number of health workers that participated in the survey was 84 in the intervention group and 49 in the control group. Out of the total health workers, 82% are male and 18% are female in the intervention group, and 84% are male and 16% are female in the control group.

In the intervention group, 79% of the health workers work in health centers, followed by primary hospitals with 14%, and general hospitals with 7%. In the control group, 84% of the health workers work in health centers, followed by primary hospitals with 12%, and general hospitals with 4%.

The mean age of health workers in the intervention group is 32.19 years and in the control group, it is 37.73 years, and the difference was statistically significant ($p<0.05$). The majority of health workers in both groups are married/living together, with 57% in the intervention group and 76% in the control group.

Regarding employment status, the vast majority of health workers in both groups (84% in intervention group and 96% in the control group) have permanent and pensionable employment. The level of education shows that 58% of health workers in the intervention group and 55% in the control group have a bachelor's degree, followed by a diploma with 33% in the intervention group and 39% in the control group.

On average, health workers in the intervention and control zones had worked for 2.54 and 4.53 years, respectively, with a significant difference ($p<0.05$) between the two groups. Similarly, the mean months worked at the health facility were 5.02 and 4.35 for the intervention and control groups, respectively, but the difference was not statistically significant ($p>0.05$).

In terms of the position of health workers, nurse (diploma and B.Sc), Health officers (HO) and midwives are the three most common positions held by health workers across all zones. In some cases, there are variations in the distribution of health worker positions between intervention and control zones. For example, intervention zones have a higher percentage of nurse (B.Sc.) compared to control zones, while control zones have a slightly higher percentage of midwives and nurse (diploma) compared to intervention zones. There were few health workers with specialized roles such as radiographer, physiotherapist, anaesthetist, laboratory technician, and laboratory technologist, with only one or two health workers in each group.

Services provided and training taken by health workers

In regard to the types of service provided in the last three months the results show that 50% of health workers in the intervention group provided consultation for children, compared to 35% in the control group. Over half (52%) of health workers in the intervention group provided consultation for adults, compared to 41% in the control group.

For family planning, ANC, PNC, and facility deliveries, higher proportion of health workers in the control group provided these services compared to the intervention group. This was also the case for vaccinations, malaria treatment, and nutrition services.

In regards to training, 65% of health workers in the intervention group reported never receiving training on IMCI, while in the control group, 71% never received training. In contrast, 13% of health workers in the intervention group received training less than 1 year ago, compared to only 2% in the control group.

Regarding malaria, 70% of health workers in the intervention group reported never receiving training, while in the control group, 78% never received training. In the intervention group, 27% of health workers received training more than a year ago, compared to 16% in the control group.

In regards to TB diagnosis and treatment, 70% of health workers in the intervention group reported never receiving training, while in the control group, 78% never received training. In the intervention group, 21% of health workers received training more than a year ago, compared to 18% in the control group.

Regarding family planning methods, 65% of health workers in the intervention group reported never receiving training, while in the control group, 57% never received training. In the intervention group, 31% of health workers received training more than a year ago, compared to 33% in the control group.

The majority (70%) of health workers in the intervention group reported never receiving training on labor and delivery, while in the control group, 80% never received training. In the intervention group, 20% of health workers in both groups received training more than a year ago.

Regarding mental health, 80% of health workers in the intervention group reported never receiving training, while in the control group, 98% never received training. In the intervention group, 19% of health workers received training more than a year ago, compared to only 2% in the control group. Health workers working in intervention facilities were more likely to take training on mental health than those in control facilities.

Close to two-third (65%) of health workers in the intervention group reported never receiving management training, while in the control group, 90% never received training. In the intervention group, 25% of health workers received training more than a year ago, compared to 8% in the control group. Health workers working in control facilities were more likely to never take management training than those in intervention facilities.

Regarding community health, 74% of health workers in the intervention group reported never receiving training, while in the control group, 90% never received training on community health, and the difference was statistically significant. ($p<0.05$).In the intervention group, 21% of health workers received training more than a year ago, compared to 6% in the control group.

Most (71%) of health workers in the intervention group reported never receiving training on pre-post natal are, while in the control group, 86% never received training. In the intervention group, 25% of health workers received training more than a year ago, compared to 14% in the control group.

Most (71%) of health workers in the intervention group reported never receiving training on HIV/AIDS care and management, while in the control group, 84% never received training. In the intervention group, 21% of health workers received training more than a year ago, compared to 16% in the control group.

The vast majority (77%) of health workers in the intervention group reported never receiving training on hypertension, while in the control group, 94% never received training, and the difference was statistically significant ($p<0.05$). In the intervention group, 19% of health workers received training more than a year ago, compared to 6% in the control group.

Three-quarters (75%) of health workers in the intervention group reported never receiving training on diabetes, while in the control group, 96% never received training. Health workers working in control facilities were more likely to never take training on diabetes than those in intervention facilities. In the intervention group, 20% of health workers received training more than a year ago, compared to 4% in the control group.

Health workers belonging control facilities were also more likely to not to take training on TB diagnosis and treatment, EmOnc (Emergency Obstetric and Neonatal Care), and LSS).

Overall, the intervention group had a higher percentage of health workers that received training on most areas, indicating that the intervention may have had a positive impact on the provision of training. However, the results suggest that the vast majority of health workers have never received training on most health services in both intervention and control areas. However, health workers belonging to the intervention facilities were more likely to take training on the different health services compared to those in the control group.

3.2. Relevance of the PBF Program

PBF is a health financing approach that aims to improve the quantity and quality of healthcare services by incentivizing healthcare providers based on their performance. In this sub-section, the

findings in regards to the relevance/ contribution of the PBF program to the government priorities and needs of stakeholders; relevance to facilities, patients and other stakeholders; and appropriateness and adequacy of inputs and strategies to achieve intended results are presented.

Relevance/Contribution of PBF to HSTP II Pillars

The PBF program has contributed to each of the five pillars of the Ethiopian Health Sector Transformation Plan II (HSTP II). The PBF program is a health system strengthening approach that aims to improve health service delivery and health outcomes by providing financial incentives to health facilities based on their performance. The PBF program's contributions to each of these pillars of the HSTP II work together to strengthen the health system, and improve utilization and availability of quality health services.

Table 6 summarizes how the PBF program has contributed to each of the pillars:

Table 6: Contribution of the PBF program to each pillar of HSTP II

Pillars of HSTP II	PBF Program Contributions
Service delivery and quality improvement	The PBF program contributed to this pillar by providing financial incentives to health facilities based on their performance. This incentivized health facilities to improve the quality and quantity of their services and ensure that they are accessible to the community. The program has also encouraged the use of performance data to identify areas for improvement and make evidence-based decisions.
Health workforce development	The PBF program has promoted health workforce development by providing financial incentives to health workers based on their performance. This has motivated health workers to perform better and improve the quality of their services. Some key informants also stated that the financial incentives and improved work environment made possible by the project's financial support has reduced health worker turnover. Additionally, the program has provided training to enhance the skills and knowledge of health workers, which in turn enhanced quality of services.
Health system financing	HSTP II has prioritized PBF program as one of the innovative financing mechanisms. The implementation of the PBF program created opportunities to demonstrate the implementation and management of the program on the ground as well as its achievements and impact to the government and other stakeholders. Although PBF as a financing mechanism is not still institutionalized, the implementation of this PBF program helped stakeholders understand how such a program improves quality and utilization of health care, which may lead to institutionalization and scaling up of the program at a larger scale.
Health information system strengthening	The PBF program has contributed to health information system strengthening by providing technical support and training, verification of declared data, and linking release of subsidy to health facilities when error margins between declared and verified data do not exceed a predefined level on each of the supported performance indicators. This data has been used to inform decision-making and to identify areas for improvement by health facilities as well as regulatory bodies. The program has also enhanced the capacity of health information systems to generate reliable and timely data for monitoring and evaluation. Improvements in data quality is also encouraging use of the data to inform decision making by health facilities and regulators. However,

	DHIS2 database was found to contain a lot of inconsistent and missing data, which suggests the need to further strengthen the system.
Leadership, governance, and accountability	The PBF program has enhanced leadership, governance, and accountability in the health system by establishing clear performance standards and regular monitoring mechanisms. This has promoted accountability and transparency by helping health facilities report actual or real performance data. The management committee established at health facilities has had a big role in the implementation business plan and effective utilization of the resource received from PBF project. The collective decision making and plan driven implementation of activities are said to have improved accountability and transparency in health facilities. PBF has promoted accountability and transparency through strong monitoring and evaluation. There was also periodic financial audit of health facilities. However, some key informants were concerned that the fund holding and the verification activities were undertaken by one organization (Cordaid in both cases). They stated that these functions should be separated to ensure transparency and accountability. The program has also strengthened the capacity of zonal and woreda level regulators to fulfill their regulatory mandates.

Overall, the PBF program in Ethiopia has contributed to all pillars of the HSTP II by promoting access to and quality of health services, improving health workforce performance and motivation, enhancing health system financing and information systems, and promoting effective leadership and governance.

Relevance to identified needs of the beneficiaries and stakeholders – patients, health workers, health authorities

The program has improved the quality of healthcare services by incentivizing healthcare providers to provide better services. PBF has also increased the utilization of healthcare services by improving access to healthcare services. Both qualitative and quantitative data analysis results suggest that the program has improved quality and utilization of health care services. The results were further confirmed by the key informants, who generally agreed the program increased the number of visits to healthcare facilities by the target population and reduced the number of home deliveries and maternal mortality. A key informant noted *"Since the service quality provided at the health facility has shown improvement, clients come to the facility abundantly and receive the service they need."* Another informant highlighted, "*health facilities have become more attractive. It has increased the desire of clients or patients to use our facilities. They have previously visited private clinics and raised complaints about our services. Now that they have access to quality and adequate services, their demand has increased." Another informant also said, "Improving the quality of healthcare facilities presents a significant opportunity for women and children to access better healthcare services as they are more vulnerable to diseases than men. Therefore, the program is highly relevant to increase the number of women that utilize health services."*

The improved availability of quality health services enabled by the program has allowed CBHI and marginalized community members, and the general public have access to such improved services in nearby locations without the need to seek services from private providers or other places. One informant stated that the program has *"…increased the availability of essential medicine demanded by the community groups who are not capable of buying medicine from private health facilities. So, it has high contribution for the CBHI users to get service."* Another informant highlighted, *"It significantly aids the marginalized community by providing sufficient medical equipment and essential medicine, enabling them to access primary healthcare services in close proximity to their homes."* Some informants also mentioned that the improved availability of quality health service has reduced complaints from the community. Informants in this regard said the following:

"…health facilities have become more attractive. It has increased the desire of client or patient to use our facilities. They have previously visited private clinics and raised complaints about our services. Now that they have access to quality and adequate services, their demand has increased. Since the advent of PBF, there are adequate medical equipment, adequate medicines and adequate health services. Therefore, the community's need to utilize service in health facilities has increased."

"The primary need of the community is access to quality and comprehensive healthcare services in close proximity to their homes. To this end, the project has implemented quality enhancement programs using the PBF approach, which has been instrumental in improving the delivery of healthcare services to the community. For marginalized communities located far from modern healthcare facilities, the availability of the PBF program has proven to be particularly beneficial."

The PBF program is also highly relevant to the needs of the health facilities and other stakeholders. It has been successful in responding to the needs of health facilities and regulators by improving regulatory capacity and data handling and reporting systems. It has improved the infrastructure of the health facilities, provided additional income to health facilities and their workers through incentives which helped them make quality health services accessible to the public.

Appropriateness and adequacy of inputs and strategies to achieve intended results

Informants provided several positive comments about the PBF program inputs and strategies in achieving the desired results. They appreciated the regular meetings held by the management committee for performance evaluation and the continuous feedback mechanism that was in place. Before the PBF program, according to many key informants, the main issue was about the coverage of the service, but the program enabled the departments, management committee, and health facility to focus on quality service provision, from which the community greatly benefited. The informants also commended the effectiveness of the checklist developed by Cordaid, which addressed issues related to the quality of healthcare, health worker discipline, and working environment, among others. They also found the inputs provided by Cordaid, such as training, supportive supervision, and finances, to be important in enhancing the quality of health service utilization. The monitoring and evaluation process was also noted as being helpful in identifying and resolving constraints faced by the health facilities.

The informants also agreed that the financial support they received from the project greatly benefited their health facility. The use of incentives was also seen as a positive motivator for health facilities and professionals to provide quality services to the community. One informant stated, *"Incentives have increased the quality of service."*

There were some criticisms, however. One informant noted that while the support provided by the project was appropriate, it was not adequate as the health facility served a huge number of clients, and the price of materials and medicines had been increasing. Another informant also stated that *"the amount of money allocated for the health facilities was not adequate to buy medicine."* [paraphrased]. Some informants also found the financial incentive provided to health workers inadequate. As a result, they recommended increasing the incentive paid to health workers.

Overall, the key informants provided positive feedback about the PBF program inputs and strategies, with some minor criticisms. They highlighted the effectiveness of the monitoring and evaluation processes, the usefulness of the tools and checklists used, and the appropriateness of the financial support provided. They also emphasized the importance of the management structure and technical assistance provided by Cordaid.

Relevance of program to health posts

HEWs were asked to assess the relevance of the selected health services or indicators for health posts for which they have been incentivized by the PBF program. The health extension workers (HEWs) shared their perspectives on the relevance and effectiveness of the selected health services and indicators for which they have been incentivized by the PBF program. The analysis reveals that the PBF program has generally been well-received by the HEWs and is seen as a significant contributor to their motivation, quality of service provision, and accurate registration and reporting. A HEW noted, *"The technical assistance provided through this program has had a significant contribution to the successful screening and referral of selected services such as pneumonia and child growth monitoring interventions."*

HEWs also highlighted the importance of the program in improving data quality at health post level. A HEWs noted the positive impact of the PBF project on the quality of their work and data: *"Before the implementation of the PBF project, we used to record activities carelessly...But currently, because an evaluation [monitoring] is carried out each month, we carefully register activities."* Another HEW stated that the program has significantly contributed to minimizing errors during referrals and improving client information registration. In another response, a HEW highlighted the discrepancy between data at the health center and their records, and stressed the importance of improving registration and documentation.

The incentives also seem to have a positive impact on HEWs' motivation and dedication to their work. A HEWs stated, *"Not only for health posts, but the finance given to HEWs has also a good contribution in motivating the HEWs and carrying out their tasks in the right way."* Another HEW mentioned the potential of the PBF project to raise the motivation of HEWs and improve the quality of healthcare.

Although some HEWs mentioned the need to increase the number of indicators or services for which health posts will have to be incentivized, they found the selected indicators/services relevant. A HEW, for example, mentioned that *"Incentivizing these two indicators [family planning and EPI] is very relevant and has many advantages, such as motivating the HEWs to strongly achieve their plans, increasing service coverage, and bringing about better healthcare improvements."*

However, the HEWs also identified several challenges and areas for improvement in the program. A common theme among their responses is the need for better resources and infrastructure at the health post level. One HEW emphasized the need for better infrastructure, stating, *"Subsequently, as the health post is not conducive to counseling, it also needs to be repaired."*

Some HEWs also mentioned challenges as related to shortage of registration books. *"At the health post level, there is a shortage of family folders, and the existing folders are also old. Additionally, there is a shortage of EPI registration books that we require,"* a HEW said. The HEW worker also suggested that *"As it affects data quality, they need to be replaced with new registration forms."* Many HEWs also stated that they are required to take the registration books to the health center for evaluation. A HEW in this regard said, *"We put all the documents there and come back home without exchanging any information or receiving feedback for the activities we have been working on."* Another HEWs stated *"Currently, we take all registration books and client data for evaluation to the health center."*

A number of HEWs highlighted the need for health post maintenance, supplying essential medical equipment, medicine, and improving the working environment. A HEW emphasized the need for essential medical equipment like a BP apparatus, and recommended that such equipment should be supplied. Others also mentioned the shortage of medical equipment and materials, particularly thermometers and registration books. Some HEWs also mentioned the shortage of family planning inputs (medicine) at the health post level as a challenge. A HEW stated, *"The provision of long-term*

family planning (IUD applying and removal) requires a standard health post, which our health post is not fit for."

Additionally, HEWs mentioned the need for more training and support, particularly in the context of the PBF program's monitoring and evaluation processes. One HEW noted, "*The monitoring and follow-up mechanisms should be modified, and any support provider should come to the health post and provide all necessary support..."* Others pointed out specific challenges, such as a lack of consistent understanding of case definitions among HEWs, which could be addressed through training and technical assistance. Another HEW highlighted that, *"The case that one HEW screens for measles may be rejected by another HEW due to a lack of adequate knowledge and skills on cases. This challenge can be resolved through training and technical assistance."*

Several HEWs mentioned the importance of feedback mechanisms, especially on the three indicators selected for the health posts. A HEW also recommended refresher training and review meetings with Cordaid to have a common understanding of the PBF project and monitor the interventions. Another HEW stated that there is a need for *"appropriate feedback … because most of the time, we don't receive any feedback after submitting the report to the health center."* Another HEW mentioned that the communication and interaction with the PBF program are poor, and suggested arranging meetings at the center level for experience sharing or providing constructive feedback during health post visits. *"Most of the time, we submit reports to the health center every month and return to our health post. However, there is no time when the health post and Cordaid review our gaps and give comments on our achievements. We provide the report to them, and subsequently, they allocate 30% of funds for us. The communication and interaction of the PBF program is poor,"* a HEW stated.

The responses from health extension workers (HEWs) also revealed a variety of experiences and perspectives regarding the funds they receive quarterly and how they have been used. Some HEWs found the funds helpful in addressing specific needs, such as purchasing medical equipment like a blood pressure (BP) apparatus, a thermometer, and new registration books. A HEW shared, *"Overall, the finance allocated for the health post is very useful and helpful in terms of successfully achieving our plan."* Another HEW mentioned that the 30% incentive was helpful in covering transportation costs and other personal expenses.

However, some HEWs mentioned receiving funds but found them insufficient. A HEW stated, "*In fact, funds have been allocated to us for the last year according to the guidelines designed by the project implementer. However, it is not sufficient."* The funds were used for purchasing stationery materials, maintenance purposes, and improving health post premises. Another HEW described using the funds for wall painting and purchasing medical equipment, but also emphasized that the amount was small.

Some HEWs reported only receiving the 30% incentive for themselves and not seeing any funds allocated for the health post. One HEW said, *"I know that there is finance available for HEWs, and that it has also been allocated for the health post. However, to date, only 30% of the finance has been distributed to HEWs, and I have not seen any finance allocated for the health post to enhance the working environment."*

Other HEWs expressed dissatisfaction with the distribution of funds. One HEW mentioned a discrepancy between the amount they were told they would receive and the actual practice: *"We have only received 15%, and the remaining 85% is used for the health post, as the health post head stated."*

There were also HEWs that reported that they *"…have not received any funds for the health post, and their requests for basic needs such as mats, repairs, and a latrine with a hand-washing facility have not been met."* Similarly, another HEW stated he had no information about the funds since

they do not have a health post office. A HEW recommended that the *"financial allocation and management need more attention and should be regulated properly."* Some HEWs blamed the health centers for using the funds allocated for health posts for health center or other purposes. A HEW noted *"Until now, we haven't received finance from the PBF program due to cash wastage happening at the health center."*

In conclusion, the PBF program has been relevant to the priorities and needs of health facilities, and patients. PBF has improved the quality of healthcare services, improved access to healthcare services, and improved regulatory capacity and data quality. Furthermore, the program's inputs and strategies were generally considered realistic and adequate to achieve its targets/objectives. The PBF program has also been well-received by health facilities and health workers, who perceive it as an important contributor to their motivation, quality of service provision, and accurate registration and reporting.

3.3. Coherence

Coherence of design, implementation and results with international and national laws and commitments to gender equality and rights

At the global level, PBF is recognized as a key strategy for achieving Universal Health Coverage (UHC), which is a key goal of the Sustainable Development Goals (SDGs) set by the United Nations. The World Health Organization (WHO) has emphasized the need to strengthen health financing mechanisms, including PBF, to achieve UHC (WHO, 2019). The PBF program can be also seen as a coherent strategy with the Sustainable Development Goals (SDGs), particularly with SDG 3: Ensure healthy lives and promote well-being for all at all ages. PBF can be an effective tool to improve the quality, accessibility, and coverage of healthcare services, which are important components of SDG 3.

At the regional level, the African Union has also emphasized the importance of health financing mechanisms, including PBF, in achieving UHC. The African Union Commission has recommended the use of PBF to improve the quality and coverage of health services in African countries (African Union Commission, 2017)[15]. As a member state of the African Union, Ethiopia has also recognized the importance of PBF in achieving UHC.

Alignment of project with national legislation and initiatives

At the national level, the Ethiopian government has introduced various health policies, laws, and instruments to improve the health outcomes of its citizens. The Second Ethiopian Health Sector Transformation Plan (HSTP II) is one such instruments/ strategies that aims to improve the health status of the population through the provision of equitable and quality healthcare services (Federal Ministry of Health Ethiopia, 2021)[16]. Indeed, the PBF program aligns with most pillars outlined in the HSTP II, including improving access to pharmaceuticals and medical devices; improving regulatory systems; improve human resource development and management; enhancing informed decision making and innovations; strengthening governance and leadership and improving health infrastructure. PBF is also mentioned as one of the innovating health financing strategies of the HSTP II to achieve this goal. The evaluation findings also confirmed that the program contributed to these priorities and targets set under each in the HSTP II. The program was also well aligned with the needs of the community and health facilities.

[15] African Union Commission. (2017). African Union Commission, UNAIDS and WHO Launch New HIV and Health Financing Platform for Africa. Retrieved from https://au.int/en/pressreleases/20170710/african-union-commission-unaids-and-who-launch-new-hiv-and-health-financing

[16] Federal Ministry of Health Ethiopia. (2021). Ethiopian Second Health Sector Transformation Plan 2020/21-2024/25.

PBF is also coherent with different health policies. The 1993 Health Policy of Ethiopia that endorses the development of an equitable and acceptable standard of health service system, for instance, aimed at reaching all segments of the population within the limits of resources. PBF incentivizes healthcare providers to provide quality healthcare services, which is in line with the objectives of the health policy and the health policy.

Complementarity, harmonization, and co-ordination with CBHI

The key informants provided insights into how the PBF program is harmonized with community-based health insurance (CBHI). They mentioned that the finance collected from CBHI is not enough to provide health services to the members. Without the PBF program, many health facilities would be closed down as it financially supports the provision of health services to CBHI members. The PBF program provides medicine and equipment to supplement the supply from CBHI. The results indicate that PBF provides a source of funding for health facilities that helped improve the quality of care and make it more accessible for CBHI members.

Another informant emphasized the role of PBF in improving the financial and service provision capacity of health facilities. The informant noted that *"the main benefits of PBF project are the enhancement of primary hospital in financial and quality service provision by resolving the constraints that mostly relate with laboratory and medicine availability issues."*

Key informants noted that PBF program has helped improve the availability of essential medicines and laboratory equipment, which is crucial for providing quality care. An informant stated that *"the availability of particularly medical [supplies] allows the CBHI members to continuously come to health facilities and to use the health services."* Another informant highlighted that *"PBF loans CBHI when a shortage happens and CBHI reimburses later on."* This demonstrates the complementary relationship between the two programs and how they support each other to fill gaps in health care financing. These highlight the importance of PBF in improving the quality and accessibility of care for CBHI beneficiaries by addressing critical gaps in the provision of essential medical supplies and equipment.

Key informants recognize the critical role that PBF plays in supporting CBHI and ensuring that essential health services are available to CBHI members. One informant noted that the fees collected through CBHI are insufficient to provide the necessary health services to the community, and that the PBF program financially supports the provision of these services. Additionally, when CBHI funds are delayed, the PBF program has been used to purchase drug and other supplies. An informant noted that *"when CBHI payment is delayed or inadequate, PBF supports CBHI by providing loans to cover the gap. It is observed that the PBF program is essential in supporting the CBHI and ensuring that the CBHI members receive the services they require."* This has led to an increase in the availability of essential medicine that is demanded by the community who are not capable of buying medicine from private health facilities. The PBF program has a high contribution to CBHI users' ability to get service.

The evaluation team also looked at the prevalence and characteristics of medical insurance coverage across different zones by intervention status and zone, although the data was not specific to CBHI. Table 7 indicates that the vast majority of the families of patients in intervention areas (78%) and control areas (79%) have medical insurance coverage. Insurance coverage is highest in Borena and lowest in West Gojjam. The percentage of families with medical insurance coverage did not vary statistically significantly ($p>0.05$) between the intervention and the control groups, which was also the case in Jimma and West Gojjam zones. In Borena zone, insurance coverage is statistically significantly ($p<0.05$) higher in the intervention group compared to the control group.

The results also show that public insurance is the dominant type of insurance across all zones (ranging from 54-100%) compared to private insurance (ranging from 0-46%). The average length of time that families have been enrolled in medical insurance was less than 27 months, but varies across zones, with the longest enrollment duration in Borena intervention area and the shortest in West Gojjam control area. Average length of time that families have been enrolled in medical insurance is statistically significantly ($p<0.05$) higher in Borena intervention areas compared to control areas, while the reverse was true in Jimma zone. The most commonly covered services under the medical insurance include sick child care, medicines, medical tests, antenatal care for pregnant women, delivery care for pregnant women, and other outpatient care. Hospital admission and inpatient care are the least commonly covered services. Most families (81-100%) have to pay premiums before using the insurance, while a very small percentage have to pay deductibles or co-insurance.

Table 7: Medical insurance coverage of patient or clients who participated in the endline survey by intervention status and zone

Questions	Response Categories	Total		Jimma		Borena		West Gojjam	
		Inter-vention	Control	Inter-vention	Control	Inter-vention	Control	Inter-vention	Control
		x̄/%	x̄/%	x̄/%	x̄/%	x̄/%	x̄/%	x̄/%	x̄/%
Is your family covered under a medical insurance?	Yes	78	79	77	84	91*	73	57	70
	No	22	21	23	16	9	28*	43	30
What type of medical insurance?	Public	96	98	100	100	100	100	54	79
	Private	4	2	0	0	0	0	46	21
How long (in months) has your family been enrolled in the medical insurance? Mean		26.63	26.48	23.12	30.40*	33.69*	19.24	29.38	15.71
What services are covered under the medical insurance?	Routine well baby visits including vaccination	42	52	24	36	75	90	69	79
	Sick child care	81	64	79	50	89	93	77	93
	Other outpatient care	53	75	35	66	82	93	85	93
	Antenatal care for pregnant women	60	50	49	33	78	90	77	79
	Delivery care for pregnant women	56	50	43	33	75	93	88	79
	Post partum care for women and new borns	47	48	32	29	75	93	69	79
	Medicines	81	96	81	98	81	97	85	86
	Medical tests	77	95	79	99	77	90	58	79
	Hospital admission and inpatient care	24	69	14	60	44	90	42	86
	Other (specify)	1	0	1	0	1	0	0	0
Does your family have to pay the following before using the insurance?	Premium	88	100	81	100	100	100	100	100
	Deductible	7	0	10	0	0	0	12	0
	Co-insurance	6	0	9	0	0	0	0	0

*p<0.05

Overall, the key informants' responses suggest that PBF and CBHI have a strong complementarity and that the PBF program played an important role in improving the quality and accessibility of care for CBHI beneficiaries. By addressing critical gaps in the provision of essential medical supplies and equipment, PBF helped health facilities provide better care for marginalized populations and increased the number of CBHI beneficiaries who can access quality health services. One informant pointed out that CBHI funds collected from the community are not enough

to provide adequate health services. Therefore, the PBF financially supports health facilities to provide services to CBHI members. The PBF program also supplies drugs that CBHI members need but cannot get due to insufficient CBHI funds. The availability of medical equipment and essential medicines has allowed insurance users to come to health facilities and use health services. This highlights how the PBF supports CBHI. An informant in this regard noted that *"the interaction between the PBF project and CBHI program is very high, and the PBF project has become a backbone for the successful implementation of CBHI."* Quantitative results suggest that medical insurance coverage of patients families did not differ by intervention status, although there was zonal variations.

Linkages between PBF and other projects and programs

Key informants provided mixed feedback on the synergy and linkages between the PBF project and other programs and projects. Some respondents stated that there were no other projects working on health issues, while others mentioned that there were other organizations working on various health interventions.

Several respondents acknowledged the importance of PBF in supporting their health facilities and highlighted the fact that PBF provided them with all-rounded support, which other organizations did not. The PBF project was also reported to have strong follow-up and evaluation mechanisms, which other projects lacked. Many informants indicated that the PBF program was unique and there was no duplication of activities. *"The [PBF] program was aligned with other projects that have been implemented in the region and we make sure the coherence of different projects to avoid duplication of efforts or redundancies,"* an informant stated. The project is also well aligned with CBHI as presented in the sub-section above.

Some respondents noted that although there was no direct interaction and linkage among organizations implementing health interventions, the activity of these organizations would support each other. An informant, for example, stated, *"The implementation of PBF project has high contribution for the other project implemented by other organizations because it built strong foundation for them. PBF primarily focuses on the institutional strengthening interventions whereas the other projects emphasized more on the specific service or program deliverance activities."* This suggests that PBF and other projects implemented in the intervention areas support each other although there may not be observable or formal collaborative work among the project implementers.

However, several respondents also mentioned that there was a lack of coordination and interaction among organizations, with each organization working separately to achieve its own objectives. For example, one respondent said, *"… there are various organization that are participating in different health interventions. But…we didn't see them working together for the community. All are running to achieve their own desired objectives rather than sharing experiences among themselves."* Another informant also noted that "*In this woreda there are organizations that work on health. …Although the interventions carried out by the PBF project supports the activities of the [other organizations/projects], we did not see them cooperate and work together to change the health problem of the society."*

Overall, while there were some reports of cooperation and indirect integration among projects (including PBF program) implemented in the intervention areas, there were also indications of a lack of coordination and interaction, with organizations working separately and focusing on their own objectives.

In conclusion, the performance-based financing program is coherent with global, regional, and national health policies, laws, and instruments. PBF aligns with the principles, priorities and objectives of the health policy, the HSTP II, the national health financing strategy of Ethiopia, and

the recommendations of the African Union Commission. The project is also aligned with other health and nutrition projects implemented in intervention areas. However, cooperation among organizations that implement different health and nutrition projects was either absent or unsatisfactory.

3.4. Effectiveness and Impact

The PBF program had three desired outcomes namely: improved health service delivery, improved governance of health service delivery, and enhanced health information system. In this section of this report, the findings on progress made by the program in achieving its desired outcomes and sub-outcomes are presented.

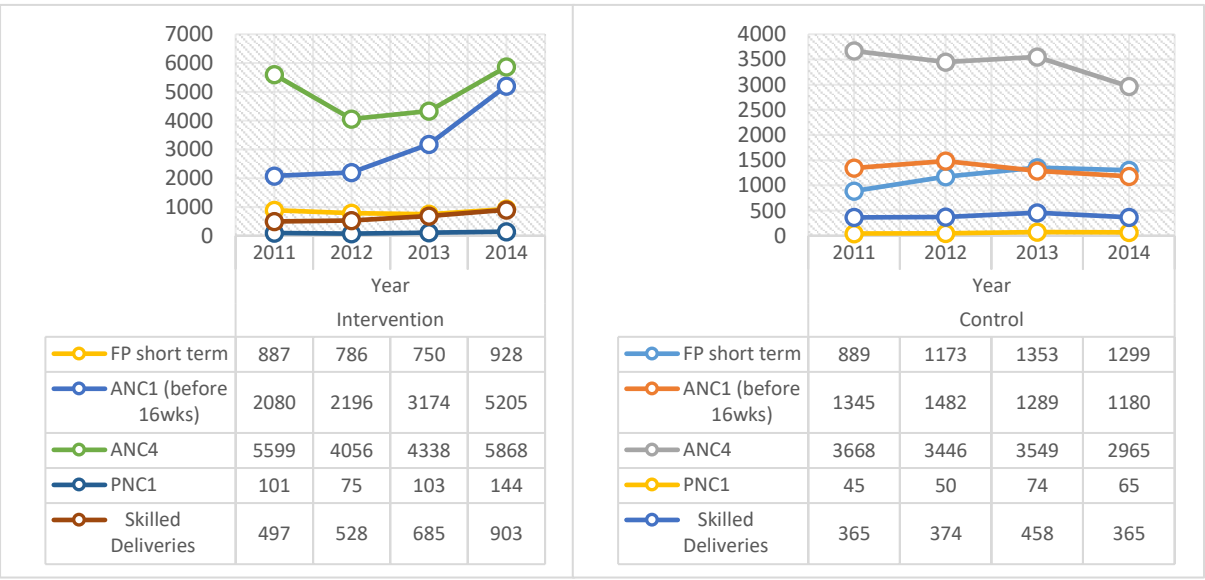
3.4.1. Outcome 1: Improved Health Service Delivery

The first desired outcome of the PBF program is improved health service delivery in selected woredas reflected in increased utilization of quality services and increased equity in access to health services. The findings are presented under these two result areas below.

3.4.1.1. Increased utilization of good quality services

i. Quantity of health service utilization

We analyzed the impact of PBF programs on the volume of health service utilization using Disability-Adjusted Life Years (DALYs) averted as a measure of health outcomes. Using DALY, we tested whether DALYs averted has increased over time in intervention compared to control facilities. Figure 1 shows the trend in health service utilization of first Antenatal Care (ANC) visit before 16 weeks, four ANC visits (ANC4), births attended by skilled health personnel (skilled deliveries), postnatal care (PNC) visit within first 7 days, and first and repeated visits for Family Planning (FP) modern methods short term from 2018/19 to 2021/22 in intervention and control health facilities.



Intervention	2011	2012	2013	2014
FP short term	887	786	750	928
ANC1 (before 16wks)	2080	2196	3174	5205
ANC4	5599	4056	4338	5868
PNC1	101	75	103	144
Skilled Deliveries	497	528	685	903

Control	2011	2012	2013	2014
FP short term	889	1173	1353	1299
ANC1 (before 16wks)	1345	1482	1289	1180
ANC4	3668	3446	3549	2965
PNC1	45	50	74	65
Skilled Deliveries	365	374	458	365

Figure 1: Trends in utilization of FP short term, ANC1, ANC4, PNC1, and skilled deliveries in DALYs averted in control and intervention groups from 2018/19 to 2021/22

The trend analysis depicted in Figure 1 above for the FP short-term showed a relatively stable trend in both groups, with control group fluctuating around yearly average of 1,178 while

intervention group fluctuating around 838. For ANC1, the control group saw a decrease from 1,345 in 2018/19 to 1,180 in 2021/22, while the intervention group showed an increase from 2,080 in 2018/19 to 5,205 in 2021/22. The trend for ANC4 visits was similar, with the control group showing a decrease from 3,668 in 2018/19 to 2,965 in 2021/22, and the intervention group showing an increase from 5,599 in 2018/19 to 5,868 in 2021/22. The number of PNC1 visits remained low for both groups over the four-year period, with the intervention group consistently showing higher usage, while, the number of skilled deliveries showed an increase in both groups, with the intervention group having higher usage compared to the control group.

The trend analysis was also conducted for other key services such as FP (long-term), outpatient department (OPD) Under-5, and all project-supported health services combined (total) in DALYs comparing the control and intervention groups from 2018/19 to 2021/22. Figure 2 depicts utilization, in DALYs averted, of Family Planning (FP) long-term, Outpatient Department (OPD) Under-5, and all project-supported health services combined (total) from 2018/19 to 2021/22 in intervention and control health facilities.

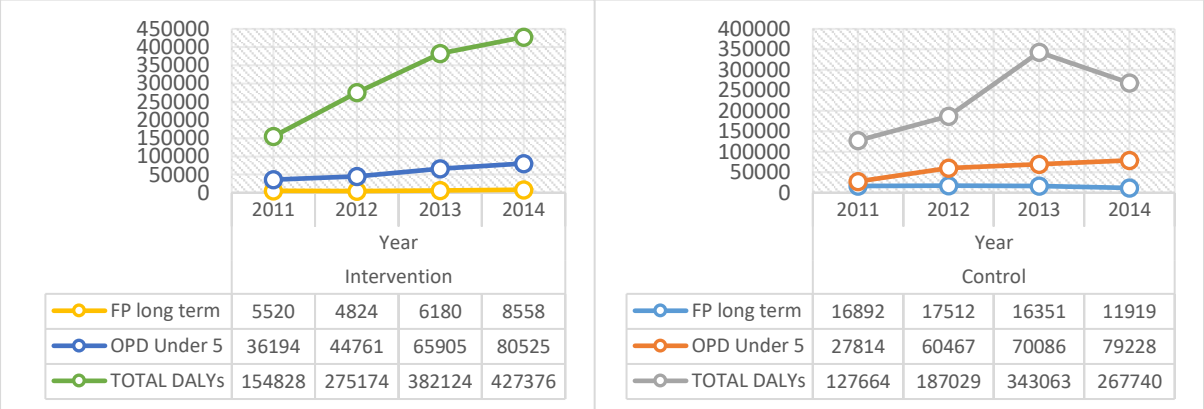


Figure 2: Trends in utilization of FP long term, OPD Under-5 and total (all project supported services combined) in DALYs averted in control and intervention groups from 2018/19 to 2021/22

The results shown in Figure 2 above indicates that long-term FP in the control group showed a decrease from 16,892 in 2018/19 to 11,919 in 2021/22. In contrast, in intervention facilities, there was an initial decrease from 5,520 in 2018/19 to 4,824 in 2019/20, followed by a sharp increase to 8,558 DALYs averted in 2021/22.

Looking at the trend in the utilization of OPD Under-5, we can see that both the control and intervention groups show an increasing trend from 2018/19 to 2021/22. The control group had a total of, in DALYs averted, 27,814 OPD visits in 2018/19, which increased to 79,228 in 2021/22. The intervention group had 36,194 OPD visits in 2018/19, which increased to 80,525 in 2021/22.

When we look at utilization trends for all project-services combined (total), the control group had a total of 127,664 DALYs averted in 2018/19 which increased to 267,740 in 2021/22. The intervention group also showed an increasing trend from 154,828 in 2018/19 to 427,376 DALYs averted in 2021/22.

In conclusion, the trend analysis of the health service indicators shows mixed results between the control and intervention groups. While some indicators, such as ANC1 and ANC4 visits and FP long-term, showed a decrease in the control group, while there was an increase in these indicators and others. In both the intervention and control facilities, there was an increasing trend in utilization of OPD Under-5, PNC1, short-term family planning, and all project supported services combined,

The evaluation also compared mean utilization in DALYs averted per year between control and intervention groups at baseline and endline. The analysis was conducted based on declared data; however, it is important to note that the declared data reported by intervention facilities may not be comparable to those reported by control facilities due to the verification activities that were carried out to discourage false reporting. The result of the comparison is described in Table 8 below.

Average utilization in DALYs averted per year grew from 10,375.47 to 16,733.74, and from 11,276.14 to 17,095.02 in the control and intervention facilities, respectively. However, the p-values for both the control and intervention are above the conventional significance threshold of 0.05, suggesting the observed differences between the baseline and endline values were not statistically significant.

Table 8: Comparison of mean utilization in DALYs averted per year between control and intervention groups at baseline and endline

	Time		Change
	Baseline	Endline	
Status of Intervention	Mean	Mean	%
Control	10375.47	16733.74 ($p = 0.129$)	61%
Intervention	11276.14	17095.02 ($p = 0.065$)	52%

The evaluation also compared average volume of service utilization measured in DALYs averted for selected health services. Table 9 below presents mean volume of utilization, along with the corresponding p-values in parentheses, for selected key health services at baseline and endline for both control and intervention groups. The p-values indicate whether the difference in mean values between the baseline and endline is statistically significant.

The results presented in Table 9 show that average volume of service utilization for ANC1 grew statistically significantly from 90.44 DALYs averted per year at baseline to 208.21 DALYs averted per year at endline (P=0.001) in the intervention group, while this was not the case in the control group. In fact, average utilization declined slightly from 86.86 at baseline to 73.77 DALYs per year at endline (P=0.823). There were no significant differences in the mean volume of utilization of FP long-term, FP short-term, ANC4, PNC1, skilled deliveries, or OPD under 5 visits between baseline and endline in both the intervention and control facilities.

Table 9: Mean volume of service utilization of key services in DALYs averted per year in intervention and control facilities at baseline and endline

	Status of Intervention			
	Control		Intervention	
Health Services	Baseline	Endline	Baseline	Endline
FP long term	1099.96 ($P = 0.383$)	744.92	265.42	342.31 ($P = 0.218$)
FP short term	57.96	81.16 ($P = 0.467$)	38.57 ($P = 0.858$)	37.11
ANC1 (before 16wks)	86.86 ($P = 0.823$)	73.77	90.44	208.21 ($P = 0.001$)
ANC4	237.18 (0= 0.518)	185.29	243.42 ($P = 0.857$)	234.70
PNC1	3.39	4.08 ($P = 0.451$)	4.38	5.74 ($P = 0.189$)
Skilled Deliveries	23.81	22.83	29.24	36.1

	($P = 0.906$)			($P = 0.348$)
OPD Under 5	2020.00	4951.74 ($P = 0.133$)	2244.76	3220.99 ($P = 0.066$)

Table 10 below compares mean volume of service utilization in DALYs for selected health services by type of health facility and source of livelihood.

As shown in Table 10, a significant increase in the DALYs averted for ANC1 visits from 104.64 at baseline to 242.21 at endline (p=0.000) in the intervention health centers. The total DALYs averted by the health centers are also found to significantly increase from 7,932.79 at the baseline to 16,677.37 at the endline (P=0.000). In the contrary, no significant variation was observed in the total DALYs averted for the selected health services provided by the hospitals between baseline and endline. However, it should be noted here that the sample size for the hospitals is too small to test for significant differences between the baseline and endline.

In the agrarian intervention areas, volume of service utilization in DALYs averted for ANC1 increased from 29.69 at baseline to 105.9 at endline, and the difference was statistically significant (P=0 .011). This was not the case in the pastoral areas. No statistically significant differences were found in DALYs averted between baseline and endline for all other selected services (FP long-term and short-term methods, PNC1, ANC4, OPD Under-5, and skilled deliveries).

Table 10: Mean volume of service utilization of key services in DALYs averted per year in intervention facilities by source of livelihood and facility type

Health services	Facility Type				Livelihood			
	Health Center		Hospital		Agrarian		Pastoralist	
	Baseline	Endline	Baseline	Endline	Baseline	Endline	Baseline	Endline
FP long term	269.34	357.70 (P=0 .228)	246.83	261.52 (P=0 .867)	329.97	397.36 (P=0 .327)	82.53	167.99 (P=0 .294)
FP short term	42.43	43.17 (P=0 .928)	20.22 (P=0 .502)	5.30	46.71 (P=0 .617)	41.77	15.49	22.33 (P=0 .249)
ANC1 (before 16wks)	104.64	242.21 (P=0 .000)	22.98	29.69 (P=0 .864)	105.90	210.79 (P=0 .011)	46.63	200.04 (P=0 .055)
ANC4	275.64 (P=0 .916)	270.54	90.36 (P=0 .681)	46.54	298.56 (P=0 .407)	251.00	87.17	183.10 (P=0 .098)
PNC1	5.07	6.61 (P=0 .137)	1.12	1.18 (P=0 .974)	5.016	6.29 (P=0 .311)	2.59	4.01 (P=0 .315)
Skilled Deliveries	25.00	32.47 (P=0 .138)	49.41	55.15 (P=0 .880)	35.30	41.19 (P=0 .503)	12.08	19.99 (P=0 .269)
OPD Under 5	1853.00	2805.51 (P=0 .024)	4105.58	5402.25 (P=0 .535)	2655.58	3754.89 (P=0 .076)	1080.75	1530.30 (P=0 .306)
Total DALYs Averted	7932.79	16677.37 (P=0 .000)	27157.03 (P=0 .618)	19287.70	13353.15	19328.60 (P=0 .126)	5391.27	10022.02 (P=0 .051)

The evaluation assessed the net impact of the project on utilization of selected services by doing DiD analysis. The DID model was fitted using a forward model fitting approach. The STATA 'stepwise' code was employed for the model fitting process, with the Akaike Information Criterion (AIC) being used to select the best-fit model that identifies the most important predictors of the outcome variables. The evaluation team started the model fitting process by using zones as a stratification variable and only considered health centers for the DID analysis. However, since the number of hospitals was insufficient to be analyzed separately, the evaluation team combined both hospitals and health centers for the DID analysis, regardless of their type.

During the model fitting process, the AIC criteria excluded crucial variables such as interaction, time, and intervention status variables from the model. Subsequently, after identifying the crucial confounding variables using the criteria, the evaluation team decided to incorporate the said variables in the models. The team included the intervention status, interaction, and time variables, along with the catchment population in the DID models, which were fitted separately for Jimma and Borena zones.

Additionally, the evaluation team utilized another DID model to measure the overall impact of the project by combining Jimma and Borena as well as West Gojjam zones. Overall, the evaluation team adopted an iterative approach in assessing the impact of the project by incorporating relevant variables to ensure the accuracy and reliability of the analysis.
Table 11 presents the results of the DID regression analysis, which examines the impact of a project on the DALYs averted by various maternal and child health services provided in health centers in Jimma zone.

The results indicate that the intervention had a significant positive impact on ANC1 before 16 weeks (221.9, P=0.043) in health facilities in Jimma zone. However, there was no significant effect on the utilization of short-term and long-term family planning, ANC4, PNC1, or all project-supported services combined (total).

Furthermore, the utilization of long-term FP was found to be significantly lower in the intervention facilities (-1038.3, P=0.003) compared to the control health facilities in Jimma zone. On the other hand, there was a statistically significant increase in the utilization of OPD Under 5 (5293.0, P=0.012) and all health services combined (total) (10329.7, P=0.021) at the endline in Jimma zone compared to the baseline.

The catchment population was found to be the main confounding factor positively affecting utilization of ANC1 before 16wks (0.00503, P=0.013), ANC4 (0.00843, P=0.000), PNC1 (0. 000169, P=0.000), Skilled Deliveries (0.0000686, P=0.000), OPD Under 5 (0.00463, P=0.038), and all health services combined (total) (0.0293, P=0.000).

Table 11: **DID Regression Analysis of the Impact of the Project on DALYs Averted by Health Facility in Jimma Zone**

	DALYs Averted by							
	FP long term	FP short term	ANC1 before 16wks	ANC4	PNC1	Skilled Deliveries	OPD Under 5	Total
Time (Endline=1)	-379.4 (P=0.301)	33.56 (P=0.274)	-94.35 (P=0.304)	-114.9 (P=0.105)	0.448 (P=0.588)	-5.696 (P=0.531)	5293.0* (P=0.012)	10329.7* (P=0.021)
Intervention status (Intervention=1)	-1038.3** (P=0.00	-32.22 (P=0.273)	-140.1 (P=0.101)	-62.31 (P=0.356)	0.788 (P=0.324)	-7.140 (P=0.389)	-1074.4 (P=0.546)	451.9 (P=0.909)

	3)							
Interaction (Intervention*Time)	411.0 (P=0.374)	-36.46 (P=0.363)	221.9* (P=0.043)	65.68 (P=0.469)	1.404 (P=0.194)	13.97 (P=0.221)	-4323.3 (P=0.080)	-5602.7 (P=0.311)
Catchment Population	-0.000057 (P=0.900)	0.000826 (P=0.311)	0.00503* (P=0.013)	0.00843*** (P=0.000)	0.000169*** (P=0.000)	0.0000686*** (P=0.000)	0.00463* (P=0.038)	0.0293*** (P=0.000)
_cons	1376.8*** (P=0.000)	59.20 (P=0.057)	87.59 (P=0.332)	121.5 (P=0.087)	-0.785 (P=0.344)	32.25*** (P=0.000)	3042.6* (P=0.044)	8542.3*** (P=0.008)
N	54	48	39	47	47	53	49	54

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12 presents DID analysis on utilization of selected health services and all project supported services combined measured in DALYs averted for Borena zone. The results show that the project had a positive impact on long-term family planning utilization (934.1; P=0.007), indicating an increase in the utilization of this service. The program did not have a significant impact on the utilization of FP short term, ANC1 before 16 weeks, ANC4, PNC1, OPD under 5 or all services combined (total DALYs averted) in Borena zone.

Table 12: **DID Regression Analysis of the Impact of the Project on DALYs Averted by Health Facility in Borena Zone**

	DALYs Averted by Key Service							
	FP long term	FP short term	ANC1 before 16wks	ANC4	PNC1	Skilled Deliveries	OPD Under 5	Total
Time (Endline=1)	-848.7** (P=0.003)	15.91* (P=0.049)	-28.36 (P=0.756)	-31.22 (P=0.695)	-0.473 (P=0.789)	2.833 (P=0.757)	-272.9 (P=0.704)	-417.4 (P=0.911)
Intervention status (Intervention=1)	-534.5 (P=0.051)	-1.107 (P=0.883)	-32.93 (P=0.719)	-46.49 (P=0.562)	-2.509 (P=0.189)	-11.12 (P=0.238)	-1170.9 (P=0.133)	-7857.0 (P=0.058)
Interaction (Intervention*Time)	934.1** (P=0.007)	-9.073 (P=0.331)	181.8 (P=0.120)	127.1 (P=0.206)	1.896 (P=0.397)	5.075 (P=0.652)	722.5 (P=0.424)	5048.1 (P=0.287)
Catchment Population	0.00505** (P=0.002)	0.000114* (P=0.016)	0.000757 (P=0.155)	0.000758 (P=0.106)	-0.00000880 (P=0.378)	-0.0000214 (P=0.678)	-0.00340 (P=0.400)	-0.0116 (P=0.577)
_cons	534.2* (P=0.035)	14.73* (P=0.049)	67.14 (P=0.430)	121.2 (P=0.116)	5.241** (P=0.007)	23.55* (P=0.014)	2307.4** (P=0.004)	13438.9** (P=0.002)
N	19	18	18	18	19	18	19	19

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 13 shows the project level DID analysis results for all intervention zones and types of health facilities. The results show that the intervention had a significant positive impact on ANC1 service utilization. The statistically significant coefficient of the interaction variable for DALYs averted for ANC1 before 16 weeks (188.7; P=0.017) suggests that the intervention had a positive impact on this particular health service. However, the project did not have a significant impact on family planning long-term (504.9; P=0.138) and short-term (-28.62; P=0.339), ANC4 (63.61; P=0.469), PNC1 (P=1.359; 0.321), skilled deliveries (10.35; P=0.251), OPD Under-5 (-1937.5; P=0.270) and all services combined (-135.3; P=0.974).

Additionally, the results show a significant decrease in utilization of long-term family planning services (-906.2, P=0.000) in the intervention group, compared to the control ones. Compared to baseline, utilization of OPD Under-5 grew statistically significantly at endline (2934.0, P=0.040).

Table 13: DID Regression Analysis of the Impact of the Project on DALYs Averted in health facilities and zones

	DALYs Averted by Key Service							
	FP long term	FP short term	ANC1 before 16wks	ANC4	PNC1	Skilled Deliveries	OPD Under 5	Total
Time (Endline=1)	-428.1 (0.107)	26.52 (0.254)	-57.82 (0.367)	-78.95 (0.251)	0.101 (0.923)	-3.135 (0.662)	2934.0* (0.040)	6095.6 (0.067)
Intervention status (Intervention=1)	-906.2*** (0.000)	-22.43 (0.298)	-47.89 (0.415)	3.134 (0.961)	1.493 (0.142)	-5.066 (0.437)	-677.5 (0.597)	-872.0 (0.775)
Interaction (Intervention*Time)	504.9 (0.138)	-28.62 (0.339)	188.7* (0.017)	63.61 (0.469)	1.359 (0.321)	10.35 (0.251)	-1937.5 (0.270)	-135.3 (0.974)
Catchment Population	-0.000027 (0.946)	-0.000043 (0.833)	0.00086 (0.071)	0.0011 (0.070)	0.000014 (0.066)	0.000071*** (0.000)	0.0040* (0.038)	0.0283*** (0.000)
_cons	1174.7*** (0.000)	68.09*** (0.000)	125.9* (0.025)	243.9*** (0.000)	3.122*** (0.001)	26.16*** (0.000)	2462.1* (0.020)	8923.2*** (0.000)
N	79	70	61	69	72	75	74	79

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In conclusion, the PBF program had a significant positive impact on ANC1 service utilization, specifically ANC1 before 16 weeks in all zones and types of health facilities. However, the program did not have a significant impact on the utilization of other health services, including family planning long-term (P=0.138) and short-term (P=0.339), ANC4 (P=0.469), PNC1 (P= 0.321), skilled deliveries (P=0.251), OPD under 5 (P=0.270), and all services combined (P=0.974), except for PNC1 in Jimma zone. Moreover, the program had a positive impact on long-term family planning utilization in Borena zone.

In addition to the quantitative analysis presented above, the evaluation drew on qualitative data collected from patients and various stakeholders involved in the implementation of the performance-based financing (PBF) program.

Qualitative data analysis results also suggest that the PBF program has made a significant difference in improving access to and utilization of quality health services in the target communities. In terms of access, the intervention health facilities were able to improve availability of medicine and laboratory tests, which, according to most patients, used to be very limited in the past. When medicine and laboratory tests become available, the community and particularly vulnerable segments of the population were able to have access to such services at nearby facilities. The program has contributed to significant changes in service accessibility, including the provision of laboratory services within health facilities, which has led to a decrease in referrals to other facilities. According to key informants from intervention health facilities, patients' complaints have declined as they have been able to improve access to quality health services. An informant said the following in connection with improved access to health services.

> *"The implementation of PBF program in this health facility significantly increased the quality of health services. Before the implementation of PBF project there was no any laboratory service in this compound, and the client goes to Yabello for laboratory services. Nowadays the lab located in our health facility gives all service demanded and there is no case that we refer to Yabello for lab testing. This a significant change that we are obtaining due to the implementation of PBF program."*

The PBF program also improved utilization of health care services by enhancing the capacity of health facilities and their health workers, improving coordination, making health facilities comfortable for patients, motivating health workers through incentives, improving availability of medical equipment and medicine, strengthening laboratory services, and encouraging healthcare providers' community outreach to provide community health education and health services, among others. This, according to many key informants, improved patients' satisfaction which also led to increased utilization of health services by the community. Here are some notable statements from some key informants which seem to be shared opinions.

> *"There have been significant positive changes in the intervention health facilities in regards to utilization rate over the course of the implementation period. I think the changes are happening because of PBF project."*

> *"...health facilities have become more attractive. It has increased the desire of patient to use our facility. They have previously visited private clinics and raised complaints about our services. Now that they have access to quality and adequate services, their demand has increased. Since the advent of PBF, there are adequate medical equipment, adequate medicines and adequate health services. Therefore, the community's desire to utilize services in health facilities has increased."*

Overall, qualitative findings suggest that the PBF program has contributed to the improvement of primary health units, and the enhancement of service access and utilization. Quantitative findings also indicated that volume of health services has increased in both control and intervention facilities between the baseline and endline. However, the percent increase in volume in intervention facilities doesn't appear to be significantly different ($p>0.05$) than control facilities.

3.4.1.2. Increased Equity and Gender Equality in Access to Health Services

Increasing equity in access to health services was also another desired outcome of the PBF program. To that end, the program paid an equity bonus to those health centers which are very remote or otherwise work under difficult circumstances.

In relation to geographic inequity, remote or poorer areas are generally at a disadvantage to reach targets and receive PBF rewards. The PBF program provided 'equity bonuses' as a means to address issues of geographic access that could affect service delivery.

According to patients and key informant responses, the PBF program has had a positive impact on promoting equity in access to health services. The program has improved the availability and quality of health services, especially in areas that previously faced challenges such as shortages of laboratory equipment, reagents, and essential medicines. One key informant highlighted the improvements brought about by the program, saying, *"Before the implementation of the project it used to be very challenging to get even laboratory services at health centers due to shortage of laboratory equipment and reagents. …there was also a significant shortage of essential medicine...the community and particularly poor families didn't get appropriate health services. However, nowadays, there is a significant change in health service provision for the community."*

Another informant noted that the PBF program has helped the more vulnerable communities to access health services by fulfilling the essential medicine requirements and improving the behavior of service providers. The informant explained that previously, lack of money and unavailability of essential medicines, particularly for communicable diseases, prevented people from seeking healthcare. However, the informant observed that *"nowadays, the number of women frequently seeking health care services, such as delivery and family planning, has shown improvement."*

The program was also instrumental in making affordable medicines available to the community. For instance, a patient remarked that, *"…before two years access to medicine was very poor. However, this year I got all prescribed medicines here without any effort."* The improved availability of medicines at health facilities allowed vulnerable people to purchase them at affordable prices. By making medicines available, poor people were able to obtain drugs for free. Those with medical insurance coverage were also able to access medicines and other health services, which used to be difficult in the past.

Overall, the responses suggest that the PBF program has contributed to improving equity in access to health services. The program's efforts to address shortages of essential medicines and laboratory equipment have had a positive impact on vulnerable communities. But some stakeholders believe that the equity matters were not adequately addressed in the program design. An informant, for example, noted that equity in access to health services is a complex issue that cannot be fully addressed by the PBF program. The informant emphasizes that the *"the PBF program is not a panacea and cannot address everything within the health system, including equity in access. It contributes to addressing some of the challenges …within the health system."* However, the program implemented different pricing strategies for pastoralist and agrarian areas, taking into account the differences in accessibility, population density, and service uptake. For instance, lower unit prices were used for quantity indicators in the Jimma Zone (an agrarian area), reflecting higher population density, a higher uptake of services, and less remote facilities compared to the Borena Zone. This differentiated pricing approach aims to address the differences and promote equity in service delivery.

The project also contributed to efforts aimed at addressing gender disparities in health. The government, in its HSTP 2[17], acknowledges gender disparities in health service utilization that it linked to the women's limited decision-making power at the household level. Factors such as needing permission to visit a health facility, obtaining money for treatment, distance to a health facility, and unwillingness to go to a health facility alone are important barriers to women's health service utilization. The responses of patients who were consulted for this evaluation also reveal various challenges and barriers to accessing healthcare services, particularly for women and children. Many patients highlighted the challenges of distance and transportation, with some having to walk for hours or pay expensive transportation fees to access healthcare services. A patient noted, *"We walk 25 to 30 hours in foot to come here...It costs us 50 birr."* This is especially challenging for women who repeatedly come for maternal services, as noted by a patient who said, *"Nowadays, the transportation fee is very expensive. This is very challenging for women who repeatedly come here for maternal services."*

The PBF program provided financial incentives to health providers and facilities for meeting specific targets related to service utilization and quality. The program was designed to address gender disparities in healthcare by targeting specific services that are important for women and children, such as maternal and child health services. A key informant noted, *"The project was designed in ways that address the key challenges faced by the health facilities and other relevant stakeholders in the health system. The interventions were significant to improve utilization of care and quality of care. The project was also relevant in improving equity in access and gender equality."*

The findings suggest that the PBF program has contributed to efforts to address gender disparities in health by ensuring quality health services accessible to the community in nearby locations, increasing women's health service utilization, supporting efforts to deal with socio-cultural and economic barriers, improving the capacity of health care workers, enforcing government laws and policies related to gender equality and human rights, and improving the capacity of health

[17] Federal Democratic Republic of Ethiopia Ministry of Health. (2021). Health Sector Transformation Plan II 2020/21-2024/25. Addis Ababa, Ethiopia.

facilities for providing quality health care services. These contributions of the project to gender equality are highlighted below.

- **Improved access to quality health services:** The PBF program has improved access to health services, particularly for women and marginalized communities. One informant noted that *"The project gives more attention to maternal health, therefore women benefit more from this project than men. As a result, the coverage of delivery, antenatal care (ANC), postnatal care (PNC), and family planning has increased."* The program has provided financial support to health facilities, enabling them to purchase essential medical equipment and medicines. This has helped to fill the gaps in health facilities, making it easier for women and men to access health services nearby their homes. An informant highlighted *"Nowadays, no women or men are turned away from the facility due to a lack of service, except in cases beyond our capacity."* Patients also expressed satisfaction with the quality of care provided at the health center, with some stating that the PBF program had made healthcare services more accessible and affordable. One patient noted, *"I'm very satisfied because it made the healthcare service more accessible for me."* The improvements in infrastructure and health services made possible through the project support also encouraged pregnant women to deliver at health facilities. One patient noted, *"Almost all pregnant women receive follow-up and treatment. However, when the time for delivery comes, some give birth at home or on the way because they cannot reach the health facility in time. But now, distance is no longer a barrier and no one has to suffer because of it."* This, according to the patient, is because *"If the time of delivery is approaching, they [patients] would come to the health centre here and wait for the [delivery] time. There is a place to stay. There is food. …I am very happy that such a service exists."* This indicates that the PBF program has improved infrastructure of health facilities and access to health services to pregnant women, which has reduced the risks associated with giving birth at home or on the way.

- **Increased women's health service utilization**: Several informants noted that the PBF program has had a positive impact on increasing women's utilization of health services, particularly for maternal and reproductive health. Women are now more willing to access health facilities due to improvements in the quality of health services and an increased awareness of the importance of maternal and child health services. For example, one informant stated that *"it significantly increased the MCH and under five treatments."* Another informant highlighted that *"the only thing that changed is the flow rate of women to health facility to get different health services."* Similarly, an informant stated *"the coverage of delivery, ANC, PNC and family planning has increased."* The program has focused on enhancing the quality of service provision for women, with improvements to delivery and ANC rooms having a significant impact on service provision, with one informant stating that *"the experience of women in giving birth, ANC and PNC follow up has been increasing from time to time following the improvement of the quality of health service care."* Another informant highlighted that *"the most significant change that we have been seen due to the implementation of this project is the change in quality of health care following the fulfilment of essential medicine and medical equipment... Due to these, nowadays, the flow rate of women for delivery, ANC and even the OPD cases are increasing."*

- **Contributed to efforts to address sociocultural and economic barriers**: There are still sociocultural barriers that hinder communities from accessing health services, particularly for women giving birth at home due to the unwillingness of husbands to send their wives to health facilities. However, the program has worked on awareness creation to change the attitudes of the society towards maternal and child health services. A key informant highlighted, *"Before the implementation of the PBF program, the number of deliveries at health facilities was very low due to the unwillingness of husbands to send their*

wives to health facilities. This issue has been significantly addressed due to the project's work on awareness creation." Key informants also noted that the PBF program has had a significant impact on improving access to health services for vulnerable communities, particularly in terms of providing essential medicines and enhancing the behavior of service providers. One informant stated, *"Most of the time patients didn't come to health facility due to lack of money and unavailability of essential medicines."* However, the program has worked to address these issues and has improved access to health services for all, regardless of gender or socio-economic status. Patients also reported that the PBF program reduced the cost of seeking healthcare services and made healthcare more accessible for them. A patient stated, *"I come to this hospital because I can get service at low price relative to the private ones and it is the nearest."* Some informants also noted that the project has contributed to increased engagement of women in community activities. One informant stated that *"the project has engaged women …at different levels like, health facility community boards, management committees..."*

- **Contributed to enforcement of existing laws and policies on the rights of women and girls:** According to HSTP 2, among the major challenges in the addressing gender disparities in health in Ethiopia are limited enforcement of existing laws and policies on the rights of women and girls. Several informants noted that the PBF intervention has contributed to national legislation and initiatives that aim to improve gender equality and human rights by providing health care services in an equitable manner without discrimination. One informant noted that *"the financial support that has been received from PBF project had a significant role in the enhancement of health service from which women and men equally benefited."* The informants noted that the program serves both genders equally without any discrimination. In fact, the program focuses on the vulnerable group, particularly mothers, to increase access and utilization of health services. Additionally, the program has enforced government laws and policies, such as requiring separate latrines for women and men and delivery rooms with functional showers, to improve the equality of access to health services for both men and women. Several informants also noted that the PBF program generates real data at health care facilities, which policy makers can use as input to make decisions. One informant stated that "*it may support policymakers, legislation, and initiatives through real information gathering to identify gaps in gender issues. Policymakers can use that information as an input to make decisions."* The key informants also noted that the PBF program's strong monitoring and evaluation system promotes accountability and transparency, which supports national legislation and initiatives aimed at improving gender equality and human rights. According to an informant, *"When you get quality data and real information about what's going on the ground, you can easily prepare policy."*

- **Improved capacity among health care workers**: The PBF program has improved the capacity of healthcare workers in providing gender-responsive health services. Health workers are now more committed to providing quality health services, and there has been a fundamental change in their willingness to provide health services. One informant stated that *"the intervention supports national legislation and initiatives that aim to improve gender equality and human rights by providing healthcare services through skilled and trained healthcare providers in health facilities and health posts, while ensuring that services are provided in an equitable manner without any discrimination."* Another informant also highlighted that *"the contribution of PBF project in improving the capacity of the health service to provide appropriate service for women is very high."* Patients also stated that the health workers are treating them respectfully, friendly, and with good care. One patient stated, *"They are so sociable and have good approach and compassionate and respectful in providing service."*

- **Improved capacity of health facilities**: The PBF program has improved the capacity of health facilities for providing quality health care services. The program has helped to reduce the shortage of medicine and medical equipment, which has enabled health facilities to provide better quality health services. A key informant noted, *"…The project supports health facilities through finance and health workers through incentives. These supports help facilities acquire necessary equipment and resolve shortages of medicine, raising the quality of health services. The availability of these inputs allows women and men to access health facilities more easily."* Another informant highlighted that *"the most significant change that we have encountered after the implementation of the PBF project is the fulfillment of necessary medical equipment, furniture, and the repair of the health facility's compound."* Informants reported that the PBF program had played a great role in improving the quality and privacy of the delivery services, which significantly increased the flow rate of women to health facilities. An informant, for example, stated, *"For women, the project creates a conducive consultation room and allows for easy access to maternal health services."*

Overall, he PBF program has made some contributions to addressing disparities in access to health services by providing financial incentives to health providers and facilities for meeting specific targets related to service utilization and quality. The findings suggest that the PBF program has improved the quality and accessibility of health services for the entire society, particularly for women. The program has contributed to efforts to address gender disparities in health by ensuring quality health services accessible to the community in nearby locations, increasing women's health service utilization, improving the capacity of health care workers, enforcing government laws and policies related to gender equality and human rights, and improving the capacity of health facilities for providing quality health care services. However, challenges still exist with regards to lack of road infrastructure, high transportation costs, distance, and socio-cultural barriers. Many patients reported shortages of medicines at health facilities, which affected their ability to access healthcare services. One patient remarked, *"Most of the time, we do not get medicine prescribed from the health center. There is always a shortage. We are villagers; we don't go anywhere."*

3.4.1.3. Improving quality of health services

The endline evaluation assessed the impact of the program on quality of health care provided by health facilities. Quality of service was measured through facility-based technical quality of care assessment, chart audit, and patient satisfaction surveys. Findings are presented under each below.

i. Technical Quality of Care at Health Centers

Another key desired outcome of the PBF program was improving access to quality health services. Table 14 presents the mean scores obtained by health centers on 16 different technical quality of care domains, including admin & planning, ANC, emergency, EPI, general appearance, general OPD, HMIS & supervision, IC & waste management, inpatient, laboratory, logistics & med supplies, maternity service, nutrition, outreach, referral, and under 5 OPD at baseline and endline, between control and intervention health centers

The results show that intervention facilities have generally higher quality of care scores compared to control facilities, both at baseline and endline. It is also important to note that mean quality of care scores on almost all domains (except nutrition domain for control facilities) grew statistically significantly ($p<0.05$) for both intervention and control health facilities. However, the magnitude of difference is very large for intervention facilities (from 24 at baseline to 81 at endline)

compared to control facilities whose score grew from 16 to 50. In the nutrition domain, average quality of care scores of the intervention facilities statistically significantly ($p<0.05$) compared to baseline, while this was not the case in the control facilities. The results suggest that technical quality of care improved more in the intervention facilities than in the control ones.

When comparing the scores among zones, we can see that West Gojjam has the lowest scores for most domains in the intervention facilities. Jimma and Borena generally have higher scores, with intervention facilities in these zones having higher scores than their control counterparts. This could be partly because West Gojjam intervention started recently compared to Jimma and Borena facilities who started receiving the program support early. Looking at the data, we can also see that the scores vary among domains in the different zones. For example, in the ANC domain, the Borena and West Gojjam intervention facilities had the highest score at baseline and endline in both intervention and control facilities, while the Jimma control facilities had the lowest score at baseline as well as endline. The same trend can be observed for other domains as well.

Table 14: Average quality score (health centers) by intervention zone, domain and term

	Jimma				Borena				West Gojjam				Total			
	Intervention		Control		Intervention		Control		Intervention		Control		Intervention		Control	
	Baseline	Endline	Baseline	Endline	Baseline	Endline	Baseline	Endline	Baseline	Endline	Baseline	Endline	Baseline	Endline	Baseline	Endline
Admin & Planning	7	86**	8	50**	52	93**	16	67**	5	80**	17	80	19	87**	13	56**
ANC	36	93**	38	81**	36	100**	35	100*	54	100	75	100	38	96**	42	86**
Emergency	1	79**	0	7	17	78**	0	39**	0	33	0	67	5	74**	0	18**
EPI	25	93**	20	49**	45	97**	28	86**	13	96*	26	83	29	94**	25	60**
General Appearance	16	63**	8	19*	42	39	19	39*	6	44**	12	12	22	55**	14	23*
General OPD	18	54**	11	27*	34	71*	19	50**	0	79*	3	50	20	61**	14	33**
HMIS & Supervision	13	71**	21	51**	50	80	13	63**	10	65*	0	70	23	73**	14	55**
IC & Waste Management	9	74**	10	29*	47	81*	4	65**	0	85**	0	52	18	77**	6	39**
Inpatient	2	47**	4	10	15	80**	4	47**	0	20	0	0	6	54**	3	17*
Laboratory	21	82**	30	58**	49	83*	13	72**	15	81*	40	77	29	82**	24	63**
Logistics & Med Supplies	34	98**	35	66**	55	97*	24	57*	10	90	30	100	38	97**	30	66**
Maternity Service	11	90**	8	42**	36	82**	12	59**	61	83	0	50	17	87**	9	46**
Nutrition	26	69**	35	48	87	97	47	87	30	60	33	60	43	75**	40	57
Outreach	0	91**	16	48*	18	83**	0	42**	0	88*	0	100	5	89**	6	50**
Referral	20	71**	0	62**	56	53	0	44**	8	67*	0	67	29	66**	0	58**
Under 5 OPD	33	86**	16	46*	70	87	21	60**	30	80	21	50	43	85**	19	49**
Total Score	18	81**	17	44**	44	82**	16	63**	11	78**	16	64	24	81**	16	50**

* $p < 0.1$, ** $p < 0.01$ and $p < 0.001$

A difference-in-difference analysis was also done to assess the effect of the project on technical quality of care score. The results presented in Table 15 show that both the control and treatment/ intervention health centers experienced an increase in mean technical quality of care scores on different quality domains from baseline to endline. The regression analysis result shows that the intervention health centers generally exhibit a greater increase in quality scores compared to the control group in most of the domains except for ANC, HMIS Supervision, Laboratory, Nutrition and Under5 OPD, indicating a positive impact of the intervention on technical quality of care.

Table 15: Difference-in-difference analysis of technical quality of care score at health centers at baseline and endline as well as intervention and control facilities by quality domain

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
	Admin Planning	ANC	Emergency	EPI	General Appearance	General OPD	HMIS Supervision	IC Waste Mgmt	Inpatient	Laboratory	Logistics Med Supplies	Maternity Service	Nutrition	Outreach	Referral	Under5OPD	Total Quality Score
Intervention	7.186	-3.527	5.303	5.753	9.039	6.169	8.151	13.00*	2.432	5.393	9.632	9.058	0.823	-1.028	28.79***	23.20***	8.282*
	(1.19)	(-0.46)	(0.95)	(1.10)	(1.67)	(1.12)	(1.17)	(2.29)	(0.27)	(1.03)	(1.63)	(1.98)	(0.08)	(-0.18)	(5.15)	(3.55)	(2.40)
Time	43.81***	45.19***	17.86**	36.11***	10.22	19.05**	40.03***	33.12***	13.97	39.51***	38.33***	37.92***	15.24	44.05***	58.33***	29.64***	33.74***
	(6.40)	(5.16)	(2.82)	(6.12)	(1.67)	(3.05)	(5.08)	(5.15)	(1.38)	(6.67)	(5.74)	(7.34)	(1.32)	(6.89)	(9.22)	(4.01)	(8.65)
Time* Intervention	24.37**	13.07	51.08***	28.85***	22.93**	21.48*	9.577	25.37**	34.06*	13.89	20.76*	32.28***	17.49	39.66***	-21.21*	12.97	23.33***
	(2.68)	(1.12)	(6.07)	(3.68)	(2.82)	(2.59)	(0.91)	(2.97)	(2.54)	(1.75)	(2.34)	(4.70)	(1.14)	(4.67)	(-2.52)	(1.32)	(4.50)
_cons	11.90***	41.07***	-3.55e-14	23.41***	12.88**	14.28***	14.97**	5.384	3.174	23.13***	28.10***	8.113*	41.90***	5.952	1.14e-13	19.64***	16.07***
	(2.75)	(7.41)	(-0.00)	(6.27)	(3.33)	(3.62)	(3.00)	(1.32)	(0.50)	(6.18)	(6.65)	(2.48)	(5.74)	(1.47)	(0.00)	(4.20)	(6.51)
N	79	79	79	79	79	79	79	79	79	77	79	79	79	79	79	79	79

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

ii. Technical quality of care at Hospitals

Table 16 presents the mean quality of care scores obtained by each intervention and control hospitals at baseline and endline for different quality of care domains.

Looking at the results, we can see that the mean quality of care scores for the intervention hospitals (those receiving the PBF subsidy) increased significantly in all domains (except laboratory) between baseline and endline assessments. The control hospitals (those not receiving the PBF subsidy) also showed some improvement, but not to the same extent as the intervention hospitals.

Specifically, in the intervention hospitals, the mean scores for all quality of care domains increased from 21.7 at baseline to 88.3 at endline. In contrast, in the control hospitals, the mean scores for most domains increased slightly or remained relatively stable, with the exception of Administration and Planning, HMIS and Supervision, IC and Waste Management, and Maternity, where the mean scores increased substantially. But only mean score achieved by control hospitals on IC & waste management domain grew statistically significantly ($p<0.05$) from baseline. When comparing the mean scores between intervention and control hospitals, we can see that the intervention hospitals generally had higher scores in all domains at endline.

Overall, the data suggests that the PBF program might have had a positive impact on the quality of care in intervention hospitals as evidenced by the significant ($p<0.05$) improvements in most quality of care domains and overall score.

Table 16: Average Technical Quality Score (Hospitals) at Baseline and Endline

	Hospital			
	Control		Intervention	
Domains	Baseline	Endline	Baseline	Endline
Gen Appearance	29	46	25	85*
Admin & Planning	21	71	20	100*
HMIS & Supervision	5	70	16	94*
IC & Waste Mngt	5	48*	1	67*
General OPD	17	33	18	92*
Maternity	32	72	16	94*
EPI & GM	40	65	10	94*
ER Services	25	56	28	98*
Inpatient	17	39	16	78*
Surgery	53	88	53	88*
Laboratory	58	54	53	85*
Radiology	5	25	14	76*
Logistics & Meds	35	73	21	100*
Total Score	26.75	58.75	21.7	88.30*

**statistically significant at p<0.05*

Qualitative data gathered from patients and key informants also suggested that the program has had a significant impact in improving the quality of health services in the health facilities where it was implemented, with health facilities experiencing improvements in the standard of care. The program has been successful in addressing issues related to financing, infrastructure, and availability of medicines and equipment. The program has also been effective in improving the capacity of health facilities and health workers, and motivating them to provide quality services to the community. It helped them make quality health services available and accessible to their

communities. PBF has complemented the government's efforts to improve the health system by addressing gaps in procurement, infrastructure, and human resources.

According to key informants, the subsidies provided by the program increased availability of medical equipment and supplies, laboratory services, medicine, water, cleaner examination rooms making health facilities more comfortable and attractive. These helped health facilities provide quality service. The program support also strengthened capacity of health facilities and health workers to provide quality services. Some notable statements shared by key informants are presented below.

> *"The most significant challenge that has been addressed with the implementation of this project are the availability of essential medicine and medical equipment which helped our facility increase number of maternity health services. The compound of this health facility has been completely changed and become more attractive, the health service provided for the community has also improved as compared to the time before the implementation of this project."*
>
> *"The most significant change that we have seen at health facilities are the increase in maternal services like family planning, delivery and ANC. In addition, infection prevention was not also simple. Previously the health facility did not have a fence and the compound was open. Nowadays, the compound is fenced and clean and attractive. Not only the compound, but the rooms are also very clean and attractive for clients. The number of chairs in treatment and consulting rooms also improved as compared to the previous time. Necessary instruments, medical equipment, and availability of medicine at health facility also show improvement and that led to quality health service provision."*
>
> *"The difference that come from PBF has providing quality of care and providing service for infection prevention by maintaining three bin system."*

Overall, stakeholders, including patients, have a positive view of the PBF program and its importance for improving the quality of health care services in the target areas. The PBF program was seen as a valuable addition to the existing health system in the intervention areas. Patients expressed satisfaction with the quality of services they received at the health centre/hospital, citing privacy, confidentiality, being listened to, getting the support/service they/their child needed, and getting necessary medicine as positive experiences. Patients also expressed that they would seek health care services at the intervention hospital/health centre in the future if a need arises and would refer a family/friend to this facility for health service. They also noted areas for improvement, such as inadequate availability of medicines and water which exposes them to additional costs, poor latrine and delivery room cleanliness, shortage of health professionals, and others.

iii. Chart Audit Index

Chart audits were conducted on selected services, including IMNCI Admission and ambulatory Diarrhoea management, IMNCI Admission and Ambulatory Pneumonia Management, Normal Deliveries Management and Hypertension checking health worker compliance to nationally and internationally acceptable standards. A total of 127 and 106 chart audits in intervention and control health facilities were done, respectively, to assess clinical completeness and determine if clinicians adhere to clinical guidelines.

To assess adherence levels, the evaluation team calculated service management adherence score by dividing the obtained scores for each chart by the total possible as per the service protocol taking into consideration procedures not applicable for each patient whose chart is reviewed.

Table 17 shows the mean scores of chart audits conducted focusing on six services to evaluate adherence of clinicians to treatment guidelines in different zones (Jimma, Borena, and West Gojjam) in both intervention and control groups. The scores range from 0 to 1, with higher scores indicating better adherence to guidelines.

Table 17: Quality adherence index by service category, intervention status and zone at endline

	Chart Audit Mean Score by Zone and Intervention Status at Endline							
	Jimma		**Borena**		**West Gojjam**		**Total**	
	Intervention	**Control**	**Intervention**	**Control**	**Intervention**	**Control**	**Intervention**	**Control**
Ambulatory Diarrhoea	0.44	0.38	0.55	0.38	0.54	0.44	0.48*	0.39
Ambulatory Pneumonia	0.57	0.52	0.58	0.54	0.6	0.41	0.58	0.51
Normal Deliveries	0.67	0.61	0.72	0.74	0.81	0.77	0.71	0.65
Diarrhoea Admissions	0.48	0.63	0.72	0.48	0.72	0.68	0.58	0.6
Hypertension Admissions	0.55	0.45	0.56	0.34	0.53	0.58	0.55*	0.45
Pneumonia Admissions	0.48	0.84	0.88	0.76	0.83	0.76	0.6	0.78

* $p < 0.05$

Source: Chart audit conducted for the end of program evaluation of PBF program

The results show that the mean scores for most services in the intervention group are slightly higher than those in the control group. The results also suggest that clinicians in intervention facilities that provided ambulatory Diarrhoea and hypertension admissions services were more likely to adhere to treatment guidelines than control groups. These difference were statistically significant at $p<0.05$.

When comparing the mean scores by zone, we see some variations. For example, the mean scores for most services (Ambulatory Diarrhoea, Ambulatory Pneumonia, Diarrhoea Admissions and Pneumonia Admissions) in Borena and West Gojjam are much higher than those in Jimma zone. The MTR also found higher scores in Borena intervention facilities in clinical management of ambulatory diarrhea, delivery and ambulatory pneumonia, compared to Jimma. Given the recent enrolment of West Gojjam zone facilities, however, the higher chart audit score found in the intervention facilities may suggest that there may be some local factors influencing adherence to treatment guidelines. But it is important to note here that the difference was not statistically significant ($p>0.05$).

In terms of specific services, the mean scores for normal deliveries are generally higher than those for other services, indicating better adherence to guidelines for this service across all zones and intervention status. On the other hand, the mean scores for Ambulatory Diarrhoea are generally lower, indicating poorer adherence to guidelines for this service.

Overall, the results indicate that the level of adherence to treatment guidelines is generally low although there are variations by service and zone. Across most services and zones, the mean scores are slightly higher in the intervention group than in the control group. Adherence rate is statistically significantly higher for ambulatory Diarrhoea and hypertension admissions services in intervention facilities than control ones. For other services, the differences between intervention and control groups were not statistically significantly. Overall, the results may suggest that there is either a capacity gap in service provision on the selected services or there is poor documentation of clinical records by health care providers.

The evaluation also assessed the impact of the project on adherence using DiD analysis. The results of the DID analysis in Jimma and Borena zones[18] are presented in Table18 below.

The results show that compared to control facilities, adherence index in the intervention facilities in Jimma zone was statistically significantly higher for ambulatory Diarrhoea management (9.211, P= 0.000), ambulatory Pneumonia management (5.347, P=0.039), normal deliveries management (19.92, P = 0.000) and hypertension (7.245, P = 0.000). In the Borena zone, adherence scores for Ambulatory Pneumonia (7.014, P=0.006), deliveries (15.33, P =0.007) and hypertension (14.47, P=0.000) were found to be significantly higher in the intervention facilities compared to the control ones.

Compared to baseline, other factors kept constant, the endline adherence index was statistically significantly higher than the baseline for deliveries (19.46, P = 0.000), and hypertension (21.80, P = 0.000). However, a statistically significant decline was observed in the ambulatory Diarrhoea score in the Borena zone (-28.32, P = 0.000).

The DID analysis was conducted by combining Jimma and Borena Zones. The result of the analysis revealed that the intervention health facilities had significantly higher adherence index compared to their control counterparts for all selected services: Ambulatory Diarrhea (7.428; P=0.001), Ambulatory Pneumonia (6.709, P=0.000), Deliveries (18.71, P=0.000), and Hypertension (12.33, P=0.000).

However, the DiD analysis results show that interaction variable (Intervention*Time) was not significant for any of the services, indicating that the changes in adherence index over time were not significantly different between the intervention and control groups.

Overall, these results suggest that the intervention had a positive effect on the chart adherence index, but this effect was not significantly different between the intervention and control facilities.

Table 18: DID regression analysis of the impact of the project on adherence to clinical guidelines in health facilities in Jimma and Borena Zones

	Jimma				Borena				Total			
	Ambulatory Diarrhea	Ambulatory Pneumonia	Deliveries	Hypertension	Ambulatory Diarrhea	Ambulatory Pneumonia	Deliveries	Hypertension	Ambulatory Diarrhea	Ambulatory Pneumonia	Deliveries	Hypertension
Time (Endline=1)	-0.951 (0.830)	3.406 (0.498)	19.46*** (0.000)	21.80*** (0.000)	-28.32*** (0.000)	1.055 (0.887)	24.84 (0.113)	-16.25 (0.101)	-10.30* (0.013)	0.498 (0.895)	21.90*** (0.000)	14.13*** (0.001)
Intervention status (Intervention=1)	9.211*** (0.000)	5.347* (0.039)	19.92*** (0.000)	7.245*** (0.000)	2.156 (0.417)	7.014** (0.006)	15.33** (0.007)	14.47*** (0.000)	7.428*** (0.001)	6.709*** (0.000)	18.71*** (0.000)	12.33*** (0.000)
Interaction (Intervention*Time)	-3.394 (0.581)	-0.266 (0.969)	-13.06 (0.080)	3.399 (0.454)	15.65 (0.115)	-2.471 (0.797)	-17.15 (0.395)	8.264 (0.516)	1.505 (0.787)	0.606 (0.905)	-13.11 (0.062)	-2.244 (0.707)
_cons	39.34*** (0.000)	48.88*** (0.000)	41.14*** (0.000)	22.99*** (0.000)	65.81*** (0.000)	52.89*** (0.000)	48.71*** (0.000)	49.96*** (0.000)	49.49*** (0.000)	50.15*** (0.000)	43.47*** (0.000)	31.03*** (0.000)
N	220	246	260	256	139	149	130	126	368	404	399	391

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

iv. Patient surveys on quality of service and satisfaction with health services

[18] DiD analysis could not be done for West Gojjam zone due to unavailability of baseline/MTR data

In this section, the results of the patients' survey pertaining to their perception of quality of service they received from the health facilities they visited are presented focusing on key selected services namely: ANC, PNC, under-5 children suffering from Pneumonia and Diarrhoea, and OPD.

ANC service

Table 19 presents the analysis about the quality of antenatal care services provided to patients in the intervention and control groups.

The percentage of clients whose medical history was recorded in the antenatal-care card/book was significantly ($p<0.05$) higher in the intervention group (99%) than in the control group (93%). The percentage of clients whose gestational age (LMP) was recorded was significantly higher in the intervention group (95%) than in the control group (85%). The differences between intervention and control groups in the proportion of ANC clients whose medical history and LMP was recorded were statistically significant ($p<0.05$). The mean number of weeks pregnant at the time of the visit was significantly ($p<0.05$) lower in the intervention group (16.23 weeks) than in the control group (20.96 weeks).

The survey shows that in the intervention group, a significantly ($p<0.05$) higher percentage of patients had their obstetric history taken during the first visit of their antenatal care booking (96% compared to 78% in the control group). A significantly ($p<0.05$) higher percentage of clients in the intervention group had their blood pressure measured (93% compared to 76% in the control group), and they were weighed (93% compared to 82% in the control group). Moreover, a significantly ($p<0.05$) higher percentage of clients in the intervention group were checked for pallor/anaemia (92% compared to 63% in the control group). Also, they had a higher percentage of patients who gave urine (75% vs. 46%) and blood (74% vs. 49%) samples for testing, and the blood sample was tested for syphilis (71% vs. 51%), and the differences were statistically significant ($p<0.05$). The intervention group had a higher percentage of patients reporting that their haemoglobin level was recorded in their card (68% vs. 41%), and their blood group and RH were mentioned in the card 81% vs, 51%), compared to the control group, and the differences were statistically significant ($p<0.05$).

A significantly higher percentage of clients ($p<0.05$) in the intervention group were checked for oedema (49% compared to 28% in the control group) and had their abdomen measured with a tape (62% compared to 40% in the control group). The number of tetanus injections received was significantly ($p<0.05$) higher in the control group (2.93) than in the intervention group (2.06).

The survey also shows that in the intervention group, a significantly ($p<0.05$) higher percentage of clients received advice on their diet during pregnancy (67% compared to 40% in the control group) and had things they should have in preparation for delivery discussed with them (39% compared to 22% in the control group).

The survey also assessed patients' knowledge of danger signs during pregnancy, and the results showed that bleeding was the most commonly known danger sign among both groups, while fever was significantly ($p<0.05$) more well-known in the intervention group (65%) compared to the control groups (45%).

Regarding counseling and family planning, the intervention group had a significantly ($p<0.05$) higher percentage of patients reporting that a provider talked with them about counseling, compared to the control group (37% vs. 18%).

In summary, the survey shows that the intervention group received better quality antenatal care compared to the control group in many areas, including obstetric history, blood pressure measurement, weight, pallor/anaemia check, oedema check, diet advice, and preparation for

delivery. The intervention group also had better knowledge of danger signs during pregnancy and received more counseling on family planning and counseling.

Table 19: ANC patients' assessment of the service they received from health facility by intervention status

Questions asked to Antenatal Care (ANC) service clients		Intervention	Control
		%	%
Is medical history is recorded pregnancy taken & recorded antenatal-care card/book? (that responded "Yes")		99*	93
How many weeks pregnant is the client, according to the ANC card? in weeks		16.23*	20.96
Is gestation age (LMP) recorded? (that responded "Yes")		95*	85
Was your obstetric history taken during the first visit of your ANC booking (that responded "Yes")		96*	78
During this visit, did someone measure your blood pressure? (that responded "Yes")		93*	76
During this visit, were you weighed? (that responded "Yes")		93*	82
During this visit did the nurses check you for pallor/anaemia (that responded "Yes")		92*	63
During this visit, did you give a urine sample? (that responded "Yes")		75*	46
During this visit, did you give a blood sample? (that responded "Yes")		74*	49
Was your blood sample tested for syphilis? (that responded "Yes")		71*	51
Is Haemoglobin recorded in client's card? (that responded "Yes")		68*	41
Does the card/book mention the client's blood group RH? (that responded "Yes")		81*	51
During this visit were you checked for anaemia? (that responded "Yes")		79*	47
During this visit, did a health worker give you advice on your diet this is, what to eat and drink during pregnancy? (that responded "Yes")		67*	40
3.32 during this visit were you checked for oedema (that responded "Yes")		49*	28
During this visit, did you schedule your delivery in the facility? (that responded "Yes")		57*	37
Average number of tetanus injections received including any TT injection you received today? (that responded "Yes")		2.06*	2.93
During this visit or previous visits, has a provider discussed things you should have in preparation for your delivery? (that responded "Yes")		39*	22
During this visit, was your abdomen measured with a tape? (that responded "Yes")		62*	40
During this or previous visits, has a provider discussed with you the side effects of the iron pill?	Yes, this visit	26*	12
	Yes, previous visit	23*	9
	Yes, both visits	11	4
Please tell me any side effects of the iron pill that you know of?	Nausea	70*	40
	Black stools	31*	9
	Constipation	35	25
	Don't know / None	21	49*
	Other	2	6

Questions asked to Antenatal Care (ANC) service clients		Intervention	Control
		%	%
During this or previous visits, has a provider given or prescribed any anti-malarial pills for you? show the client the antimalarial capsules	Yes, this visit	11*	1
	Yes, previous visit	11	6
	Yes, both visits	4	1
During this visit was fetal heart checked (that responded 'Yes')		73*	56
Urine tested for infection? (that responded 'Yes')		80*	65
Urine test for protein (that responded 'Yes')		71*	54
Was haemoglobin test conducted? (that responded 'Yes')		66*	51
Blood group and RH (that responded 'Yes')		80*	56
What danger signs do you know?	Bleeding	86	95
	Fever	68*	45
	Genital injuries	33	33
	Other	4	01
During this or previous visits, did a provider talk with you about counseling?	Yes, this visit	37*	18
	Yes, previous visit	27	29

**statistically significant at p<0.05*

Mothers/ Guardians of Under- 5-year-old children suffering from Pneumonia (IMNCI Pneumonia Clients)

Table 20 presents the analysis of data collected from mothers/guardians of under-5-year-old children suffering from pneumonia who were interviewed at the time of their health facility visit to get treatment for their children.

In terms of child health care card, 61% of the intervention group and 71% of the control group had a card. Among those who had a card, 59% of the intervention group and 42% of the control group reported that growth monitoring was charted, while 63% of the intervention group and 42% of the control group reported charting on developmental milestones.

Regarding immunization status, 85% of the intervention group and 42% of the control group reported being up to date, while 15% of the intervention group and 58% of the control group reported missing some immunizations. The difference was statistically significant ($p<0.05$).

In terms of recording important data, the intervention group performed better than the control group. For example, 98% of the intervention group and 83% of the control group reported that the date of the visit was recorded on the OPD/IMNCI clinic, and 100% of both groups reported that the age of the child was recorded.

The intervention group also outperformed the control group in terms of recording physical examination data. For instance, 80% of the intervention group and 33% of the control group reported that the child's weight was recorded and plotted in a growth chart during the visit, while 70% of the intervention group and 33% of the control group reported that the child's height was recorded and plotted in a growth chart.

The intervention group also reported a higher percentage of clinical history documentation, symptoms and duration recorded, child temperature recorded, pulse recorded, respiratory rate recorded, and signs of respiratory distress recorded than the control group.

Overall, the survey results suggest that the intervention group received more comprehensive and thorough care for their children suffering from pneumonia compared to the control group.

Table 20: Assessment of health service received by mothers/ guardians of Under- 5-year-old children suffering from Pneumonia by intervention status

Questions to mothers/ guardians of Under- 5-year-old children suffering from Pneumonia (IMNCI Pneumonia clients)		Intervention %	Control %
Do you have a child health care card? "If yes: ask to see the card/book?	Yes	61	71
	No, card kept with facility	35	24
	No, card/book used	4	6
Check child health card that indicate charting on growth monitoring (that responded 'Yes')		59	42
Check child health card that indicate charting on developmental milestones (that responded 'Yes')		63	42
Is health card on immunization status up to date?	Yes, up to date	85*	42
	No some immunizations missed	15	58*
Was date of visit recorded on opd/imnci clinic? (that responded 'Yes')		98*	83
Was age of child recorded (that responded 'Yes')		100	100
Was child weight recorded and plotted in a growth chart on this visit (that responded 'Yes')		80*	33
Was child height recorded and plotted in a growth chart (that responded 'Yes')		70*	33
Was child temperature recorded (that responded 'Yes')		96*	67
Was pulse recorded (that responded 'Yes')		80	58
Was respiratory rate recorded (that responded 'Yes')		89*	58
Was signs of respiratory distress recorded (that responded 'Yes')		85*	50

*p<0.05

<u>Mothers/Guardians of Under- 5-year-old children suffering from Diarrhoea (IMNCI Diarrhoea Clients)</u>

Table 21 presents data from a survey of mothers/guardians of under-5-year-old children suffering from diarrhea who visited a health facility to receive treatment for their children.
Overall, the intervention group appears to have higher percentages of positive responses for most questions compared to the control group.

In terms of child health care card, both intervention and control groups have high percentages of respondents with a card (69% and 77%, respectively). The majority of respondents in both groups have their card with them (rather than it being kept at the facility or not being used).

For questions related to charting on growth monitoring and developmental milestones, the intervention group has higher percentages of positive responses compared to the control group (84% vs 53% for growth monitoring and 81% vs 53% for developmental milestones).

Regarding immunization status, the intervention group has a significantly higher percentage of respondents ($p<0.05$) with up-to-date immunizations compared to the control group (89% vs

59%). Additionally, a smaller proportion of the intervention group reports missing some immunizations compared to the control group (11% vs 41%).

For questions related to recording information during the facility visit, the intervention group has higher percentages of positive responses compared to the control group for most questions. For example, all respondents in the intervention group report that the date and age of the child were recorded, compared to 88% and 88% in the control group, respectively. The intervention group also has higher percentages of positive responses for recording child weight and temperature, as well as signs and characteristics of diarrhea, although the difference was not statistically significant ($p>0.05$).

Overall, the data suggests that the intervention group has better adherence to recommended practices related to child health and diarrhea treatment compared to the control group.

Table 21: Assessment of health service received by mothers/guardians of Under- 5-year-old children suffering from Diarrhoea (IMNCI Diarrhoea Clients)

Mothers/Guardians of Under- 5-year-old children suffering from Diarrhoea (IMNCI Diarrhoea Clients)		Intervention %	Control %
Do you have a child health care card? " if yes: ask to see the card/book?	Yes	69	77
	No, forgot card at home	31	23
Check if child health card indicate charting on growth monitoring	Yes recorded	84*	53
	No record	16	47*
Check child health card that indicate charting on developmental milestones	Yes recorded	81*	53
	No record	19	47*
Check child health card on immunization status	Yes, up to date	89*	59
	No some immunizations missed	11	41*
Date of visit recorded on opd/imnci clinic (that responded "Yes")		100	88
Age of child recorded (that responded "Yes")		100	88
Child weight recorded and plotted in a growth chart on this visit (that responded "Yes")		81	65
Child height recorded and plotted in a growth chart (that responded "Yes")		62	53
Child temperature recorded (that responded "Yes")		78	76
Respiratory rate recorded (that responded "Yes")		70	65
Pulse recorded (that responded "Yes")		73	65
Symptoms and duration recorded at a minimum absence/presence (that responded "Yes")		68	65
Duration of fever (that responded "Yes")		57	71
Characteristics of stool (blood) (that responded "Yes")		62	59
Characteristics of stool mucus (that responded "Yes")		62	53

*p<0.05

Out-Patient Department (OPD) clients

The survey result shown in Table 22 presents data on the Out-Patient Department (OPD) clients, covering several aspects of the clients' experience, including their possession of an outpatient card, the recording of their personal and medical information, and the services provided to them during their visit.

In terms of possession of an outpatient card, the majority of the clients (69%) have one with them, while 27% have their cards kept with the facility, and 4% do not have any card or book used. The data also show that all clients who responded "yes" to having an outpatient card had their date of visit, age, and weight recorded.

For medical information, a higher percentage of intervention clients had their blood pressure (91% vs. 79%), pulse 65% vs. 56%), and any other abnormalities noted (30% vs. 19%), compared to the control group. However, the control group had a slightly higher percentage of patients with their weight (58% vs. 53%), temperature (77% vs. 65%), and respiratory rate (63% vs. 59%) recorded, although the difference was not statistically significant ($p>0.05$).

Overall, the survey results suggest that patients' experience did not vary statistically significantly ($p>0.05$) between the intervention group and control groups.

Table 22: Assessment of health service received by outpatient department (OPD) visitors/clients by intervention status

Out-Patient Department (OPD) clients		Intervention	Control
		%	%
Do you have an outpatient card with you?	Yes	67	73
	No, card kept with facility	29	25
	No, card/book used	5	2
Date of visit recorded on card (that responded "Yes")		100	100
Age of patient recorded (that responded "Yes")		100	100
Weight of patient recorded (that responded "Yes")		53	58
Blood pressure recorded (that responded "Yes")		91	79
Patient temperature recorded (that responded "Yes")		65	77
Respiratory rate recorded (that responded "Yes")		59	63
Pulse recorded (that responded "Yes")		65	56
Any other abnomalities noted (that responded "Yes")		30	19
Medicines resupplied (that responded "Yes")		85	88
Review date given (that responded "Yes")		71	60

Postnatal Care Clients (PNC)

Table 23 provides results from a survey conducted with postnatal care (PNC) patients. The survey asked a range of questions related to the quality of care provided during the PNC visit.

The first question asks whether the patient has an integrated antenatal, labor, delivery, newborn and postnatal care card. The majority of patients in both groups responded "Yes," with 63% in the intervention group and 40% in the control group. The next most common response was "No, card kept with facility," with 35% in the intervention group and 60% in the control group.

The survey also asked whether certain procedures were conducted during the PNC visit. The responses vary by procedure and group, but in general, the intervention group had a higher percentage of positive responses than the control group. For example, 88% of the intervention group reported having their temperature, pulse, and respiration checked, compared to 67% in the control group. In both intervention and control groups, the majority of clients had their blood pressure measured during their visit (78% and 93%, respectively). Around two-thirds of clients

in both intervention and control groups had their uterine contraction assessed or checked for postpartum hemorrhage (71% and 73%, respectively). The majority of clients in both intervention and control groups were asked about dribbling or leaking urine during their visit (71% and 53%, respectively).

The majority of clients in the intervention group were checked for pallor or anemia (76%), compared to only 60% in the control group. Similarly, the majority of clients in both intervention and control groups were examined for vaginal discharge during their visit (69% and 60%, respectively). A higher proportion of patients in the control group (93%) had their breasts examined during their visit, compared to the intervention group (73%).
The survey also asked about counseling and education provided during the visit. Again, the responses vary by topic and group. For example, 81% of both groups reported being counseled on child immunizations, while only 50% of the intervention group and 60% of the control group reported being counseled on the use of insecticide-treated nets (ITN).

The survey also asked about the patient's knowledge of danger signs related to their health and their baby's health. The responses again vary by topic and group, but in general, the intervention group had a higher percentage of correct responses than the control group.

Overall, the data suggests that the intervention group received higher quality care and education during their PNC visit than the control group.

Table 23: Assessment of health service received by PNC clients by intervention status

Questions to Postnatal Care Clients (PNC)		Intervention	Control
		%	%
Do you have an integrated antenatal, labor, delivery, newborn and postnatal care card" if yes: ask to see the card/book.	Yes	63	40
	No, card kept with facility	35	60
	No, card/book used	2	0
Date of visit recorded (that responded "Yes")		97	100
Was your temperature, pulse, respiration checked today if yes: ask to see the card/book. (that responded "Yes")		88	67
During this visit, did someone measure your blood pressure? (that responded "Yes")		78	93
During this visit did someone assess uterine contraction /look for PPH IF YES: ASK TO SEE THE CARD/BOOK. (that responded "Yes")		70	73
During this visit did someone ask you about Dribbling/leaking urine IF YES: ASK TO SEE THE CARD/BOOK (that responded "Yes")		70	53
During this visit did the nurses check you for pallor/anaemia (that responded "Yes")		76	60
During this visit did someone exam you for vaginal discharge IF NO: ASK TO SEE THE CARD/BOOK (that responded "Yes")		69	60
During this visit did someone conduct a pelvic examination (that responded "Yes")		65	73
During this visit did someone exam your breasts (that responded "Yes")		72	93
During this visit, did the provider give you Vitamin A pills, or give you a prescription for them? (that responded "Yes")		72	87
Ask to see the clients vitamin a pills	Saw pills	82	92
	Saw prescription	15	8
	No pills or prescription	3	0
During this visit did the nurse counsel you on danger signs (that responded "Yes")		59	73
Please tell me any danger signs that you know of?	Increased vaginal bleeding	75	91
	Fits	9	0

Questions to Postnatal Care Clients (PNC)		Intervention	Control
		%	%
	Fast or difficulty in breathing	25	18
	Fever and too weak to get out of bed	50	45
	Severe headaches with blurred vision	53	45
	Calf pain	56	45
	Red or swelling calf	44*	9
	Shortness of breath or chest pain	25	27
During this visit, did someone counsel you on child immunizations (that responded "Yes")		81	80
During this visit, were you counselled on use of ITN? (that responded "Yes")		50	60
During this visit was the baby's breathing checked? (CHECK clients CARD TO ASSESS IF THIS IS RECORDED) (that responded "Yes")		80	80
During this visit was the baby's breast-feeding ability assessed (that responded "Yes")		78	87
During this visit did someone assess the baby's weight? (that responded "Yes")		89	73
During this visit was the babys immunization status checked? ASK TO SEE THE CHILD IMMUNIZATION CARD (that responded "Yes")		81	67
Was the baby tested for HIV? (that responded "Yes")		48	53
Was the HIV test result R/NR (non-reactive) (that responded "Yes")		85	88
Was ARV prophylaxis dispensed to the mother? (CHECK clients CARD TO ASSESS IF THIS IS RECORDED) (that responded "Yes")		77	88
Was ARV prophylaxis dispensed to the newborn? (CHECK clients CARD TO ASSESS IF THIS IS RECORDED) (that responded "Yes")		73	88
During this visit, did someone from this health facility discuss baby feeding options with you (that responded "Yes")		44	60

*p<0.05

Impact of program on patients' overall satisfaction

Table 24 compares the responses of the intervention group and the control group at endline regarding their satisfaction with various aspects of their visit to a health facility. The percentages represent the proportion of respondents who strongly disagree, slightly disagree, slightly agree, and strongly agree with each statement.

Overall, the intervention group appears to be more satisfied than the control group with their visit to the health facility. For example, a higher percentage of the intervention group strongly agrees that it is convenient to travel from their house to the health facility (63% compared to 52% in the control group), the health staff are courteous and respectful (72% compared to 67% in the control group), and the overall quality of services provided was satisfactory (64% compared to 49% in the control group).

There are some exceptions to this trend, however. For example, a higher percentage of the control group strongly agrees that they had enough privacy during their visit (82% compared to 75% in the intervention group), and the amount of time you spent waiting to be seen by a health provider was reasonable (63% compared to 61% in the intervention group).

In terms of specific aspects of the health facility, the intervention group is more likely to perceive the health facility as clean, trust in the skills and abilities of the health workers, and believe that the health workers did a good job of explaining their illness and treatment. The intervention group is also more likely to find the fees for registration, lab tests, medication, and transport to be reasonable.

Overall, the results suggest that the intervention group had a more positive experience at the health facility than the control group, particularly in terms of convenience, staff courtesy, quality of services, and reasonable fees. However, the control group had a slightly better perception of privacy and time spent with the health worker.

Table 24 also compares the responses of the intervention group and the control group regarding their perceptions of the security and trust patients have in health care providers.

Overall, the intervention group appears to have a more positive perception of the security and trust they have in health care providers compared to the control group. For example, a higher percentage of the intervention group strongly agrees that the health workers are extremely thorough and careful (70% compared to 58% in the control group), they completely trust the health workers' decisions about medical treatments (66% compared to 56% in the control group), and they trust the health worker completely (70% compared to 61% in the control group).

The intervention group also perceives the health workers as more approachable and easy to make contact with than the control group. A higher percentage of the intervention group strongly agrees that the health workers are very friendly and approachable (72% compared to 61% in the control group) and easy to make contact with (61% compared to 57% in the control group). Additionally, a higher percentage of the intervention group strongly agrees that health workers care about their health just as much or more than they do (53% compared to 46% in the control group).

However, both groups have concerns about the fairness of treatment provided by health workers. A higher percentage of the control group believes that health workers act differently toward rich people than poor people (66% compared to 49% in the intervention group).

Overall, the results suggest that the intervention group has a more positive perception of the security and trust they have in health care providers, particularly in terms of their thoroughness, decision-making, and trustworthiness. Both groups have concerns about the fairness of treatment provided by health workers, with the control group being more likely to perceive differences in treatment based on socioeconomic status.

Table 24: Patients' regarding their satisfaction with various aspects of their visit to a health facility, and their perceptions of the security and trust patients have in health care providers by zone and intervention status

Client/Patient Satisfaction		Strong disagreement		Slight disagreement		Slight agreement		Strong agreement		Not applicable	
		Inter-vention	Control	Inter-vention	Control	Inter-vention	Control	Inter-vention	Control	Inter-vention	Control
Patient satisfaction	It is convenient to travel from your house to the health facility	10	8	10	15	18	25	63	52	0	0
	The health facility is clean	2	5	5	16	22	35	71	44	0	0
	The health staff are courteous and respectful	0	2	10	5	17	26	72	67	0	0
	You trust in the skills and abilities of the health workers	1	2	8	6	19	33	73	59	0	0
	The health workers did a good job of explaining your illness	2	3	10	7	19	29	70	61	0	0
	The health workers did a good job of explaining your treatment	2	3	9	11	20	34	68	52	0	0
	It is easy to get medicine that health workers prescribe	6	3	19	14	26	32	48	49	1	1
	The registration fees of this visit to the health facility were reasonable	3	1	5	3	15	8	70	48	6	41
	The lab fees of this visit to the health facility were reasonable	3	1	7	2	16	9	62	44	11	45

	The medication fees of this visit to the health facility were reasonable	3	0	8	3	14	9	67	44	8	44
	The transport fees for this visit to the health facility were reasonable	12	6	13	8	15	8	48	35	12	44
	The amount of time you spent waiting to be seen by a health provider was reasonable	5	3	13	5	21	27	61	63	0	2
	You had enough privacy during your visit	1	0	8	2	16	15	75	82	0	0
	The health worker spent a sufficient amount of time with the patient	1	1	11	11	25	31	62	57	0	0
	The hours the facility is open is adequate to meet the needs of the community	2	2	12	8	18	30	68	60	0	0
	The overall quality of services provided was satisfactory	2	4	11	10	23	37	64	49	0	0
	Your overall visit was satisfactory	1	1	9	8	21	34	68	58	0	0
Security and trust of health care providers	The health workers are extremely thorough and careful	1	2	8	8	21	32	70	58	0	0
	You completely trust the health workers decisions about medical treatments	1	0	10	9	23	35	66	56	0	0
	The health workers are very friendly and approachable	1	1	9	11	18	28	72	61	0	0
	The health workers are easy to make contact with	1	1	13	12	24	30	61	57	0	0
	The health workers care about your health just as much or more than you do	2	9	18	22	27	22	53	46	0	0
	The health workers act differently toward rich people than toward poor people	49	66	15	14	12	6	23	14	1	0
	All in all, you trust the health worker completely	2	1	12	8	16	31	70	61	0	0

Additional tests and analysis was conducted to determine the impact of the program on patients' satisfaction. The overall satisfaction score was obtained by summing the items results of all the participants and compared to the general average obtained. A participant was classified as "Satisfied" if his/her rate was above the general average; otherwise he/she was "Unsatisfied". All items included to measure client satisfaction were also aggregated.

Table 25 presents the results of the regression analysis done to assess the impact of the program on the satisfaction of the clients for selected service categories.

Table 25: Clients' satisfaction by service type

	Coefficient for the interaction variable (β_3)					
Service type	B	S.E.	Wald	Df	Sig.	Exp(B)
ANC	0.844	0.136	38.530	1	0.000	2.325
PNC	1.014	0.282	12.918	1	0.000	2.758
OPD	1.163	0.159	53.206	1	0.000	3.200
IMNCI Diarrhea	0.537	0.226	5.621	1	0.018	1.710
IMNCI Pneumonia	1.012	0.252	16.064	1	0.000	2.751

Based on the logistic regression results presented in Table 25, the intervention has a statistically significant effect on client/patient satisfaction for all service types, including ANC, PNC, OPD, IMNCI for pneumonia and IMNCI Diarrhea. The coefficient for interaction variable β_3 is positive for each service type, indicating that there is an increase in client satisfaction from baseline to endline in the intervention facilities at $p<0.01$ compared to the control facilities. For all health service types, the intervention has a positive impact on client satisfaction across different types of health services. The details of client's satisfaction by the service type is described as follows:

- **ANC**: The odds of client satisfaction with ANC services were 2.325 times higher in the intervention group compared to the control group, after adjusting for other variables. The p-value for this coefficient is 0.000, indicating that the effect of the intervention on client satisfaction with ANC services is statistically significant at a 1% level of significance.

- **PNC**: The odds of client satisfaction with PNC services were 2.758 times higher in the intervention group compared to the control group, after adjusting for other variables. The p-value for this coefficient is 0.000, indicating that the effect of the intervention on client satisfaction with PNC services is statistically significant at a 1% level of significance.

- **OPD**: The odds of client satisfaction with OPD services were 3.200 times higher in the intervention group compared to the control group, after adjusting for other variables. The p-value for this coefficient is 0.000, indicating that the effect of the intervention on client satisfaction with OPD services is statistically significant at a 1% level of significance.

- **IMNCI Diarrhea**: The odds of client satisfaction with IMNCI diarrhea services were 1.710 times higher in the intervention group compared to the control group, after adjusting for other variables. The p-value for this coefficient is 0.018, indicating that the effect of the intervention on client satisfaction with IMNCI diarrhea services is statistically significant at a 5% level of significance.

- **IMNCI Pneumonia**: The odds of client satisfaction with IMNCI pneumonia services were 2.751 times higher in the intervention group compared to the control group, after adjusting for other variables. The p-value for this coefficient is 0.000, indicating that the effect of the intervention on client satisfaction with IMNCI pneumonia services is statistically significant at a 5% level of significance.

In summary, the quantitative results show that the intervention had a statistically significant positive impact on client satisfaction for all service types, with odds ratios ranging from 1.710 to 3.200. These findings suggest that the intervention was effective in improving client satisfaction with health services in the intervention group compared to the control group.

Qualitative data gathered from patients and other key informants also suggest that the same. Patients that participated in the FGDs expressed satisfaction with the quality of services they received at the health centre/hospital, citing privacy, confidentiality, being listened to, getting the support/service they/their child needed, and getting necessary medicine as positive experiences. *"I'm very satisfied with the attractiveness of the compound, the quality of health care provided by the service provider, and counseling activities,"* a patient said.

Key informants and patients consulted for the end of program evaluation highlighted several changes that the program has brought about in improving quality health services in the target communities. One of the key benefits of the program has been the improvement of health facilities' infrastructure. These changes include infrastructure improvements such as the construction of fence doors, ceramic floors, toilets, and shelves, as well as the purchase of necessary equipment and medical supplies. A key informant in this regard stated that *"It helped us fulfill the more than 80% of our facility gap. This enabled us to improve the quality of health care services."* They provided specific examples, such as constructing ceramic floors, buying chairs, constructing quality toilets, and constructing water line pipes. Another informant emphasized that the program helped them *"fulfill the necessary equipment, strengthen capacity of health workers, we fulfill the medical equipment from this funding, we constructed infrastructure for the health center that can help us to provide quality of care."* Improved infrastructure made possible by the project was also mentioned as a key factor behind increased quality of service.

Almost all FGD participants have also expressed satisfaction with the health staff's friendliness, understanding, respectfulness, trustworthiness, sympathy, and caring. They have noted that the staff are respectful and provide excellent service, with no discrimination based on identity. *"I am very satisfied with the health workers, because of their respectfulness friendliness and understanding*

…. The only problems is the long the waiting time," a patient that participated in one of the FGDs said.

All participants have also expressed that they would seek health care services at this hospital/health centre in the future and would refer a family/friend to this facility for health service. They have cited reasons such as the quality of service, respectfulness and counseling approach of the service provider, the availability of services, and the improvement in the registration area. Of course, many patients will continue using the services of the health facilities due to lack of alternatives and/or lack of money to pay for health services provided at private facilities. Since the health facilities are the only ones in many cases that are found nearby and some of the services provided with no payments, many patients would seek health care services in the future and refer a family/ friend to these facilities. *"I think the only facility that is nearby to our residences is this one and the quality of service is also good, so why would I not come to this facility in the future,"* a patient argued.

Although patients expressed satisfaction with the quality of services they received at the health centre/hospital, many also noted some areas of improvement. The areas of improvement relate to availability of medicine and water, quality delivery room and toilet, and adequacy of health workers to meet growing demand. Some patients noted that they will continue using the services of the facility if the shortage of medicine is addressed. *"If the unavailability of medicine is resolved, I will always use this health center because the health care providers are very nice,"* a patient said. Some patients also suggested the need to address shortage of health workers. In relation to this, a patient said the following:

> *"In fact, there is a change, but not adequate. The number of clients who need service has been increasing from time to time whereas the number of staff didn't fit the demand. So, this needs attention and resolution mechanisms."*

Overall, the PBF program has significantly improved the quality of health care service provision. The PBF program has led to the enhancement of the appearance of health facilities, including the construction of infrastructure and the improvement of the cleanliness and attractiveness of health centers, and improvements in availability of medicine.

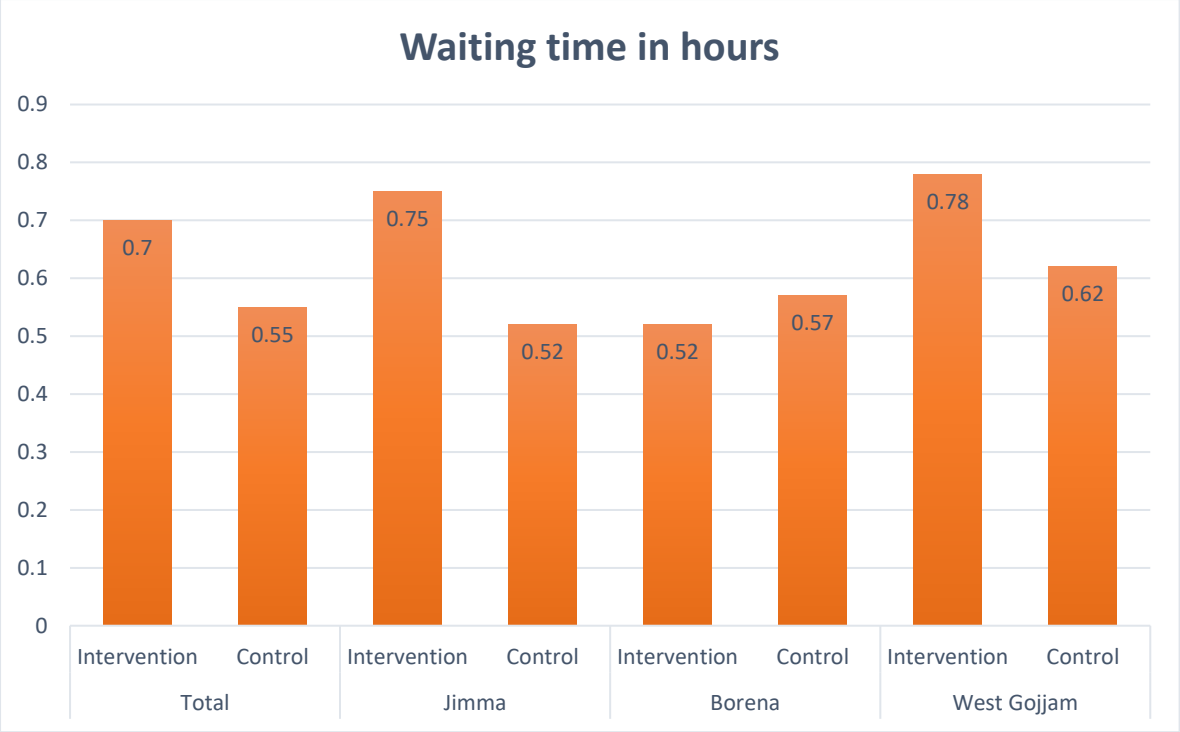


Figure 3: Patients' waiting time in hours by intervention status and zone

Although patients are satisfied with various domains of the quality of care, the increased quality of care has also increased patient flow or demand for health services in the intervention health facilities. The increased in demand for health services, according to some patients, has resulted in long-waiting time as some intervention health facilities do not seem to be adequately staffed to meet the growing demand for their services. For example, the mean waiting time for patients in the intervention group (0.70 hours) was higher than that of the control group (0.55 hours). This difference was statistically significant ($p<0.05$). See Figure 3.

A higher percentage of patients in the intervention facilities felt that the waiting time was too long (26%) compared to the control group (15%). This difference was also statistically significant ($p<0.05$). See Figure 4.

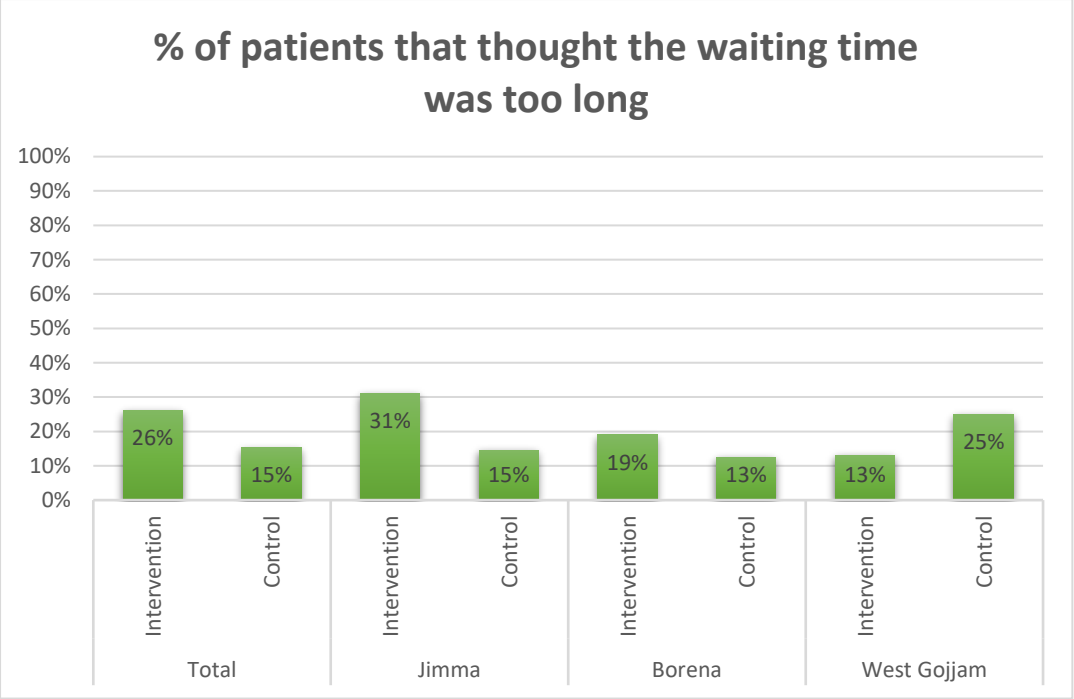


Figure 4: Patients that thought waiting time was too long by intervention zone and status

Although they are in the minority, a significant proportion of patients had long waiting times. *"The time that we spent at the card room is disappointing. Because it took a long time, their card room workers have no speed as well as not available in their room on time"*, a patient said. Another patient said *"There is long waiting time. It takes more than 2 hours to get services. There is large queue in card room."* Many patients and key informants believe that intervention health facilities are not adequately staffed to meet the growing demand for health services. Additionally, the high patient flow is forcing health workers to rush service delivery in some cases to the extent of compromising quality of care.

It is also important to note here that views of patients on waiting time vary, ranging from those that reported short waiting time to those that waited for over half a day. One of the patients that participated in the FGDs in this regard said *"Currently the time that we spent here is very short. Even we didn't stay here more than 30minute. The patient waiting time is significantly reduced within these two years. This is very impressive,"* A similar view was also shared by another patients who said *"In this health facility there is no time that we waste by looking for clinician. All health workers come just at their right time and provide services. Particularly in the last two years there is no overcrowded condition that we saw and consume our time."*

V. Program's impact on motivation and satisfaction of health workers

Among the strategies that the project employed to improve utilization of quality care services included motivating the health workforce at the intervention health facilities so as to improve their performance. The project tried to enhance motivation of health workforce by letting facilities retain part of their quarterly subsidies as staff performance incentives. Additionally, the project targeted to enhance motivation of health workforce by improving the regulatory capacity of WorHOs and ZHDs to strengthen their facility supervision, coaching and quality assessment.

Overall health worker staffing levels at health facilities

In light of the aforementioned targets, the end term evaluation assessed the level of staffing, and motivation and satisfaction of health workforce at both intervention and control health facilities.

As depicted in Figure 5 below, average health worker staffing rates computed for each type of health facility vis-a-vis the recommended staffing levels for HCs and hospitals.

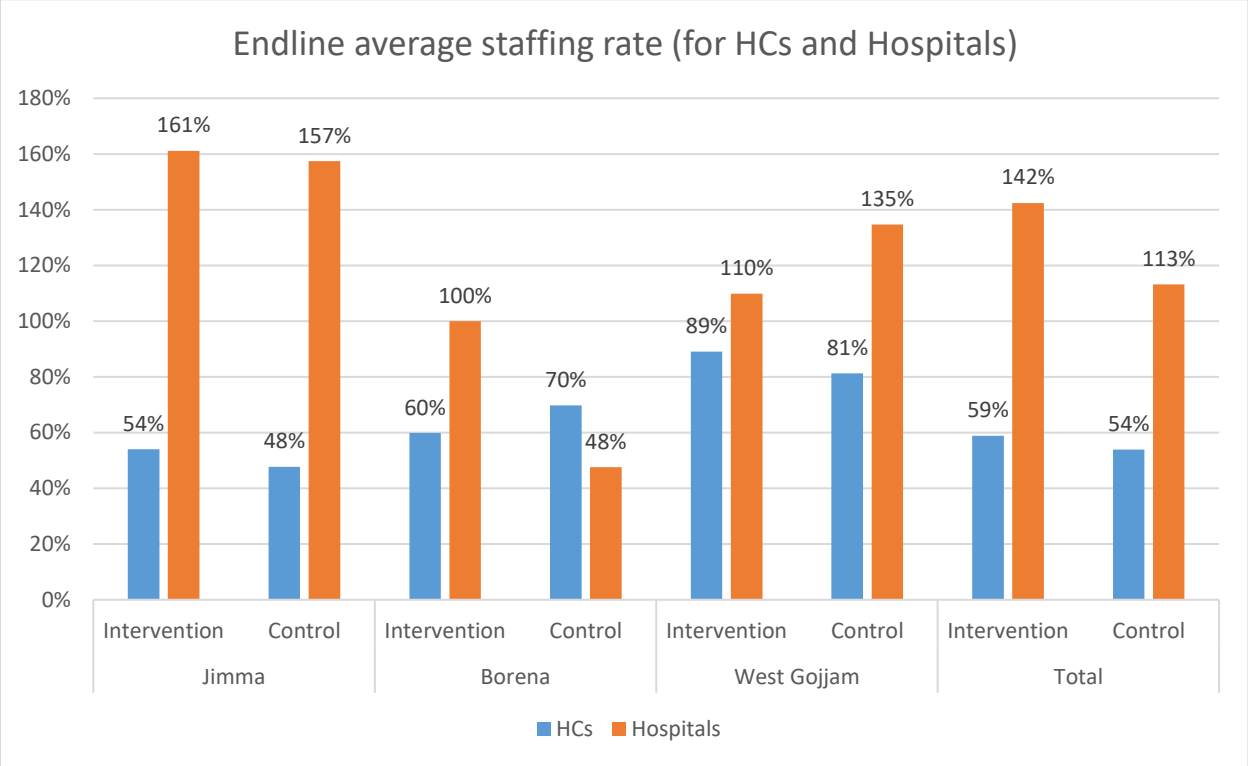


Figure 5: Average staffing rate at the health facilities

The results depicted in Figure 6 show existence of significant variation ($p<0.05$) in staffing rates between health centers and hospitals. Staffing level in intervention and control hospitals is higher than the staffing levels recommended by MoH. In contrast, health centers in intervention and control had 59% and 54% of the recommended number of staff, respectively. A similar pattern was also reported in the mid-term evaluation. Staffing rate did not vary statistically significantly ($p<0.05$) between intervention and control health facilities and across zones. The results suggest that health centers remain understaffed which is having significant implications. The shortage of health workers seems to have resulted in longer wait times for patients seeking medical attention. This can have a negative impact on the health outcomes of patients, particularly those with chronic conditions or in need of emergency care. A shortage of health workers can lead to overworked and burnt out health workers who may be more likely to make mistakes or provide suboptimal care. This, according to some patients, has resulted in decreased quality of care and lower patient satisfaction. A shortage of health workers can result in inefficiencies in healthcare service delivery, including delayed diagnoses and treatments, and increased hospital stays. This can impact the overall efficiency of healthcare services and have a negative impact on health outcomes. The results also suggest that the potential of the PBF program to increase demand and volume of service utilization could be curtailed if the facilities are not adequately staffed.

Staffing level per position

Table 26 shows the HC staffing rates by position. Looking at the overall staffing rate, the intervention group had a higher mean staffing rate compared to the control group (0.59 vs. 0.54). In terms of specific staff positions, the intervention group had a higher staffing rate for health officer position compared to the control group (1.15 vs. 0.91). The intervention group had a higher staffing rate for midwives position compared to the control group (0.67 vs. 0.65). The staffing levels were relatively high for nursing positions for both intervention and control groups, with no

significant differences (p>0.05) between the two groups. Similarly, the control group had a higher staffing rate for laboratory technicians/technologist positions compared to the intervention group (0.91 vs. 0.74). The staffing levels for pharmacists/pharmacy technician positions were also relatively high for both intervention and control groups, with no significant differences (p>0.05) between the two groups. The staffing levels were low for environmental health professionals both intervention and control groups, with the intervention group having slightly higher staffing levels overall. None of the health facilities had a general practitioner, maintenance officers and morgue attendants, and ophthalmic and psychiatry.

Overall, the intervention group had higher staffing levels compared to the control group in most staff positions, with statistically significant differences (p<0.05) in some cases. These findings are more or less consistent with the mid-term review. When comparing zones, the West Gojjam zone generally had higher staffing levels than the other two zones, particularly for Health Officers, while the Jimma and Borena zones had similar staffing levels for most positions.

However, there were still some key staff positions such as midwives where staffing levels are low for both intervention and control groups, indicating a need for improvement. A key informant, for example, highlighted, *"We have been working with small number of health workers."* Another informant also mentioned that *"There is a shortage of human powers like; delivery, health officer experts."* The shortage of midwives can lead to a situation where there are not enough skilled healthcare providers to adequately provide antenatal care (ANC), intrapartum care, and postnatal care (PNC) to pregnant women and newborns. Additionally, the shortage of midwives may have resulted in inadequate time and attention being given to each patient, leading to missed opportunities to provide appropriate education, counseling, and referrals. For example, chart audit results on normal deliveries showed adherence index that is below the required standard, which suggests gaps adhering to clinical guidelines, which might be partly due to shortage of midwives.

Table 26: Staffing Rate at Intervention and Control Health Centers by Position

Required staff positions	Zone							
	Jimma		Borena		West Gojjam		Total	
	Intervention Status		Intervention Status		Intervention Status		Intervention Status	
	Intervention	Control	Intervention	Control	Intervention	Control	Intervention	Control
Health Officer	1	0.88	1	1.17	2.75	0.5	1.15	0.91
General Practitioner	0	0	0	0	0	0	0	0
Midwife	0.57	0.47	0.71	1.22*	1.17	1	0.67	0.65
Nurse	1.14	1.22	1.09	1.4	1.7	2.2	1.17	1.31
Ophthalmic nurse	0	0	0	0	0	0	0	0
Psychiatry nurse	0.14	0	0	0	0	0	0.09	0
Environmental Health professional	0.07	0	0	0	1	0	0.13	0
Laboratory technician or technologist	0.64	0.75	0.79	1.33	1.25	1.5	0.74	0.91
Pharmacist or	0.6	0.61	0.81	0.56	1	1	0.7	0.63

Pharmacy Technician								
Cleaners	0.49	0.38	0.54	0.53	0.2	0.4	0.48	0.41
Archive Workers	0.33	.12*	0.48	0.39	0.58	0.5	0.4	0.2
Maintenance officer	0.14	0	0	0	0	0	0.09	0
Morgue attendant	0	0	0	0	0	0	0	0
Average staffing rate (HC)	0.54	0.48	0.6	0.7	0.89	0.81	0.59	0.54

* $p < 0.05$

Working hours and duties

Health workers' working hours differ based on the intervention status and the zone they are located in. The health workers in the intervention group generally work fewer hours compared to those in the control group. For instance, on weekdays, the mean working hours per day for health workers in the intervention group are 6.00 hours, while those in the control group work an average of 6.54 hours per day. Similarly, on weekends and public holidays, the mean working hours per day for health workers in the control group is much higher than those in the intervention group.

Moreover, there are significant variations ($p<0.05$) in the working hours across different zones. For example, health workers in Borena intervention facilities work fewer hours on weekdays (mean = 0.69 hours) compared to those in Jimma (mean = 8.02 hours) and West Gojjam (mean = 6.26 hours). Similarly, health workers in Jimma work longer hours on Saturdays (mean = 10.5 hours) and Sundays (mean = 10.4 hours) compared to other zones. Health workers in intervention health facilities work fewer hours than those in control health facilities in Jimma and West Gojjam zones.

Close to two-third (64%) of health workers in intervention group were never late in the last month, compared to 47% in the control group, and the difference was statistically significant ($p<0.05$). In the last month, the most common reasons for being late were “Authorized late arrival – training” and "Travelling to work takes too long".

Regarding the number of days absent from work, the mean number of days was highest in West Gojjam for the intervention group (2.78) and lowest in Borena for the control group (0.08). The mean number of days absent from work was significantly different ($p<0.05$) between some intervention and control groups, such as in Jimma (0.41 vs. 1.35). But the mean number of hours worked by health workers in Borena does not seem to make sense.

In terms of the number of days worked at the facility in the last week, the mean number of days was highest in the intervention group in Borena (6.29) and lowest in the control group in West Gojjam (5.0). The mean number of days worked at the facility was not significantly different ($p>0.05$) between intervention and control groups in all zones.

The mean number of hours worked during the last week was not significantly different ($p>0.05$) between intervention and control groups in all zones.

Regarding whether the respondents worked yesterday, the percentage of respondents who worked yesterday was highest in the intervention group in Borena (100%) and lowest in the in Jimma and West Gojjam (78%).

Regarding the number of hours worked yesterday or the last working day, the mean number of hours was higher in the control group (8.8 hours) than the intervention group (7.29) and the difference was statistically significant ($p<0.05$). Mean working hour was highest in the control group in Jimma (9.49) and lowest in the intervention group in Jimma (6.50).

Salary and Bonus

It appears that the mean net salary and top-up[19] amount vary across intervention and control groups. The mean net salary for the intervention group is slightly lower than that of the control group, while the mean top-up amount for the intervention group is higher than that of the control group. But the differences were not statistically significant ($p>0.05$).

Additionally, there is some variation in mean net salary and top-up amount across the different zones. For instance, the mean net salary is highest in the Borena zone for both the intervention and control groups, while the mean top-up amount is highest in the Jimma zone for both the intervention and control groups. Salary and top up is lowest in Jimma and West Gojjam zones.

The percentages of respondents who received their salary in the past <7 days are as follows: 20% (intervention) and 23% (control) in Jimma; 0% (intervention) and 25% (control) in Borena; and 0% (both intervention and control) in West Gojjam. For the time period of 1-2 weeks ago, the percentage of respondents who received their salary is 17% (intervention) in Jimma, and 0% in all other groups. For the time period of 3-4 weeks ago, the percentage of respondents who received their salary is between 57% and 100% in different groups.

Health workers were also asked about the factors that determine the size of the salary change. In the intervention group, the most common factor for salary change is routine increment (30%), compared to 53% in the control group. In Jimma, the most common factor for salary change in the intervention group is routine increment (20%), compared to 68% in the control group, and the difference was statistically significant ($p<0.05$). In West Gojjam, the most common factor for salary change in both the intervention and control groups is routine increment (67% for each group). In Borena, the most common factor for salary change in both the intervention and control groups is individual performance (43% and 33%, respectively). It is also important to note that higher proportion of health workers in intervention group (42%) reported no change in in their salary than the control group (29%) and the difference was statistically significant ($p<0.05$).

Health workers were also asked whether an increase in salary or top-ups based on performance improves the quality of work. A higher proportion of health workers in the control group (75%) felt that increase in salary based on performance improves their quality of work, compared to only 43% in the intervention group. But equal proportion (50%) of health workers in both groups reported that increase in top-ups based on performance improves the quality of work.

The percentage of respondents who received all their due top-ups for the last 12 months is higher in the control group (82%) than in the intervention group (60%). In Jimma, the percentage of respondents who received all their due top-ups is higher in the control group (87%) than in the intervention group (67%). In Borena, the percentage of respondents who received all their due top-ups is higher in the control group (75%) than in the intervention group (43%). In West Gojjam, the percentage of respondents who received all their due top-ups is higher in the control group (67%) than in the intervention group (56%).

[19] Salary top-ups mean official cash payments or in-kind benefits that a civil servant receives over and above what colleagues in the same grade and pay scale receive.

Overall, the data suggests that there are some differences between the intervention and control groups and different zones in terms of salary and bonuses. But it's worth noting that the sample sizes are different across the different zones and intervention groups, which could impact the accuracy of the mean estimates.

Other Compensation, benefits and supplemental Income

The results show that the percentage of employees who receive free housing provided by the facility or the community is higher (33%) among the intervention group than the control group (27%), although the difference was not statistically significant ($p>0.05$). Over a third of health workers (37%) in intervention facilities in Jimma and Borena zones reported receiving free housing provided by the facility or community, while none of the health workers in Borena control group and both groups in West Gojjam received these benefits.

The percentage of employees who receive health care benefits and medicines is generally low in both groups, with the exception of West Gojjam, where all intervention and control employees receive these benefits.

The majority of employees in both groups receive uniform or uniform allowance for their work. However, in Jimma, the control group has a much higher percentage (87%) compared to the Intervention group (41%). The percentage of employees who receive risk allowance is generally low in both groups.

Most health workers in Borena intervention (95%) and control (67%) areas receive hard to reach allowance, compared to none in Jimma and West Gojjam zones. Regarding the payment period for the hard to reach allowance, most employees who receive it are paid monthly.

Regarding bonuses, most health workers (80%) in the intervention areas reported receiving bonuses that are based on the volume of services their facility provides or the quality of services (87% in Jimma and 67% of health workers in Borena and West Gojjam intervention facilities), almost all received the payment quarterly.

Regarding the question "How much did you receive in the last period for bonuses?" we can see that the intervention group in Jimma had the highest average bonus with 3491.305 ETB (Ethiopian Birr), followed by the intervention group in Borena with an average of 2798.715 ETB. Meanwhile, the intervention group in West Gojjam had the lowest compensation with an average of 983.34 ETB.

Regarding the question "Do you currently receive any other allowances?", the percentage of health workers receiving allowances in both intervention and control groups is relatively low, with only 2% in the intervention group and 4% in the control group. Finally, almost none of the health workers have another job to supplement your income from this health facility.

Supervision

Based on the data presented, it appears that the majority of respondents across all zones had a supervisor responsible for providing feedback on their performance (ranging from 65% to 95% across zones). However, there was a notable difference in the proportion of respondents who had a supervisor between the intervention and control groups, with a higher proportion of respondents in the intervention group reporting having a supervisor than in the control group (89% vs. 76%, respectively). The chi-square test shows a statistically significant difference ($p<0.05$), indicating that the intervention might have had an effect on the likelihood of having a supervisor.

In the intervention group, the most common supervisor was WorHO (31%), followed by head of a health facility (24%), and chief medical officer (9%). In the control group, the most common supervisor was WHO (38%), followed by head of a health center (35%), and chief medical officer (5%).

The data also suggests that the frequency of supervisory meetings varied across the different zones, with the average number of meetings per year ranging from 7.3 to 26.0 across zones. The highest frequency of meetings was reported in the West Gojjam intervention zone (26.0 meetings per year), while the lowest frequency was reported in the Borena control group (7.3 meetings per year). On average, in the intervention group, supervisors met with their staff 13.13 times per year, while in the control group, supervisors met with their staff 10.84 times per year. But the difference was not statistically significant ($p>0.05$). In the intervention group, 71% had met with their supervisor within the past 30 days, while in the control group, 73% had met with their supervisor within the past 30 days.

In the last supervisory interaction, the supervisor mostly checked records, and observed consultations in almost all groups. The most common action was to check records, followed by observing consultations and providing health instruction.

Regarding job difficulties, most of the participants reported discussing them with their supervisor in both intervention (94%) and control (92%) areas. More participants in intervention group (77%) reported improvements after discussing job difficulties with their supervisors compared to the control group (73%).

In terms of job satisfaction, the majority of respondents across all zones reported being satisfied or very satisfied with their last supervisory meeting (ranging from 57% to 85% across zones). However, there was some variation in the proportion of respondents who reported being motivated by these meetings across zones, with the highest proportion reported in the West Gojjam intervention zone (89%) and the lowest in the Borena control group (36%). Most participants in both intervention and control health facilities reported being satisfied or very satisfied. Finally, the majority of participants reported that the supervisory meetings motivate them, with percentages ranging from 76% to 100% across different groups and zones. Overall, the results suggest that the intervention group had higher rates of having a supervisor and more frequent meetings with supervisors compared to the control group. Additionally, the WHO was the most common supervisor across the zones. In conclusion, the intervention had a statistically significant effect ($p<0.05$) on the likelihood of having a supervisor and the frequency of meetings, but no significant effect on the types of supervisors or the timing of the last supervisory interaction.

Health Workers' Satisfaction

The end-term evaluation assessed the job satisfaction of health workers in both intervention and control health facilities. For each aspect of their work, the survey participants were asked to rate their level of agreement or disagreement with a series of statements using a Likert scale that ranged from "Strongly Agree" to "Strongly Disagree." Workers' responses to statements were measured from 1 to 5 where 1 was "strongly agree" and 5 was "strongly disagree". Prior to the analysis, we first reversed statements written in negative form. We also reversed the responses to each of the statements to make sure that higher values such as "5" represent strong agreements while lower values such as "1" represent strong disagreements. We then determined average and grouped the responses in to two categories: "Dissatisfied" and "Satisfied".

Table 27 compares the intervention and control groups' responses to various statements under nine job satisfaction domains namely: nature of job and responsibilities, work load, compensation and benefits, organizational practices and functioning, working environment, career development

and job security, performance, recognition, and overall well-being. Percentages are presented for each statement under different levels of agreement (strongly disagree, moderately disagree, neither agree nor disagree, moderately agree, and strongly agree).

Nature of Job and Responsibilities: In both groups, the majority of respondents moderately or strongly agreed that they liked their current responsibilities and tasks (Intervention: 94%; Control: 82%). The intervention group showed higher satisfaction regarding the scope to apply creative skills (84% moderately or strongly agreed) compared to the control group (69%). Both groups found their work meaningful and felt their responsibilities demand high morale, with slightly higher agreement in the intervention group.

Workload: The intervention group showed higher satisfaction with their work schedule (79% moderately or strongly agreed) compared to the control group (74%). A considerable portion of both groups had to involuntarily do extra work (Intervention: 47%; Control: 33%). A higher percentage of intervention group need to compromise their personal and social time due to their work schedules (69% moderately or strongly agreed) compared to the control group (61%).

Compensation and Benefits: The intervention group was more satisfied with their fair compensation compared to the control group (33% moderately or strongly agreed vs. 21%). The proportion of health workers that are not satisfied with non-monetary allowances is higher in the control group than those in the intervention group (71% moderately or strongly agreed) compared to the control group (66%). Satisfaction with performance-based allowances was higher in the intervention group (50% moderately or strongly agreed) than in the control group (19%).

Organizational Practices and Functioning: The intervention group showed a higher level of satisfaction with the opportunity to express opinions and engage in decision-making (78% moderately or strongly agreed) compared to the control group (71%). The intervention group reported less cordial and mutually supportive relationships with co-workers (86% moderately or strongly agreed) compared to the control group (90%). The intervention group was more satisfied with the relationships between the health facility and the community (29% moderately or strongly agreed) compared to the control group (27%).

Working Environment: Overall, the intervention group shows a higher level of satisfaction in the working environment domain. More intervention group members appreciate the overall leadership (80% vs. 78%) and client friendliness (89% vs. 81%) of their facility than the control group. The intervention group also perceives their facility as more dynamic and innovative (62% moderately or strongly agree vs. 58%) than the control group. More control group members do not appreciate the infrastructure status (62% vs. 48% moderately or strongly agree), and are not satisfied with the status of drugs (71% vs. 58% moderately or strongly agree), and status of equipment (42% vs. 54% moderately or strongly agree) of their facility than the intervention group. A higher percentage of the control group is not satisfied with their ability to provide high-quality care (37% vs. 34% in the intervention group).

Career Development and Job Security: The intervention group reports a higher chance of getting supervision and performance assessment (69% vs. 50% moderately or strongly agree), more opportunities for skill and knowledge upgrade through training (49% vs. 43% moderately or strongly agree), and feeling safer and more secure in their facility and community (77% vs. 71% moderately or strongly agree). The majority of health workers from both the control and intervention group members perceive fewer chances for promotion (67%in the intervention group and 69% in the control group) and those in the intervention group have relatively higher opportunities to upgrade their skills than the control group (49% vs. 43%). A higher percentage of the control group feels their job is not secure (72% vs. 63% moderately or strongly disagree).

Supervision and Performance Assessment: Members of the control group experience more difficulties discussing work issues with their supervisors (44% vs. 58%) and receiving regular supportive supervision on technical matters (49% vs. 63%). The intervention group feels more negatively about managing human relations, supervision, and performance assessment fairness and transparency.

Performance and Recognition: Both groups reported having high levels of competence in their tasks, with 93% of the intervention group and 94% of the control group moderately or strongly agreeing. More members of the control group feel they receive adequate recognition from patients and the community (90% vs. 87%). The intervention group reports less recognition from supervisors (54% vs. 55% moderately or strongly disagree) but receives more appreciation and recognition from co-workers (85% vs. 77% moderately or strongly agree).

Overall Well-being: The intervention group is more likely to feel burnt out, with 54% moderately or strongly agreeing, compared to 48% in the control group. Both groups report similar levels of feeling active and vigorous in the past two weeks, with 81% and 76% moderately or strongly agreeing in the intervention and control groups, respectively. The intervention group experiences a lesser extent of daily life being filled with interesting things (74% vs. 71%).
Overall job satisfaction: The intervention group has a higher level of overall job satisfaction, with 88% moderately or strongly agreeing, compared to 78% in the control group.

In conclusion, the intervention group has a higher overall job satisfaction perceives having better infrastructure and drug status in their facilities. The intervention group has also a more positive working environment in terms of leadership, management, client-friendliness, and innovation.

Table 27: Satisfaction of health workers with their job by intervention status, percentages

Job Satisfaction Domains	Statements	Strongly Disagree		Moderately Disagree		Neither Agree Nor Disagree		Moderately Agree		Strongly Agree	
		Inter-vention	Control	Inter-vention	Control	Inter-vention	Control	Inter-vention	Control	Inter-vention	Control
Nature of job and responsibilities	I like my current responsibilities and tasks that I have in this facility	2	6	2	3	3	10	40	29	54	53
	I am not satisfied with the accomplishments I could secure from this current job	27	34	24	26	11	9	30	21	8	10
	My current tasks give me a scope to apply my creative skills	2	8	4	11	10	12	54	36	30	33
	I believe my current work is meaningfulness as it meets social needs and an effective use of my time	7	7	4	10	8	11	39	35	42	37
	I have fabulous variety in tasks	1	5	5	8	18	18	51	37	26	33
	I do not have a chance to face challenges at this job	26	35	23	29	13	9	28	18	11	10
	My current responsibilities demand keeping up a high morale	2	6	2	11	9	7	50	40	36	37
	The goals of this organization are not clear to me	63	66	11	16	5	3	10	7	11	8
	I am fully aware of what are my responsibilities in this facility	0	2	2	6	3	3	28	23	67	66
	This job does not provide me an opportunity to utilize my skills and talents considerably	34	44	25	25	14	10	22	16	6	6
Work load	My current volume of work in this facility is too much	2	6	9	8	8	3	37	28	44	55
	I am satisfied with the schedule of my work at this facility	6	9	9	8	7	10	44	34	35	40
	Often I have to involuntarily do extra work in this facility	28	35	18	24	7	8	22	12	25	21
	I need to compromise my social and personal time due to work demanding work schedule	7	11	13	13	12	13	40	38	29	25
	My working hours are not too long	32	39	21	21	8	5	27	23	11	12
Compensation and benefits	I feel I am being paid a fair compensation for the work I do.	37	58	18	14	12	7	23	14	10	7
	I am not satisfied with the non-monetary allowances	19	19	12	10	4	0	21	13	45	58
	I am satisfied with the performance based allowances that I receive in this job	25	57	14	15	11	10	28	11	22	8
	I have many opportunities to be rewarded for my hardship, financially or otherwise	25	41	18	16	12	7	30	26	15	10
	I cannot appreciate the living accommodations provided to me by this job	12	27	18	14	22	5	29	23	20	31
	I am satisfied with the available schooling facilities for my children.	37	46	19	14	11	5	22	23	12	12
Organizational practices and functioning	I am not satisfied with the current organizational structure of this facility	23	26	26	25	8	7	33	27	10	15
	Many of our organizational rules and procedures make doing a good job difficult in this facility	14	17	16	23	11	16	46	32	13	13
	I do not have the autonomy to execute my tasks	31	34	25	32	17	8	22	17	6	10
	We have a participatory decision making process in this facility	7	11	10	9	8	9	41	41	34	31
	I get a fair opportunity to express my opinion and engage in decision making	4	12	7	10	10	7	46	34	33	38
	My professional relationship with the co-workers is very cordial and mutually supportive	2	3	6	3	6	5	41	32	45	58
	We do not share our personal matters among co-workers and render any personal affinity	36	43	19	27	8	12	28	12	8	7
	My Working relationships with the Management staff within the health facility is not that cordial	45	30	22	33	6	6	19	15	9	16
	I have a respectful and cordial relationship with the District/ Ministry of Health staffs	6	13	5	13	16	11	47	30	27	33

	I am not convinced about the relationships between this health facility and the community	38	40	23	23	10	10	23	18	6	9
	I appreciate the existing conflict resolving practice among staffs and supervisors in this facility	4	6	8	9	13	7	46	39	29	38
	The staffs and the facility- in -charge are on a regular communication	8	11	10	12	11	10	48	37	24	31
	The communication between this facility and the higher authorities is satisfactory to me	5	12	8	12	13	11	53	43	21	23
	The overall functioning of this facility is transparent	6	10	7	15	9	6	51	48	27	21
	The promotion practices are not fair in this facility	24	17	26	23	14	14	29	20	7	25
	The payment of performance incentives are transparent and fair enough in this facility	13	15	15	30	17	13	34	25	21	18
	The prevailing feedback mechanisms is not transparent in this facility	21	31	27	23	19	14	27	19	6	13
Working environment	I appreciate the overall leadership of this facility	6	11	7	7	8	4	52	46	28	32
	I am satisfied with the management of the health facility by the internal staffs	6	11	7	6	10	4	47	46	30	34
	I do not appreciate the infrastructure status of this facility (e.g. beds, waiting rooms, toilets etc.)	14	12	28	17	11	10	31	23	17	39
	I am not satisfied with the status of drugs (i.e. availability and quality) in this facility	13	5	14	16	15	8	40	34	18	37
	I am satisfied with the status of equipment (i.e. availability and quality) in this facility	14	28	14	23	18	8	42	25	12	16
	I am not convinced about the availability of supplies in this facility	13	11	21	14	19	13	38	33	8	29
	I am satisfied with the overall physical condition of this facility	12	29	15	18	11	4	47	37	16	12
	I often have to work harder because of the incompetence of the people I work with.	20	24	13	14	12	13	37	35	18	13
	I appreciate the client friendliness of this facility	2	3	3	9	6	8	44	43	45	38
	My facility is very dynamic and an innovative place. People are willing to take risks to do a job well-done.	6	12	12	17	20	12	45	44	17	14
	I am not satisfied about my ability to provide high quality of care given the current working conditions in this facility	34	28	22	25	9	10	25	26	9	12
	This facility has frequent turn over of staffs	12	25	16	20	16	12	43	24	14	20
	The changes (functioning, staffing, infrastructure etc.) happening in this facility are acceptable to me	8	21	10	20	16	11	55	35	12	13
Career development and job security	There is really too little chance for promotion on my job	8	11	15	18	11	3	38	31	29	38
	I have opportunities to upgrade my skills and knowledge through training in this job	20	29	17	22	14	6	31	29	18	14
	I feel my job is not secured	35	41	27	30	12	8	20	11	6	10
	I feel safe and secured to work in this facility and community	3	13	10	6	10	10	42	38	35	33
	Given an opportunity, I wish not to continue my job with this facility	14	18	15	17	15	9	33	25	23	32
	Supervision and performance assessment	6	21	11	17	15	13	57	32	12	18
	I rarely have the opportunity to discuss work issues with my immediate supervisor	10	27	16	15	15	9	38	32	20	17
	I receive regular supportive supervision on technical matters	9	27	12	14	16	10	45	31	18	18
	When it comes to managing human relations, I do not get regular and encouraging supervision	20	24	24	32	19	12	27	22	10	11
	My supervisor shows too little interest in the feelings of subordinates	22	23	22	21	17	20	34	22	6	14
	The performance assessment in this facility is regular	9	22	13	14	14	11	41	29	23	24
	I feel, the performance assessment in this facility is unfair	33	27	26	27	12	15	23	17	6	16
	I am convinced that the performance assessment in this facility is transparent	6	16	11	13	18	13	44	39	21	19

	I receive optimum professional support in this job, especially from my supervisor	6	18	15	15	21	12	44	36	15	19
Performance	I am competent to undertake tasks demanded by this job	1	0	2	4	4	2	43	42	49	52
	My level of performance as of today in this job is not satisfactory to me	35	34	22	25	9	7	22	20	12	14
Recognition	My supervisors do not recognize my work adequately	29	32	25	23	7	12	30	19	9	14
	I receive much appreciation and recognition from my co-workers	2	9	6	11	8	4	50	47	35	31
	I feel, I receive adequate recognition from the patients and community	2	3	8	7	3	1	39	46	48	44
Overall well-being	I often feel burnt out while at work	12	21	20	21	14	10	42	31	12	17
	In the past two weeks, my daily life has been filled with things that interest me….	8	12	6	10	12	7	56	44	18	27
	In the past 2 weeks, I have felt active and vigorous…	7	8	6	6	7	10	61	48	19	28
	In the past 2 weeks, I could not wake up feeling fresh and rested…	27	20	22	36	6	11	35	19	11	14
Overall satisfaction on job	Overall, I am satisfied with my job	2	6	5	7	6	9	46	48	42	30

Table 28 presents the results of a regression analysis conducted on Likert scale scores comparing health workers in intervention and control groups at baseline and endline on various domains of job satisfaction, which include nature of the job and responsibilities, workload, compensation and benefits, organizational practices and functioning, career development, job security, performance, recognition, and working environment. The results show that the proportion of workers who expressed satisfaction with compensation and benefits significantly ($p<0.05$) increased from 42.2% to 70.5% between baseline and endline. Additionally, the percentage of workers in the intervention facilities who reported satisfaction was significantly ($p<0.05$) higher than their control counterparts in the domains of compensation and benefit, career development and job security, overall satisfaction on job domains, and overall total satisfaction.

On the other hand, the analysis shows a statistically significant ($p<0.05$) decline in the workload domain from 57.1% at baseline to 40.9% at the endline. A similar decline was also observed in the performance domain of workers' satisfaction, with the percentage of workers reporting being satisfied declining from 62% at the baseline to 50% at the endline.

Table 28: Comparing job satisfaction between intervention and control facilities: Results of a bivariate analysis on various satisfaction domains

	TERM		Intervention Status	
Satisfaction Domains	Baseline	Endline	Control	Intervention
Nature of job and responsibilities	55.5%	54.5%	50.9%	57.7%
Work load	57.1%	40.9%*	41.2%	55.2%
Compensation and benefits	42.2%	70.5%*	38.1%	65.6%*
Organizational practices and functioning	54.0%	59.1%	50.0%	60.2%
Working environment	46.0%	55.3%	45.6%	53.0%
Career development and job security	51.2%	62.1%	41.2%	65.4%*
Performance	62.0%	50.0%*	57.9%	55.8%
Recognition	48.8%	56.1%	51.8%	52.2%
Overall well-being	49.4%	42.4%	47.4%	45.6%
Overall satisfaction on job	76.8%	69.7%	64.0%	79.7%*
Overall Total Satisfaction	51.8%	64.4%*	47.4%	63.7%*

The evaluation team fitted a logistic regression model incorporated into the DID framework to assess the impact of the project on worker satisfaction. The results of the DID model are presented in Table 29 below.

The results presented in Table 29 suggest that the program had positive and statistically significant impact on workers' satisfaction with compensation and benefits ((0.827, AOD=EXP (0.827) = 2.29, p<0.01)), work load (1.272, AOD=EXP (1.272) = 3.57, p<0.01) and overall satisfaction on the job ((0.911, AOD=EXP (0.911) = 2.49 p<0.05)). Health workers' satisfaction with compensation and benefits is understandable given the fact that part of the program subsidy was utilized to incentivize health workers. However, the impact of the program on other satisfaction domains was not statistically significant (p>0.05), implying that the intervention did not have a significant impact on nature of job and responsibilities, organizational practices and functioning, working environment, career development and job security, performance, recognition, and overall well-being.

Table 29: Impact of a project on job satisfaction: Results from a logistic DID regression analysis on multiple domains

	Satisfaction Domains										
	Nature of job and responsibilities	Work load	Compensation and benefits	Organizational practices and functioning	Working environment	Career development and job security	Performance	Recognition	Overall well-being	Overall satisfaction on job	Overall Total
Interaction (Time*Intervention)	-0.303	1.272**	0.827**	-0.566	0.623	0.0110	-0.368	-0.353	0.253	0.911*	0.0470
	(-0.83)	(3.10)	(2.10)	(-1.49)	(1.68)	(0.03)	(-1.00)	(-0.96)	(0.69)	(2.33)	(0.12)
Term (1=Endline)	0.123	-1.451***	0.670*	0.552	0.0574	-0.472	0.509	0.582	-0.450	-0.903**	0.552
	(0.37)	(-3.81)	(2.00)	(1.59)	(0.17)	(-1.43)	(1.52)	(1.74)	(-1.34)	(-2.61)	(1.62)
Sex of health worker's (male=1)	-0.0614	0.574	-0.305	-0.427	-0.0755	-0.0414	0.0406	-0.287	-0.139	0.378	0.131
	(-0.21)	(1.86)	(-0.98)	(-1.45)	(-0.25)	(-0.14)	(0.14)	(-0.97)	(-0.47)	(1.05)	(0.44)
Marital status (Married=1)	-0.0309	-0.315	-0.167	-0.0964	0.0954	-0.359	-0.199	0.183	0.0769	-0.0867	0.122
	(-0.14)	(-1.38)	(-0.71)	(-0.43)	(0.43)	(-1.59)	(-0.89)	(0.82)	(0.35)	(-0.34)	(0.55)
_cons	0.374	0.0821	0.351	0.842	-0.0144	1.110*	0.250	-0.110	0.0189	0.905	-0.285
	(0.71)	(0.15)	(0.64)	(1.59)	(-0.03)	(2.06)	(0.48)	(-0.21)	(0.04)	(1.46)	(-0.54)
N	294	293	291	293	294	293	294	293	294	294	294

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Impact of program on health workers' motivation

The evaluation also assessed the impact of the project on workers' motivation. For this purpose, a total of 42 items were used to measure the level of worker motivation, which was then divided into two categories: "motivated "and "not motivated." The resulting dichotomized data was then aggregated to determine the overall level of motivation among workers.

The logistic regression results presented in Table 30 demonstrate a statistically significant difference (B = 0.779, SE = 0.395, Wald = 3.894, df = 1, p = 0.048) in worker motivation between control and intervention facilities. This suggests that, after accounting for other factors in the model, health workers in the intervention group have 2.18 times higher odds of exhibiting greater motivation compared to those in the control group after the project intervention. The result of the regression analysis also shows that, other factors, such as sex, marital status, and employment type, did not exhibit a significant influence on worker satisfaction. Hence, the PBF program appears to have improved workers' motivation.

Table 30: Logistic regression result: the effect of the PBF program on health workers' motivation

Variables in the Equation							
		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Interaction	.779	.395	3.894	1	.048	2.180
	Sex of the health worker	-.387	.328	1.389	1	.239	.679
	Marital status	-.021	.239	.007	1	.931	.980
	Type_of_employment_cat_2	-.777	.677	1.315	1	.251	.460
	Years worked as a health worker at this facility	.050	.035	2.099	1	.147	1.052
	TERM	-.675	.362	3.475	1	.062	.509
	Constant	.841	.821	1.050	1	.306	2.319
a. Variable(s) entered on step 1: Interaction, 1.01 sex of the health worker, 1.03 marital status, type_of_employment_cat_2, 1.07 Years worked as a health worker at this facility?, TERM.							

In conclusion, the quantitative findings suggest that the intervention had a positive impact on the motivation, and satisfaction of health workers with salary and bonus.

Improvements in motivation and performance of health workers in intervention health facilities did not also go unnoticed by patients and other key informants. Most patients that participated in the FGDs observed improvements in the motivation, friendliness and respectfulness of health workers. This has attracted many patients to the health facilities and increased utilization of health services in those facilities. Availability of health workers is also reported to have increased over the past few years. *"There is a great difference in terms of health staff respectfulness and sympathy as compared to before. Before two years it was very challenging even to get health workers particularly on Monday and Friday, but currently there is no problem to get health workers every day,"* a patient said suggesting increased availability of health workers to provide services.

The program, according to some key informants, has contributed to the retention of health professionals by providing incentives and making the health facilities to become attractive to work. A key informant in this regard said *"There was a professional turnover before. Since PBF program started, they have become stable in the same place and performing well."*

Acknowledging improvements in access to quality services, some patients noted shortage of health workers in the intervention health facilities, which may also be leading to delivery of compromised quality care. *"…There is good change. The approach of health care provider is friendly*

and consulting is also interesting. But when there is large number of patients, they do hurry and skip some examination," a patient stated. Another patient also emphasized that *"… there is change, but not adequate. The number of clients who need service has been increasing from time to time whereas the number of staff didn't grow in line with the demand. So, this needs attention and resolution mechanisms."* The results suggest that although the project has increased accessibility and utilization of health services, health facilities are not increasing their workforce to keep up with the growing demand.

3.4.2. Outcome 2: Improved Governance of Health Service Delivery

The second target outcome of the PBF program was improved governance in health service delivery through increased capacity at the level of WorHO and ZHDt to perform their regulatory tasks and provide supportive supervision; and institutionalization of PBF in the Ethiopian health system. The findings under this outcome area are presented below.

3.4.2.1. Increased capacity at the level of WorHO and ZHD to perform their regulatory tasks and provide supportive supervision

Qualitative data analysis results suggest that program has improved the capacity of the local government to perform their regulatory tasks and provide supportive supervision to the health facilities. The capacity of Zonal and woreda health experts to provide technical feedback and constructive support has also been enhanced. After the implementation of the program, a key informant noted that *"there has been improvements in the capacity of local government to perform their regulatory tasks and provide supportive supervision to the health facilities."* Another informant also acknowledged that *"there are some improvements regarding the capacity of Woreda Health Offices and Zonal Health Department to perform their regulatory tasks and provide supportive supervision."* But this same informant emphasized that *"it is too soon to tell whether this capacity is transformed to institutionalize PBF in the Ethiopian health system."*

Apart from capacity building, the provision of incentive enabled WorHOs/ZHDsto fulfil their regulatory mandates, which used to be difficult to do due to financial resource limitations. An informant, for example, highlighted that *"PBF is necessary to regulate the health practices of our woreda."* The informant also noted that the PBF program has *"solved finance problems and improved institutional efficiency."* According to the informant, *"it was difficult to regulate due to finance problems."*

An informant noted that the PBF program has *"strengthened the capacity of the leadership"* and facilitated monitoring of health activities. Another informant emphasized that the PBF program improved the governance of health service delivery, particularly through the effective utilization of allocated finances and regular performance monitoring. Several informants highlighted the role of the PBF program in strengthening monitoring and evaluation mechanisms. For instance, one informant noted that the program's monitoring and evaluation tools have increased the capacity of the woreda and health facilities in assessing the quality of health care services. The improved capacity of WorHOs/ZHDs seems to also be recognized by health facilities. According to an informant, *"…the Zone as well as woreda experts ….provide constructive support for the health facilities." During the supervision, the comments provided by the regulatory team were considered "very helpful in terms of improving the quality of health services particularly in delivery room, nutrition departments and drug storage."*

The program has provided valuable lessons at the office and institutional levels, such as the importance of regular performance follow-ups, checklist-supported supervision, and customer satisfaction assessments for quality service provision. The guidelines and standards developed for the PBF program monitoring and evaluation have been useful tools for both regulators and health

facilities, serving as reference materials during activity performance monitoring and regulation, and enabling them to understand what quality health service mean. One informant explained that the program's quality assessment tools have helped health workers better understand health service quality and improve their practices.

The project has also improved governance of health service delivery by strengthening the capacity of health facility management. A key informant in this regard said *"Previously there was no good leadership starting from lower to higher levels. But, currently the PBF program has created strong leadership."* Good leadership and manager led to good commitment from health care provider and improved quality of services. The program was also said to have reduced bureaucracy or hierarchy and management negligence by promoting regular evaluation and monitoring. The program has also promoted institutional efficiency by enabling health facilities prepare integrated plan, which was not the case in the past.

In conclusion, the informants' responses indicate that the PBF program has made notable contributions to improving the regulatory capacity of WorHOs and ZHDsby addressing financial challenges, strengthening their leadership and governance capacity, and improving monitoring and evaluation mechanisms.

3.4.2.2. Institutionalization of PBF in the Ethiopian health system

One of the pathways adopted by the PBF program to achieve its overarching objective was by improving governance of the health service delivery through increased capacity and institutionalization of PBF in the Ethiopian health system.

In regards to in institutionalization, opinions differ among stakeholders about the current state of institutionalization of the PBF program in the Ethiopian health sector. But many agree that there is growing ownership of the program by government and health facilities. The PBF program has been well-received by health facilities, creating good awareness and buy-in among different stakeholders. An informant, for instance, highlighted that *"the PBF program is well received by the health facility. ...the program has been increasing the utilization of health care and quality care thereby improving maternal and child health."* The program is also well received by health workers. *"The staff love this project and we are working to improve health care quality together with PBF,"* an informant stated.

The PBF program has contributed to increasing the number of facilities providing quality health services. One informant explained that the program's quality assessment tools have helped health workers better understand health service quality and improve their practices. Another informant highlighted the role of PBF in ensuring the availability of essential medical equipment, generators, and medicines, stating, *"I wish it could continue as it started."* Stakeholders also agree that the program has been beneficial in terms of increasing utilization of health care, improving quality of care, and improving health outcomes such as decreasing maternal and neonatal mortality. For example, an informant mentioned that the program has brought about *"radical change on reducing home delivery and increasing institutional delivery as well as decreasing neonatal death in the community."* PBF is identified as one of the strategies that MoH plans to consider to address funding gaps in implementing its second Health Sector Transformation Plan (HSTP 2). This could indicate government recognition of the potential of the PBF program. The fact that PBF is indicated as one of the health care financing mechanisms could be taken as an important step towards institutionalization of the program in the Ethiopian health care system. The project was appreciated by many key informants as it introduced and demonstrated the potential of the PBF

program in improving access to and utilization of quality health services. A key informant in this regard said *"... I can for sure say that the program has played vital roles in introducing PBF to the health sector."*

The contribution of PBF to supporting government strategies and policies was also emphasized. As one informant pointed out, *"PBF plays an important role in supporting government strategies and policies. ... PBF has a big part to play in supporting that."* Another informant reiterated the importance of PBF in improving health institutions and aligning with government systems: *"It plays an important role in improving health institutions and establishing the systems required by the government. So it is going in line with the government system."* As a result, almost all stakeholders at all levels reported having positive attitudes towards the program. Hence, the fact that the program is widely supported and embraced by stakeholders within the government system may suggest growing opportunities for the program to be institutionalized. Growing ownership of the PBF program by policy makers and other stakeholders could be one indicator of institutionalization. Growing recognition of the program by the government can be exemplified the statements quoted from a key informant:

> *"At the start of the program, there was some hesitation [from government] in accepting the essence and importance of the program. But, after the apparent result registered by the PBF program in Borena, we became actively involved in the program design, implementation, follow-up, and review process."*

Although informants generally acknowledge that the PBF program aligns with the government's strategies and policies, they also point out that much still remains in fully institutionalizing the PBF program in the Ethiopian health care system. Many doubt the sustainability of the outcomes/results achieved by the program when/if the program phases out because of their belief that PBF is not institutionalized in the health care system in Ethiopia. Although the program is very well received and has motivated health facilities to improve quantity and quality of health care services, and enhance information management and governance, many question if these outcomes will continue in the absence of continued support from the project or other bodies with such a financing mechanism. An informant emphasized *"I believe that it is too soon to talk about the issue of institutionalization of PBF in the Ethiopian health system at this stage of the project."* This is a shared view by many key informants.

Several challenges were identified as impediments to the institutionalization of the PBF program. The major challenge mentioned by many stakeholders relate to the inadequate ownership of the program by local and federal government bodies. The role of key federal government bodies in the program was not also as strong as expected importance of the program. A key informant in this regard noted that *"the government is not [paraphrased] owning and institutionalizing PBF as one mechanism of health care financing."* The fact that the program had short duration in some health facilities was also identified as a factor that could constrain institutionalization of the program at even institution level. An informant noted that "*If institutionalized, it can brings excellent results. But currently we can't say it is institutionalized fully… because we joined [paraphrased] the program recently."*

Sustainability concerns may also challenge the institutionalization of the PBF program. Some informants express doubts about the sustainability of the PBF program, questioning whether the financial benefits and incentives provided to health facilities and professionals can be maintained in the long run. This concern may pose a challenge to the program's institutionalization, as stakeholders may be hesitant to commit to a financing mechanism with uncertain long-term

viability. An informant, for example, noted that *"if the program stops operating there will be a great chance of going back to previous status,"* suggesting that health facilities' reliance on the program.

Overall, the findings suggest that institutionalization of PBF in the Ethiopian health system is in its early stages. But the program played important roles in demonstrating PBF as an effective health care financing mechanism in Ethiopia. Apart from indicating PBF as one of the alternative financing mechanisms for the HSTP II, we learned that the government is considering to initiate PBF at a national level, which can be considered as the first step in the institutionalization process.

3.4.3. Outcome 3: An Enhanced Health Information System

The third outcome of the program is enhanced health information system that supports data-based decision making at woreda, zonal and regional level, and additional financing potential for the health system through enhanced transparency. The end-term evaluation results on these two result areas are presented below.

3.4.3.1. Enhanced health information system that supports data based decision making at Woreda, Zonal and Regional level

Quality health data is essential for the successful implementation and evaluation of Performance-Based Financing (PBF) programs. It is with this recognition that the program also aimed to enhance the health information system. The end of program evaluation assessed the contribution of the program in improving health information system using quantitative and qualitative approaches.

The data reliability, which is the quotient of verified and declared health service utilization data, was the quantitative approach used to assess the program's contribution to generating and reporting quality data. This is based on the assumption that data reliability can serve as an indicator of data quality, as a higher percentage of verified data from the declared may suggest accuracy in data collection or reporting. If data reliability is high, it implies that there is a high level of data quality, with minimal inaccuracies or errors in data collection and reporting. Service utilization data reporting data reliability were calculated for all quantity indicators excluding waivered indicators (OPD Adults, OPD under 5, Major Surgery) for all intervention health facilities by comparing declared data with verified data. Trends in data reliability are presented and analysed below for each zone.

Figure 6 depicts the quantity of service utilization for all indicators along with the data reliability and the declared and verified quantity data from 2015 to the 3rd quarter of 2022 for Borana health facilities by phase and year.

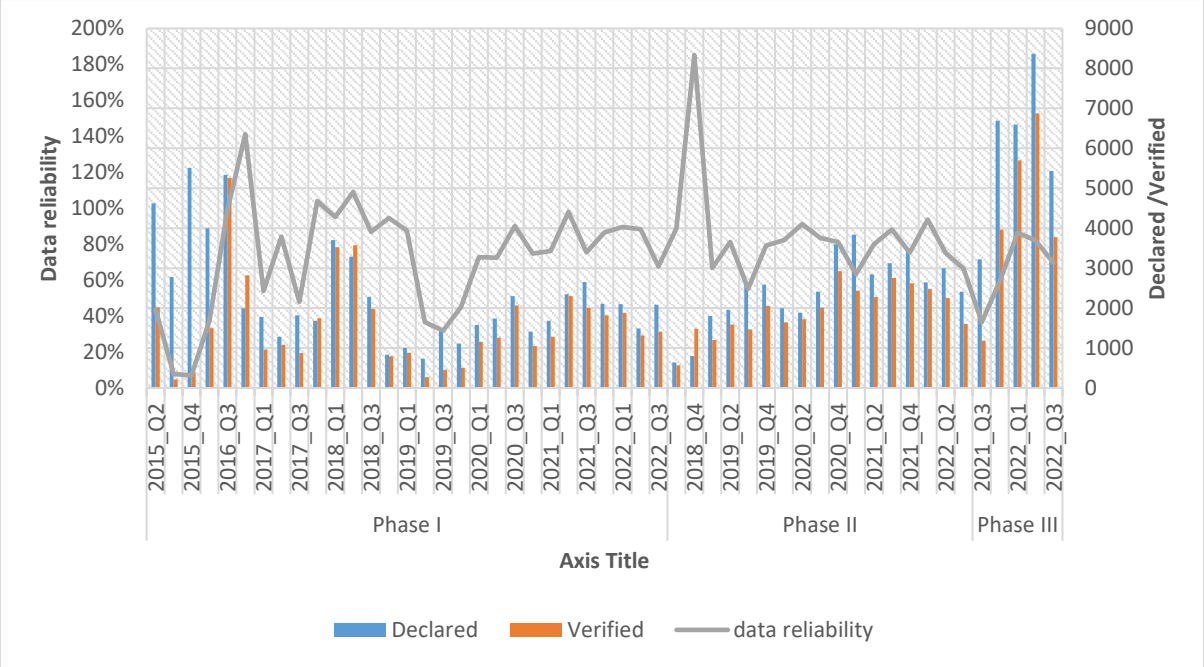


Figure 6: Comparison of declared and verified quantity data with data reliability by phase and year in intervention health facilities: Borena

The results depicted in Figure 6 above shows that, in the first phase, the data reliability ranged between 7% (2015_Q4) and 141% (2016_Q4). The highest declared figures were reported in 2015_Q2 (4,625) while the lowest in 2022_Q3 (644). The highest verified figures were also recorded in 2015_Q2 (2021) and the lowest in 2022_Q3 (572). It should be noted here that declared figures showed steady decrease just after the project inception indicating that the facilities reduced overstating the number of patients who have received their services.

In the second phase, an increasing trend in both declared and verified data was observed. The declared data values ranged from 802 in 2018_Q3 to 8,367 in 2022_Q2, whereas the verified data ranged from 1,485 in 2018_Q3 to 6,871 in 2022_Q1. In comparison to Phase I, Phase II data reliability is generally higher, with values fluctuating between 37% in 2019_Q1 and 185% in 2018_Q3. In the third phase, which started from 2021_Q3, data reliability remained relatively stable compared to the previous phases, with values varying between 70% in 2022_Q3 and 82% in 2022_Q1. The highest declared figures were reported in 2021_Q1 (6,595) while the lowest in 2022_Q3 (5,428). The highest verified figures were also recorded in 2022_Q1 (5,694) and the lowest in 2022_Q3 (3,777). (For details, please refer Appendix 2).

Figure 7 below shows the data reliability and compares declared and verified quantity data for the Jimma intervention zone, divided into two phases and various quarters of the Gregorian calendar from 2019 to 2022.

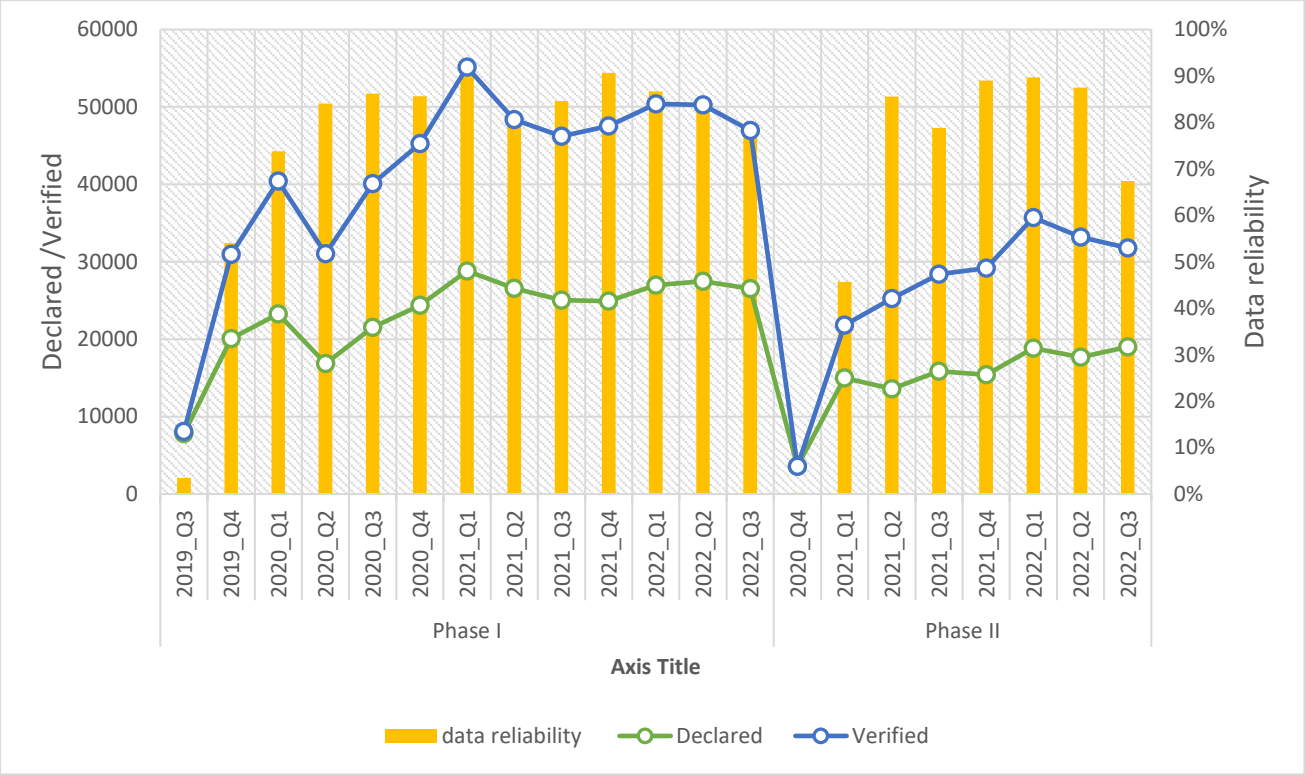


Figure 7: Comparison of declared and verified quantity data with data reliability by phase and year in intervention health facilities: Jimma

As Figure 7 above shows, for facilities that entered into the project during the first phase of implementation in Jimma zone, declared data ranged from a low of 7,820 in 2019_Q3 to a high of 28,823 in 2021_Q1, while verified data ranged from a low of 274 in 2019_Q3 to a high of 26,299 in 2021_Q1. Data reliability in Phase I improves over time, with an initial value of 4% in 2019_Q3, peaking at 91% in 2021_Q1 and 2021_Q4.

During the second phase of project implementation in Jimma zone, an overall increasing trend in both declared and verified data was observed, with some fluctuations. The declared data values ranged from 3,589 in 2020_Q4 to 18,989 in 2022_Q3, while the verified data ranged from 10 in 2020_Q4 to 16,888 in 2022_Q1. Data reliability in Phase II displays significant improvement, with values ranging from 46% in 2021_Q1 to 90% in 2022_Q1. (For details, please refer to Appendix_2).

Figure 8 shows the declared and verified quantity data along with data reliability in West Gojjam zone.

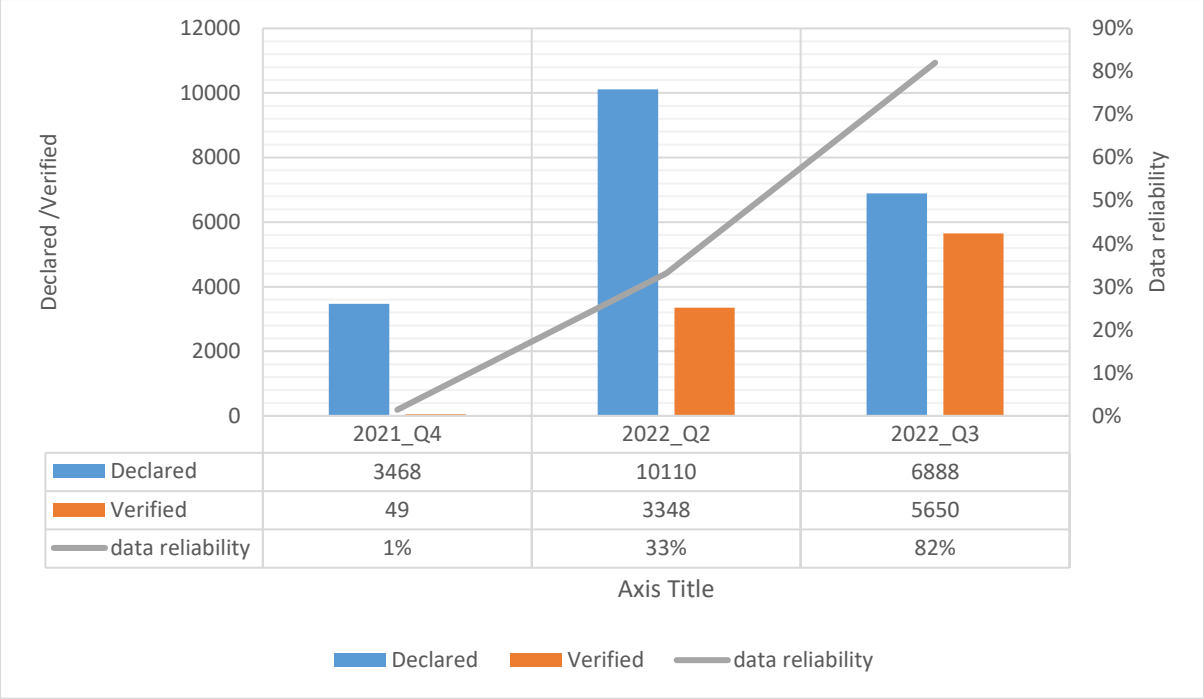


Figure 8: Comparison of declared and verified quantity data with data reliability by phase and year in intervention health facilities: West Gojjam

The results depicted in Figure 8 above shows that, in the fourth quarter of 2021 (baseline), there were 3,468 declared data, of which only 49 were verified. This resulted in a very low data reliability of 1%. In the 2022_Q2 reporting period, the declared data increased to 10,110, with 3,348 of these data points being verified. As a result, the data reliability improved to 33%. During the 2022_Q3 reporting period, the declared data amounted to 6,888, with 5,650 of these data points verified. This led to a significant improvement in data reliability, reaching 82%. The substantial increase in data reliability compared to the previous quarters suggests that efforts to enhance data quality processes or strengthening HMIS have been effective.

Table 31 below presents the zonal and grand total declared and verified data, as well as data reliability. As shown in Table 31, it appears that Jimma had a higher data reliability of 79%, indicating that the declared and verified quantities were more closely aligned in that zone. Borena had a relatively high data reliability of 73%. Overall, the results indicate that data quality has shown significant improvements when compared to the year health facilities entered the program. On the other hand, West Gojjam had an extremely low data reliability of 44%, signifying a significant discrepancy between the declared and verified quantities in that zone, which could be due to the fact that health facilities from this zone joined the program lately.

Table 31: Total declared and verified data, and reliability percentages for all indicators in all intervention health facilities by zone

Intervention Zone	Declared	Verified	Data reliability
Boreana	138,631	100,915	73%
Jimma	419,340	330,220	79%
West Gojjam	20,466	9,047	44%

The contribution of the program to enhancement of health information system was also assessed qualitatively by gathering opinions of different key informants. Most informants provided positive feedback regarding the impact of the PBF program on health information systems, data-based decision making, and transparency. The program has helped change the attitude of health workers towards health information management and value of data, resulting in proper data handling and

management. One informant stated, *"Before the program started, there was negligence in handling health information in our health center. PBF program has changed this attitude through data related evaluation."* As a result, *"The staff register each piece of information on registration book carefully because they will be evaluated later,"* an informant stated. Following the implementation of the project, health facility staff "*only report the real data because they understand the benefit of the data,"* according to an informant. This is partly because *"The knowledge and skill of health workers on importance of data and full recording of patient history was significantly enhanced due to the implementation of this project,"* according to another informant.

The PBF program has also improved data management and documentation, with an incentive paid for the timeliness of HMIS reports. *"There is an incentive paid for timeliness of HMIS report. Before, you do not find report in hard copy in woreda offices. But now you can get reports of every month in hardcopy,"* a key informant said. The positive impact of the project in improving data quality is also well recognized at health facility level. A HEW, for example, noted *"At the health post level, the implementation of this program has significantly contributed to minimizing errors that may occur during referral and has improved client information/data registration. As a result, we place more emphasis on working on the right things and carefully registering cases than usual."*

The program has also enhanced information-based decision making at all levels. The key informants noted that the quality of data recording has greatly improved, which has allowed for more accurate and reliable data to be used for decision making at various levels. One informant noted, *"... the improvement in documentation and appropriate registration of patients' data contributed to the availability of quality health information. The information is also used for various decision-making purposes."* Another informant also noted, *"... the improvement in quality of data has significantly helped the zone make evidence-based decisions."* A similar view was also shared by another informant who noted that *"There is a big change in terms of data recording and organization at health facilities. The usage of information for designing as well as decision making intervention had shown a big change as compared to the previous time." "...the improvement in quality of data was aiding us in preparing an appropriate plan,"* another informant stated.

Key informants also noted that the program has contributed to the coordination of health-related data in health facilities, and there are now various HMIS toolkits available to monitor the top 10 diseases. Performance monitoring is now conducted regularly, and identified problems are dealt with an action plan to enhance the quality of health service, according to informants. Following the program implementation, key informants indicated that data has received more attention from decision makers and health workers. *"Data had not been given as much attention as it is now,"* an informant noted.

In conclusion, qualitative findings suggest that the PBF program has had a significant impact on the quality of health data and evidence-based decision making, and promoted transparency through the provision of technical and financial support. The program has changed the attitude of health workers towards the importance of quality data, and increased their knowledge and skills of health workers in data management. The program has also strengthened the information management system which led to a significant improvement in the registration and documentation of health-related data. These contributed to generation and reporting of quality health data that also led to increased use of the data to inform decision making. In sum, key informants noted that the PBF program improved the collection, documentation, and organization of health data, which were previously lacking.

Overall, qualitative findings suggest that there has been improvements in generation and reporting of quality health data. Error margins or differences between declared and verified data also narrowed after the health facilities started receiving support from the program.

3.4.3.2. Enhanced health information system that supports additional financing potential for the health system through enhanced transparency

The PBF program has enhanced transparency and could lead to the mobilization of additional financing for the health system through the following ways:

- **Improved quality and quantity of service**: The positive impact of the PBF program in improving quality of health service provision is well established by this evaluation study. According to many informants, the improved quality and quantity of health services at the intervention facilities will more likely attract more income for them in the future. For example, an informant highlighted *"… as you perform well and more, the volume of finance allocated for you also increases. So, this project directly …gives a better chance to get an additional income because the financial capacity of the health facility is most of the time determined by the quality as well as the scope of health service provided."* Another informant also highlighted, *"From this project what I learned is how quality data handling has a significant role in quality service provision,"* a key informant stated. A similar view was also shared by another informant who noted *"The project has resulted in quality data. When there is quality data and there is quality service."*

- **Improved data management and reporting**: The program emphasized accurate, timely, and complete data collection and reporting. Better data management would help health facilities and policymakers make informed decisions, identify gaps in service delivery, and allocate resources more effectively, potentially attracting additional funding from various sources. An informant noted, *"The program also enhanced transparency through real [data] reporting system which will help to mobilize additional finance."*

- **Incentivizing performance:** The program linked financial incentives to the achievement of specific performance indicators, encouraging health workers and facilities to improve their performance. This results in better results, which may attract further investments from both public and private stakeholders.

- **Enhanced accountability and transparency:** The program established clear targets and expectations, promoting accountability among healthcare providers. Health facilities prepared business plans and the implementation of their plans were regularly monitored. An informant in this regard said *"PBF promoted accountability and transparency through strong monitoring and evaluation."* The business plan along with regular monitoring and evaluation of performance has helped ensure that resources are used efficiently and as intended, which could increase the confidence of donors and government in the system's effectiveness. Moreover, *"the autonomously usage of finance and management by management committee allows the finances to be used for appropriate interventions,"* a key informant mentioned.

  The program is also said to have enhanced transparency by improving data quality at health facilities. The program has enhanced the data management and reporting system and has made it difficult for staff of health facilities to report false or inaccurate data. One informant stated, *"It enhanced our reporting system. Previously the staff were reporting what was not done, but now they can't do that. They can only report the real data because they understand the benefit of the data."* This suggests that the project has helped health facilities to become more transparent by ensuring that they manage their data properly and report accurate data. The program's approach of requiring health facilities spend the subsidies in line with the priorities set in their business plans and establishment of committees also helped facilities to become more transparent. One informant noted, *"Almost all of the decisions have been made by the management committee and based on the*

business plan…, so this has enhanced transparency …and will aid the facility in mobilizing more finance." The improved transparency, according to many key informants, will help health facilities mobilize finance and other resources in the future.

- **Increased financial autonomy**: The PBF program granted health facilities greater financial autonomy, allowing them to allocate resources based on local needs and priorities. This flexibility and adaptability might have led to improved efficiency and better performance, making them more attractive to potential funders.

- **Strengthening governance**: The program involved the establishment of management committees drawn from different sectors. These committees have been a crucial role in overseeing the allocation and use of resources, implementation of planned activities, enhancing transparency, and building trust among stakeholders. The program also strengthened capacity of health facilities. *"They gave various trainings on how to improve internal revenue of the health centre. The training has ensured budget transparency in all activities carried out as a health centre,"* an informant said.

- **Demonstrating results**: By focusing on measurable outcomes, the program helped demonstrate the impact of investments in the health system. This evidence of effectiveness can be used to attract additional financing from donors, government agencies, and private investors.

In conclusion, the program has enhanced transparency and may lead to the mobilization of additional financing for the health system by improving data management, incentivizing performance, enhancing accountability, increasing financial autonomy, strengthening governance, and demonstrating results. These factors will help build trust and confidence in the health system, making it more attractive to potential funders.

3.5. Efficiency

3.5.1. Efficiency of processes in achieving results

The management processes in place for the program were appropriate in supporting delivery. According to most key informants, the implementation and management arrangement is effective in achieving desired results. The program's M&E mechanism has worked well in improving service quality and utilization, enhancing capacity of regulatory bodies and health facilities, and in identifying and resolving challenges faced within health facilities. Regular monitoring and evaluation has made health institutions work harder. One informant noted that *"finance and resource management are very good"* due to the program's efficient use of resources. Additionally, the program's approach, which involves providing funds and monitoring their implementation, is considered effective by the informants.

To most stakeholders, the program strategies and tools used in the implementation of the program have been effective, and the program inputs and strategies were realistic, appropriate and adequate to achieve the results. The management committee established at health facilities facilitated planning, monitoring of performance and improvement in quality of health services as well as resolving the challenges of the health facilities.

The overall efficiency of the PBF program is very good, according to key informants. They support their claim by stating the program has been implemented with a small number of staff and has been able to achieve significant results. But it is important to also note here that shortage of program staff was blamed for inadequate follow-up support after verification of performance. The monitoring and evaluation system has also been strong and scheduled, which has led to improved

performance. The actual or expected results (outputs and outcomes) justify the costs incurred, according to many stakeholders. The participants also noted that the program has been successful in achieving its objectives within a short period of time with a small number of staff. Overall, the PBF program was considered efficient in terms of resource utilization and achieving desired results by most stakeholders. But the program could have been more efficient if it strengthened the linkage between health post and health centers, capacitated the management committees in financial knowledge and skills, centrally bought and distributed medical equipment and medicine. Furthermore, the management structure was good because it rewards those who do the work and punishes those who do not.

3.5.2. Collaboration with the government

The program has maintained good collaboration with the government at all levels, CBOs, Steering Committees. CBOs and the Steering Committees have been successful in contributing to program implementation and its achievement of results. The CBOs have been instrumental in assessing the quality of health services, collecting feedback from the community, and verifying administrative data. The CBOs have direct contact with the community and have been able to identify gaps in the health service delivery system, which simplifies planning for better implementation of the program. They collect comments and complaints from the community regarding drug supply and service quality, and assess patients' satisfaction with health service. The Steering Committees that are comprised of members from finance, health and various offices have also been actively engaged in providing direction, evaluating progress, and finding solutions for identified challenges.

The strong collaboration has been beneficial in terms of program implementation and improving the chances of sustainability of some of its results, with all stakeholders showing enthusiasm and interest. The program has also been able to exploit available resources through collaboration with the local government.

3.5.3. Duplication of effort

The findings suggest that the program was designed to be unique and to avoid any overlap or duplication with other initiatives. The program office worked with the regional health bureaus and other stakeholders to ensure that the program was aligned with other projects and that there was no duplication of efforts. The program was also designed to support other programs and initiatives, such as the national immunization campaign and community-based health insurance. There were no donor or project dedicated to the quality of care delivery in the same way as the PBF program, according to many key informants.

The PBF program does not appear to overlap with other interventions or programs of similar nature in the implementation areas. It has been indicated by stakeholders that Cordaid has taken steps/measures to ensure there are no overlaps including working closely with government bodies and health facilities, and other stakeholders during the design and implementation.

3.5.4. Efficiency of Program Management and Accountability Structures

The management and accountability structures of the program appear to be effective and efficient, with good collaboration between the different offices and levels, and there were a clear structures and roles and responsibilities in place. The external collaboration with the regional, zonal and woreda authorities has been strong. There was also good collaboration with federal bodies although it may not be as much as it was desired. The work is governed by a management committee and there is transparency in the process. Health facilities put in place a strong management and accountability structure in place for the PBF project.

All work has been done according to plans and there has been a clear structure in place to ensure accountability. There was clarity in roles and responsibilities of every actor in the project which promoted accountability. Decisions were also being taken based on evidence.

However, some key informants were of the view that transparency and accountability could have been enhanced if the program fund holding/administration and verification responsibilities were separated. Cordaid, on its part, argued that there was separation of fund management and performance verification works in some way stating that the program fund has been administered by the country office while verification works as well as performance procurement were undertaken by its field offices established for each intervention zone. But still the field offices still belong to Cordaid and key informants believe that there could have been more transparency if the fund administration and performance verification roles are played by separate entities.

3.5.5. Financial Management Processes and Procedures

The PBF program's financial management processes and procedures appear to be well and in accordance with the rules of finance, adapting to changing circumstances. It was considered successful in managing and monitoring the budget and spending. Moreover, there were no financial management issues or problems reported that were based on tangible evidences but areas of improvement were indicated around auditing.

Health authorities and facilities indicated that the system has improved over time ensuring timely payments and reporting. The program strengthened roles and responsibilities within the existing system and provided clear financial management procedures that were monitored and regulated every three months, according to many key informants. Several training courses were provided to implementing entities and finance workers to ensure understanding of the system, and effective utilization of resources towards planned activities. The system was also transparent and responsible, and allocated money depending on the business plan. The program has also been able to adapt to changing circumstances, such as currency fluctuations; allowing for extra budget to be reallocated to address COVID-19 and the entire woredas and health facilities zones in Jimma and Borena.

However, stakeholders have indicated a potential for improvement in the area of financial audit of health facilities which needs to be done more frequently and strictly. Improvements are also needed in bringing attitudinal change among health facilities toward the financial audit, according to key informants. Management committee members' knowledge and skill gaps in financial management and auditing principles were also frequently mentioned as bottlenecks to financial management processes at health facilities.

3.5.6. Cost-effectiveness

In this section, the findings of the endline evaluation on cost-effectiveness of the program are presented focusing on cost per health facility and estimated cost per disability adjusted life years (DALY) averted. Methodological issues are addressed in the methodology section of this report.

Overall, the PBF program budget and expenditure data obtained from the Cordaid Country Office shows that the program budgeted Euro 21,978,476, and utilized Euro 17,061,242 until end of 2022, suggesting utilization rate of 78%.

For intervention and control health facilities, financial data on their budget, revenue and expenditure was gathered from health facilities and WorHOs/ZHDs. Whenever possible, the expenditure data was taken as cost of services provided by the facilities. But when expenditure data was unavailable, budget was taken as proxy indicator for cost of service. For intervention

health facilities, PBF program costs obtained from Cordaid Country Office covering the period until end of 2022 were added to determine their total cost of service. PBF program costs included direct payments made to health facilities and administrative costs apportioned/allocated to each intervention facility by taking into account their month of entry.
Table 32 compares the average program costs per intervention health facility by type. The results show that subsidies received by a health center averaged ETB 732,582, which was lower than the subsidies received by a hospital which averaged ETB 1,000,170. But the average subsidies received did not vary statistically significantly by type of facility. When administration costs are apportioned and included, program cost per health center averages ETB 1,217,368, compared to average program cost of ETB 1,491,335 per hospital The total program cost (including admin costs) per facility averaged 1,260,420.

.

Table 32: Program cost per intervention health facility in ETB

	Facility Type		Total
	HC	Hospital	
Mean program cost per facility (PBF direct program cost allocations only)	732,581.64	1,000,169.72 ($P = 0.115$)	777,179.65
Mean administration cost per facility (PBF admin cost only)	596,535.59 ($P = 0.812$)	582,089.88	594,265.55
Mean total PBF cost per facility (PBF program and admin. cost)	1,217,367.49	1,491,335.08 ($P = 0.121$)	1,260,419.54

Table 33 presents program cost per capita by type of intervention facility. The results show that health centers have a substantially higher cost per capita of ETB 48.35 compared to hospitals, which have a cost per capita of ETB 13.44. The total cost per capita for both health centers and hospitals averaged ETB 42.86. Cot per capita differed statistically significantly (P=0.000) between health centers and hospitals.

Table 33: Program cost per capita by type of intervention health facility

	Facility Type		Total
	Health Center	Hospital	
Cost per Capita	48.35 ($P = 0.000$)	13.44	42.86

To measure efficiency and cost-effectiveness of the intervention, we used cost (in ETB) per DALY averted. Table 34 shows the mean cost per DALY averted by intervention status. The results indicate that mean cost per DALY averted is much higher in a control health facility (ETB 938.83 or USD 17.4[20]), compared to the intervention facilities (ETB 267.17 or USD 4.9). The difference was found to be statistically significant (P=0 .044). Average cost per DALY averted per facility in the pastoralist area (ETB 307.04 or USD 5.7) is lower than a facility in agrarian area (ETB 533.02 or USD 9.9), but the difference was not statistically significant (P=0.536). If the World Health Organization's (WHO) cost-effectiveness threshold of USD 100-150 is taken as a benchmark, these figures suggest that the PBF program and regular health service financing system in the intervention and control areas are cost-effective.

[20] Current buying exchange rate of 1 USD = 54.08 ETB was taken, which was sourced from National Bank of Ethiopia's website for April 20, 2023. Retrieved from: https://nbe.gov.et/#1669381958033-b17b5c36-73ed

Table 34: Mean cost per DALY averted across different zones and livelihoods in ETB

	Intervention zone			Intervention Status		Livelihood	
	Jimma	Borena	West Gojjam	Control	Intervention	Agrarian	Pastoralist
Total Cost per DALY Averted	544.85	307.04	59.92	938.83 (P=0 .044)	267.17	533.02	307.04 (P=0.536)

Table 35 compares cost per DALY averted between control and intervention facilities at baseline and endline. The results show that cost per DALY averted increased in the control facilities (although sample size was small), while there was a decline in cost per DALY averted in the intervention facilities from baseline to endline. But the differences were not statistically significant (P=0.729; and P=0.838). It is again important to note here that the cost per DALY averted is higher among intervention facilities compared to control ones at baseline

Table 35*: Comparison of mean cost per DALY averted between control and intervention groups at baseline and endline in ETB

		Time		Percentage Change
		Baseline	Endline	
Total Cost per DALY	Control	266.52	313.53 (P=0.729)	18%
	Intervention	329.44 (P=0 .838)	295.31	-10%

Difference-in-difference analysis was also conducted to evaluate the impact of the PBF program on cost-effectiveness in health facilities, while stratifying by intervention zone (Jimma and Borena) and livelihood (pastoral or agrarian context). The results presented in Table 36 indicate that the cost per DALY averted did not differ statistically significantly (35.62; P=0.872) between baseline and endline for the intervention group compared to the control ones.

Table 36: *Cost per DALY averted **in health** facilities: Difference-in-Differences Analysis in ETB*

	Cost Per DALY Averted				
	Jimma	Borena	Agrarian	Pastoral	Total
Time (Endline=1)	33.84 (0.866)	120.4 (0.780)	33.84 (0.863)	120.4 (0.780)	45.84 (0.784)
Intervention (1=Intervention)	-112.4 (0.578)	-71.15 (0.907)	-115.7 (0.558)	-71.15 (0.907)	-99.39 (0.583)
Interaction (Time*Intervention)	49.20 (0.848)	0 (.)	29.94 (0.904)	0 (.)	35.62 (0.872)
Catchment	0.00059** (0.007)	-0.00204 (0.919)	0.00060** (0.005)	-0.00204 (0.919)	0.00058** (0.003)
_cons	261.0 (0.073)	290.4 (0.751)	260.7 (0.067)	290.4 (0.751)	249.6* (0.047)
N	36	9	38	9	47

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Incremental Cost-Effectiveness Ratio (ICER)

Incremental Cost-Effectiveness Ratio (ICER) is a commonly used measure to assess the efficiency of a health intervention or program. It is calculated by taking the difference in costs between the endline and baseline, and between the control and intervention facilities and dividing it by the difference in their effectiveness (measured in terms of Disability-Adjusted Life Years (DALY) averted.

The ICER provides a measure of the additional cost required to achieve an additional unit of health outcome (compared to an alternative intervention). In other words, it indicates how much it costs to produce a given health benefit. If the ICER is lower than a predetermined threshold, the intervention is considered to be cost-effective. If the ICER is higher than the threshold, the intervention may not be considered cost-effective.

The analysis is done using the total cost and the total DALYs averted by the health facilities. Table 37 below shows the total cost and total DALYs averted disaggregated by intervention status (intervention, control), and zone.

The results also show the total cost of the intervention health facilities higher than the control health facilities in both zones (Jimma and Borena), which is expected given the additional resources required for the intervention. Additionally, the total DALYs averted also found to be higher in the intervention facilities when compared with their control counterparts.

Table 37: Total cost and DALYs averted by intervention status in Jimma and Borena Zones in ETB

	Control		Intervention		Total	
	Jimma	Borena	Jimma	Borena	Control	Intervention
Total Cost	67,166,886.90	31,065,207.00	298,132,385.65	38,001,062.82	98,232,093.90	339,663,718.60
Total DALYs Averted	770,716.61	161,805.26	1,009,479.23	170,141.89	983,430.54	1,239,501.74

$$= \frac{Total\ cost\ of\ intervention\ facilities - Total\ cost\ of\ control\ facilities}{Total\ DALYs\ averted\ by\ the\ intervention\ facilities - Total\ DALYs\ averted\ by\ the\ control\ facilities}$$

We used the above formula to calculate the ICER, and the result of the analysis is presented in Table 38 below. The results indicate that the incremental cost of the PBF program between the intervention and control facilities in Jimma was 230,965,499ETB, whereas in Borena, it was 6,935,856 ETB. Additionally, the incremental benefit or the additional DALYs averted in Jimma was 238,763, and in Borena, it was 8,337.

Table 38: Incremental Cost-Effectiveness Ratio (ICER) of the PBF program in Jimma and Borena Zones, Oromia Region, Ethiopia in ETB

	Jimma	Borena	Total
Incremental cost	230,965,498.75	6,935,855.82	241,431,624.70
Incremental Benefit (DALYs Averted)	238,762.62	8,336.63	256,071.20
ICER	967.34	831.97	942.83

To calculate the ICER, the incremental cost was divided by the incremental benefit. This results indicate an ICER of ETB 967and 832 in Jimma and Borena, respectively. These ICER values indicate that it costs approximately ETB 967 and ETB 832, respectively, to avert one additional DALY in Jimma and Borena. Overall, it costs Birr 943 to avert one additional DALY.

The WHO recommendation on cost-effectiveness states that interventions with an Incremental Cost-Effectiveness Ratio (ICER) less than the Gross Domestic Product (GDP) per capita are considered cost-effective[21]. As the ICER of the project intervention in both zones is significantly lower than the 2021 national GDP per capita of USD 925.1 (at current prices) estimated by the World Bank[22] or 40,524[23] ETB reported by the Ethiopian Statistics Service, the interventions may be considered cost-effective in both Jimma and Borena zones.

Overall, the analysis suggests that cost per DALY averted is not statistically significantly different ((35.62; P=0.872)) between intervention and control facilities, which means that there is no significant difference in the cost-effectiveness of the intervention compared to the control. In other words, the results suggest that the intervention is not more cost-effective than the control. The DiD analysis results also suggest that the cost per DALY averted did not vary statistically significantly between pastoralist and agrarian communities. The findings of the cost-effectiveness analysis can inform future program designs by highlighting the need for reevaluating the intervention approach and strategies to identify potential improvements that can lead to more significant differences in cost-effectiveness; or the need to conduct further research to understand the factors that may have influenced the cost-effectiveness of the PBF program .The results may also suggest the need for exploring alternative interventions that may yield better results, which may include conducting further research and pilot testing of innovative approaches to healthcare financing and service delivery to identify strategies that can effectively improve health outcomes and cost-effectiveness.

3.6. Sustainability

The likelihood of continuation and sustainability of program outcomes and benefits after completion of the program is uncertain. While some respondents believe that the outcomes can continue with the usual support and monitoring, others are more skeptical and believe that the program results will not be sustained without the continued support of the PBF project.

Although in the minority, those that argue the program benefits can be sustained mentioned the system strengthening, improved infrastructure, and capacity strengthening work done by the program as key elements that could ensure sustainability. These are described below.

System strengthening: Several informants expressed the belief that the PBF program had a lasting impact on system strengthening. Another informant mentioned that *"the project mainly strengthened the system of the health facility than providing …. So, I think it will sustain."* Another informant highlighted that *"PBF project didn't give us cash only, it also established a system."* Some key informants argued that the program helped them learn about strict regular monitoring and evaluation that they will continue to employ to sustain the program outcomes. An informant highlighted that *"good work has been done in terms of monitoring and evaluation, so we may sustain the outcome in this way."* Many informants pointed out the importance of support and monitoring

[21] World Health Organization (WHO). (2001). Macroeconomics and health: Investing in health for economic development: Report of the Commission on Macroeconomics and Health. Geneva: World Health Organization. Retrieved from https://apps.who.int/iris/handle/10665/42435

[22]GDP per capita (current US$) – Ethiopia: Accessed from: https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=ET

[23] Determined by multiplying the GDP per capita in current US dollars by the average exchange rate of 2021, which was 1 USD equals 44 ETB.

for the continuation of the program. One informant said, *"The program [benefits] can continue if the usual support and monitoring can be continued by the government and other donors."* Another informant emphasized the need for *"regular supervision and monitoring"* to sustain the program outcomes.

Capacity building: The program has also strengthened the capacity of individuals and health institutions. It enhanced leadership and management capacity and commitment, improved the capacity of the regulatory bodies such as WorHO and ZHDs, and promoted professional ethics. The capacity of clinicians to provide professional health service has also shown improvement as compared to the time before the implementation of the PBF project. The nature of the project by itself encourages the health workers to know more about the service they provide for the clients. This helped health workers to enhance their capacity to provide quality health service. One informant noted, *"…from this project, we gain a lot of knowledge and skill on quality service provision and deliverance of services."* Another informant mentioned *"nowadays, as a health facility, the capability in problem identification, prioritizing and developing action plans is significantly increased."*

Improved management and infrastructure at facilities: The health facilities were able to acquire important infrastructure that they will be able to use to continue providing the service. Few others believe that the program has laid good foundation that improves sustainability of its results. As a result of the program, health facilities currently have better capacity in terms of finance, management, and infrastructure to continue providing quality service. . An informant noted, *"I have no doubt on the sustainability of this project's results because they resolved the health facilities bottlenecks."* Another informant emphasized that *"the health facility has developed a good teamwork and cooperation among different departments."*

However, the overall sentiment of the responses is that the likelihood of the continuation and sustainability of program outcomes and benefits is low. Many informants expressed concerns about the sustainability of the program due to its short duration, lack of institutionalization of the program in the Ethiopian health care system, staff turnover, lack of alternative revenue sources, staff motivation. They believe that the program's impact might not last without continued support and monitoring. These are described below.

Short duration of the program: In some areas, the program implementation had short duration. Health facilities from these areas are particularly concerned that the benefits of the program could not be sustained. An informant stated that *"The duration of the project is too short to talk about the sustainability of the benefits that come from the program."* Another informant argued that *"it is too soon to discuss the likelihood of continuation and sustainability of program outcomes and benefits."* As a result, an informant suggested that *"It is necessary to extend the implementation period further to improve the continuation and sustainability of program outcomes and benefits after completion of the program."*

Lack of institutionalization and government ownership: Several informants emphasized the need for government involvement and ownership to sustain the benefits of the program. An informant stated that *"… sustainability is not realistic at this time because PBF is not institutionalized as a system in the region."* Another informant mentioned that there is limited awareness in the local government and that the issue of ownership needs to be addressed before talking about sustainability. Many informants emphasized the importance of government support and ownership for the program's sustainability. They pointed out the limited awareness and involvement of local governments and the need to strengthen their capacity.

Staff turnover and the need for further training: Informants identified staff turnover as a significant challenge for the program's sustainability. They stressed the importance of continued financial support and capacity building for key stakeholders to ensure the continuation of

benefits. An informant noted a potential hurdle, saying, *"Given the turnover rate of staff in this area, it could be a challenge for us to ensure that we have an adequate number of healthcare workers."*

Lack of alternative revenue sources: Many key informants mentioned that the sustainability of program outcomes and benefits are dependent on the financial support provided by the PBF project. Without this support, they fear that health facilities will struggle to provide services and maintain the gains made during the program. One informant highlighted that *"if PBF project stops the financial support that it was providing until to date, the health service provision will collapse."* Another informant shared similar concerns, stating, *"if there is no another organization that supports the health facility with finance, the capability of the health facilities in doing their job efficiently will become a challenge."* Many informants noted the critical role of financial support and resource allocation in the program's continuation. They expressed concerns about potential challenges in service delivery if the program ceased to provide financial support. *"[name of health facility] health facility has gained many benefits from the PBF program. However, it has not yet reached a level of maturity and self-sufficiency to continue its operations without the support of PBF. While there are some activities that may continue in the absence of the project, medicine and medical equipment purchasing remain significant and challenging issues in the health sector,"* an informant said. Another informant highlighted that *"... if the PBF project stops, it [sustainability] will face challenges. The majority of the interventions were dependent on the finance supplied by the PBF project."* The responses indicate that the health facilities will need on-going support from the government and other donors in order to continue and sustain the quantity and quality of health services that they managed to provide through the support of the PBF program. Without the continued support of the PBF program, most stakeholders are concerned about the shortage of financial and other resources that they particularly need to buy medical equipment and medicine, and decline in staff motivation at health facilities if the program stops.

Staff Motivation and Incentives: Many informants mentioned the importance of staff motivation and incentives in maintaining the program's outcomes. They expressed concerns about the potential loss of staff motivation if the program ceased. A key informant, for example, mentioned, *"I am suspicious about the sustainability of the outcomes because ...the staff already got used to incentives; if this stops the likelihood of continuation will be at risk unless competent alternative is in place."* Another informant said *"The staff motivation may be lost, and resource shortage occur,"* if the program support does not continue.

The findings suggest that the continuation and sustainability of program outcomes and benefits after the program phases out is uncertain, if not low. The sustainability of program outcomes is dependent on commitment of government bodies at all levels and facility leaders, the availability of financial support, and the capacity of the health facilities, among others. Health facilities' limited financial resources, lack of awareness by the local governments and the lack of institutionalization of the program as a system in the region further complicate the issue of sustainability. Without alternative revenue generating mechanisms for health facilities, it is unlikely that most of the program benefits will continue. But some outcomes or benefits such as enhanced capacities of health workers and regulatory bodies, infrastructure built through the program support, and awareness and commitment to regular monitoring, quality of care and professionalism may continue in the absence of the program.

4. ANALYSIS OF OPPORTUNITIES AND CHALLENGES

Performance-based financing (PBF) is a health financing strategy that has gained popularity in recent years, particularly in low- and middle-income countries. PBF involves providing financial incentives to healthcare providers based on their performance in achieving specific health outcomes or service delivery targets. This approach is intended to improve the quality and

efficiency of healthcare services and increase access to care, particularly for underserved populations.

The findings of the endline evaluation provided insights into the potential opportunities and challenges of PBF for the health sector in Ethiopia and other developing countries as described below.

Opportunities

One key accomplishment of PBF has been its ability to improve the quality of care. The findings revealed that the PBF program lead to improvements in the quality of healthcare services, including improved patient satisfaction, and better health outcomes. The PBF program has also been found to increase access to care, particularly for vulnerable populations, by incentivizing providers to deliver services in underserved areas or to target specific populations.

All stakeholders consulted for this endline evaluation had positive views about the Performance-Based Financing (PBF) program. According to them, this program has brought energy and enthusiasm back into the system, empowered people, improved data quality, and increased the quality of healthcare at health facilities. The program's focus on quality and standardized service delivery, efficient and disciplined staff, and its incentive system for better performance are some of its strengths. They also appreciated the program's ability to identify real health system challenges and design appropriate interventions to improve the quality of health care services.

Another advantage of the PBF program is its contribution to promotion of accountability and transparency in the healthcare system. By linking payment to performance, the PBF program was able to promote transparency in resource allocation and ensure that funds are directed towards priority areas. This improved the accountability of healthcare providers by creating clear targets and incentives for performance.

The design of the interventions, specifically the readiness fund, and the management structure of the project can also be considered a strength. The PBF program's implementation process included the use of an integrated approach or system that brought good results to the community, such as giving incentives to health facilities and health workers, enhancing regulatory and service delivery capacity, auditing system, financial flexibility, continuous monitoring and support on services provided, system strengthening, supportive supervision, and ability to identify real challenges. The availability of additional finance allowed health facilities to provide appropriate and quality health services. The program's support for the delivery of quality community-centered healthcare were some other strengths mentioned by the stakeholders.

The PBF program is also directly aligned with the government's policies and priorities. The program’s alignment with government priorities helped it obtain government’s support at all levels. Hence, the engagement of the government bodies and their willingness to support the program were opportunities for the program. Another opportunity for the program was the existence of established health system/facilities, which supports the implementation of the program. The program was able to establish collaboration with government and health facilities, and engage government experts to program implementation. This built capacity while also helping the program to utilize government resources to support program implementation. Another opportunity presented by the program is the potential for further funding for health facilities as the program strengthened transparency and accountability.

Overall, the PBF program has the potential to improve healthcare outcomes and strengthen the health system in Ethiopia. The PBF program has improved accountability and transparency in the health sector, increased motivation and performance of health workers, enhancing data quality, increased access to health services for the population and particularly for the disadvantaged or

vulnerable communities, improved quality of healthcare services, and enhanced the potential for health facilities' to mobilize funding in the future. There is a growing commitment from the government to institutionalize PBF in the existing health system of the country. Given its strong results, there is also a potential for increased investment in the health sector by donors and other stakeholders when the results and lessons are widely disseminated.

Challenges and threats

However, PBF also presents some challenges that need to be addressed in order to ensure its effectiveness. One challenge is the potential for unintended consequences, such as the focus on achieving certain targets at the expense of other aspects of care, or the potential for providers to manipulate data in order to meet targets. There is also a risk that PBF may lead to fragmentation of the healthcare system, with providers focusing on specific services or populations in order to maximize their incentives.

The PBF program also faces several other challenges that need to be addressed to ensure its long-term sustainability and success. PBF program has remained dependent on external funding. Government has limited capacity to manage and monitor a similar PBF program, and has insufficient funding to scale up the program to cover the entire country. There is limited progress in the institutionalization of PBF within the Ethiopian health system. Continuity of program results will be at risk unless the program continues, or the government institutionalizes PBF in its existing health system. This could lead to reduced staff motivation, drug shortages, availability of quality health services and. *"If the program is interrupted, it will not be possible to purchase medical equipment and medicines from the government's regular budget,"* a key informant said. There is still no guarantee that the government will institutionalize PBF in its health system, which could jeopardise the sustainability of results and scaling up objectives.

There have been various external factors such as drought, food insecurity, drug shortages, inflation, and shortage of water that pose a threat to the program's implementation as well as continuity. Political instability and conflicts, inflation and economic downturn could also disrupt the program and kill the momentum gained so far. Prolonged drought and inflation of medical equipment could lead to sustainability issues and a shortage of drugs. Instability in some intervention areas have also made, and may continue making access to some of facilities difficult. Staff turnover particularly at health facilities and lengthy government procurement processes have also t challenged the program implementation and its effectiveness, although the program has contributed to reduction in turnover, according to key informants.

In summary, the PBF program has faced several threats and challenges. Stakeholders have emphasized the importance of the government's commitment to facilitating the program's continuation to prevent the potential risks to sustainability and effectiveness. Additionally, addressing operational and procurement challenges and ensuring adequate staffing at the health facility level could also help overcome some of the program's challenges.

In conclusion, PBF presents both opportunities and challenges for the health sector. Evaluation findings suggest that it has the potential to improve the quality and efficiency of healthcare services and increase access to care, particularly for underserved populations. However, careful attention must be paid to the design and implementation of PBF programs in order to contextualize and ensure their effectiveness and avoid unintended consequences.

5. KEY FINDINGS, INCLUDING LESSONS LEARNED

5.1. Key Findings

The key findings of the end of program evaluation and lessons learned are highlighted below.

The PBF program has shown promising results in improving the quality of healthcare services. Although the program improved quality of health services, no evidence was found that the program had statistically significant positive impact on health service utilization compared the control group. But it is important to note here that utilization measured by DALYs averted increased in both intervention and control health facilities compared to baseline. Qualitative findings also suggest that the PBF program has contributed to an increase in the number of women who received postnatal care, had deliveries attended by skilled birth attendants, number of outpatient departments (OPDs) visits by adults. The increased availability of quality health services also helped the poor and vulnerable people to get treatment that otherwise could not have been possible. In other words, the program improved access to quality health service which particularly benefited the poor and disadvantaged segments of the population who do not have financial capacity to seek health services from private facilities or travel to other locations to get health services.

The end of program evaluation also found statistically significant differences in overall patient satisfaction between intervention and control health facilities. It appears that the program has improved patients' overall satisfaction.

The findings of the end of program evaluation also suggest that the PBF program improved health worker motivation and job satisfaction with their compensation and benefits as well as work load. By linking pay to performance, the PBF program was able to create a sense of recognition and reward for health workers. This, in turn, improved job satisfaction and motivation as well as quality of service delivery. The results also suggest that financial incentives are powerful tools to improve health worker motivation and satisfaction.

The PBF program has also produced a number of outputs, including the introduction of a regular and rigorous system for measuring and monitoring healthcare performance, and the introduction of performance-based contracts that involved a payment system that rewards health providers for achieving specific targets. These outputs have helped to deliver quality healthcare services, and improve access to health services and essential medicines.

The evaluation did not find statistically significant difference between PBF and non-PBF facilities at endline compared to baseline in terms of efficiency (cost per DALY in ETB).

The partnership strategy of the PBF program mainly involved collaboration among the project implementer (Cordaid) the government at all levels, donor and healthcare providers. The government provided leadership and oversight of the program and strengthened its regulatory activities, while the Embassy of the Netherlands provided financial support. Health facilities have been responsible for delivering healthcare services and achieving specific outcomes, in exchange for financial rewards. Cordaid managed and implemented the program working closely with health facilities and relevant government bodies. It was also responsible for signing performance contracts with health facilities and providing subsidies after doing verifications of reported data. This partnership strategy has helped to ensure that the program is aligned with national health policies and priorities, and priorities and needs of health facilities, while also leveraging the expertise and resources of government and health providers. But some key informants thought that the program fund management and verification tasks may need to be performed by different organizations to deal with any conflict of interest.

Overall, the PBF program has shown promising results in improving quality of health service delivery and increasing access to quality care, particularly to disadvantaged people. The program's focus on incentivizing healthcare providers to deliver quality care, its strong regular and monitoring support aided by well-designed checklists, implementation of activities based on plans, capacity building of health care providers and regulators, its partnership strategy and others have all contributed to its success.

5.2. Unintended outcomes

PBF programs are often criticized for their potential to create perverse incentives if the performance indicators are not well-designed, or if the rewards are too narrowly focused. One common concern is that providers may prioritize incentivized services over other important health services, which can lead to the neglect of non-incentivized services, and ultimately harm patient health outcomes.

Some key informants raised concerns that the PBF program implemented in the three zones in Ethiopia created perverse incentives, such as providers prioritizing incentivized services over other important health services. An informant, for example, emphasized, *"There are indicators that were not included in this program. Thus, there is poor in performance. This need attention."* This may suggest that providers may be neglecting non-incentivized services.

Many key informants also mentioned the need to add indicators and expand its coverage of health services. For example, an informant noted, *"The indicators should be added and cover more areas of health care delivery."* Another informant stated, *"I think if indicators were added…, more quality work would be done."* The need to add more indicators was mainly mentioned as related to health posts. The fact that stakeholders suggested the need to add more indicators into the program may not, however, necessarily mean that they are neglecting non-incentivized services.

The findings seem to also suggest that health facilities have increasingly become dependent on the program to maintain their improved service delivery. Most informants mentioned that the program benefits would not be sustained and they would more likely go back to their baseline situation if the support is interrupted or the support stops. While sustainability of program benefits remain uncertain, the reliance of health facilities on external funding support to maintain their improved services could be a worrying sign.

Although not conclusive, the program might have created perverse incentives, such as providers prioritizing incentivized services while neglecting non-incentivized services. But it is also important to note here that some of the support provided by the program were also meant to strengthen their systems which could benefit all programs regardless of whether they are incentivized directly or not.

5.3. Lessons Learned

The implementation of a PBF program taught stakeholders several important lessons, which are highlighted below.

Context matters: The success of PBF programs largely depends on the local context, including the level of resources, and the capacity of the health system. Many informants felt that the quality of care assessment checklist was very long and questioned applicability of some of the assessment criteria to local realities. An informant noted *"… they ask you to do what is not applicable, you cannot demolish and construct what was already constructed at once."* Hence, it is essential to adapt the program design to the local context and address the existing challenges.

Clear and measurable indicators and assessment checklists: For PBF programs to be effective, it is crucial to define clear and measurable performance indicators. Besides, assessment checklists used to verify or assess performance levels should be customized to local context. An informant, for example, said, *"…quality assessment should be objective, not subjective. Verification officers are subjective in their evaluation."* The informant, as a result, suggested the need *"standardize quality [assessment] checklist that can avoid bias and subjectivity based on our context."*

Strengthening Health Management Information Systems (HMIS): The PBF program has largely relied a lot on performance indicators that are based on health service utilization data. Cognizant of this, the program invested in strengthening the existing HMIS, which was the right decision. However, data quality has remained an issue in the existing HMIS. A lot of inconsistencies were observed between the data available at health facilities and the data entered into the HMIS, which might be due to non-reporting, and other reasons. It is true that the program worked a lot to improve data quality particularly at health facility level. But the fact that data quality is improved at health facility level might not necessarily mean that that all that data is entered into the system. Another important lesson learned from this program implementation of the PBF program is the need to strengthen HMIS at all levels to ensure that comparable data is available at intervention and control health facilities for measuring impact of the program. For example, although this evaluation utilized declared data for assessing impact and cost effectiveness, the quality of data declared by the intervention and control health facilities is believed to be significantly different. Intervention facilities face penalties (non-release of subsidy) if their declared data is significantly different from the one obtained through verification. Hence, they have the incentive to ensure that their declared data is close to the reality. On the other hand, non-intervention health facilities do not have such incentives. The error margin analysis presented in this report showed that data quality was poor at baseline and declined after they start receiving support from the project, suggesting that the declared data gathered from non-intervention health facilities might still involve that high error margin. Hence, lessons learned emphasize the need to invest more in strengthening HMIS at all levels. Improved data quality also encouraged use of the data by stakeholders in making decisions.

Aligning program indicators with existing HMIS indicators: One of the important lessons learned from the program implementation was the need to harmonize PBF indicators with existing HMIS indicators to make it easier for healthcare providers to collect and report data, reducing the burden on health facilities. Using the same indicators across PBF and HMIS also ensured consistency and comparability of data, which is vital for monitoring progress, evaluating the effectiveness and impact of the program, and making evidence-based decisions at different levels of the health system. Moreover, when PBF indicators are integrated into existing HMIS, it enhances the sustainability of the PBF program by embedding it within the existing health system.

Regular and robust monitoring and evaluation system: PBF program links financial incentives to the achievement of predefined performance indicators, aiming to improve health service delivery and health outcomes. This calls for a strong and robust M&E system to ensure that financial incentives are allocated based on actual performance, ensuring transparency and accountability in the distribution of resources. Most informants were also very appreciative of the M&E system in place for the program. The importance of having a strong and regular M&E system was identified as major lesson learned by informants at all levels. *"PBF promote accountability and transparency through strong monitoring and evaluation,"* an informant said. Regular M&E, according to another informant, *"have made health institutions work harder."* Comparing the time before the project implementation and after, another informant also noted *"Before stating of this project, work done went unrecorded. The registration and the report would not match. When they come for follow-up, they look at plans, look at reports and actually evaluate what has been done. Their monitoring and evaluation…has brought quality data and timely reporting."* The well-organized monitoring system put in place by the program supported by checklists also helped health facilities improve their standard quality of care, according to many key informants. An

informant, for example, noted *"The tool designed for PBF monitoring and evaluation purpose is very smart and helpful to bring the health facilities to a better standard level and to improve the health facilities' gaps or bottlenecks."*

Data audits and verification: To maintain the quality of health service utilization data, regular data audits and verification exercises conducted by the program were found to be important. This helped identify discrepancies, detect potential manipulation of data, and improve data quality. As some informants also suggested the need to give data verification and funding holding tasks to different bodies, this may call for the need to examine the matter further and/or implement a third-party verification system to contribute to the credibility and reliability of the data. The fact that the program incorporated data quality issues measured by error margins to decide to release (or not to release) subsidy associated with each indicator was also an important lesson to encourage health facilities work harder to improve data quality.

Engaging stakeholders: The need to actively engage all relevant stakeholders, including healthcare providers, local communities, and government bodies at all levels, was another important lesson learned from the implementation of the program to ensure ownership, achieve objectives, and improve chances of sustainability. Their involvement in the design and implementation of the program helps ensure that the program is tailored to the local context, reduce duplication of effort, and so on. The need to involve central government bodies throughout the program cycle is particularly critical to institutionalize PBF in the exiting government health care financing system, and sustain its results.

Capacity building and technical assistance: Implementing the PBF program requires a strong capacity for management, monitoring, and evaluation and others. Provision of technical assistance and capacity-building support to local health systems was essential for the successful implementation of PBF program.

Ensuring sustainability: The long-term success of the PBF program depends on its sustainability. The evaluation findings suggest that sustainability of the program results remain uncertain. Future PBF programs need to address sustainability issues starting from the very design of the program and all other phases of the program implementation cycle. Among others, strategies to ensure sustainability include integrating the program into the national health system by engaging central and lower level government bodies, securing consistent funding until PBF to pilot the program at large scale and demonstrate its values, and fostering ownership at all levels.

Program management and components: Key informants appreciated the program's management and its components, emphasizing the importance of the readiness fund, continuous support, and transparent system. They believe these aspects should be replicated in future programs. *"…the components of the program intervention, like that of the readiness fund, continuous support, and transparent system, are very good and need to be replicated in the future,"* an informant stated. Respondents highlighted the importance of staff involvement and motivation, noting the positive impact of incentivizing good work. They suggested that the PBF program can serve as a model for other health facility programs. An informant stated *"PBF is a role model program for other program in our health facility."* Another informant said *"If you incentivize someone for what he has done, he will be motivated to work hard to yield more good result."* Health facilities also appreciated the autonomy they enjoyed in utilizing the subsidy. They reported that the autonomy helped them to allocate the fund to address their critical gaps. *"From this project we learned that financially supporting the health facility and letting them manage it autonomously can improve the quality of health care."* Informants also emphasized the importance of a conducive working environment and strong management committees for better service delivery. The establishment of management committee at health facility level with member drawn from different departments and with different professions is said to have promoted better decision making and resource utilization, and promoted transparency and accountability.

In conclusion, the main lessons learned from the implementation of the PBF program include effective program management and components, staff involvement and motivation, contextualization of intervention, stakeholder engagement, clear indicators, audit and verification of data, strong M&E system and HMIS, capacity building, and sustainability are all critical factors for the effectiveness of the program.

6. CONCLUSIONS AND RECOMMENDATIONS

6.1. Conclusions

On the basis of the findings presented above, the following conclusions have been drawn.

i. The program has made significant contributions to the improvement of quality of healthcare services. By aligning its efforts with the Ethiopian Health Sector Transformation Plan II (HSTP II), the program has enhanced service delivery, health workforce development, health system financing, health information system strengthening, and leadership, governance, and accountability in the health sector. The program has been addressing the needs of patients, health workers, and health authorities, the program has contributed to each of the five pillars of HSTP II. The program's inputs and strategies were generally considered realistic and adequate to achieve its targets/objectives. The program has also been largely well-received by health facilities and health workers.

ii. The PBF program is relevant and coherent with global and regional instruments, declarations and development agenda, the national government's policies and priorities, needs of health facilities and their clients, and portfolio of Cordaid. The PBF program is complementary to the CBHI scheme, providing financial support and resources to improve health service delivery and accessibility for CBHI beneficiaries. But there is a lack of coordination and interaction among organizations implementing health interventions.

iii. The PBF program has contributed to efforts to address gender disparities in access to and utilization of maternal and child health services, and improved quality of care. The program's focus on maternal health has specifically helped vulnerable groups, including mothers, to access and utilize healthcare services. The program's efforts to address shortages of essential medicines and laboratory equipment have had a positive impact on vulnerable communities, particularly women seeking healthcare services. It also contributed to improving the capacity of health care workers and facilities to provide quality health care services. However, challenges related to distance, transportation, and the quality of care remain, highlighting the need for continued efforts to promote gender-sensitive healthcare delivery.

iv. The impact of PBF program has been mixed. The program appears to have substantial positive impacts on most dimensions of quality of care. The percent increase in quality index between baseline and endline in intervention facilities is significantly higher than control facilities. The quality of care impacts extend beyond infrastructure and equipment and include instances of improvements in the quality of service delivery, such as doing laboratory tests prior to prescribing medicine, and better customer service.

v. However, the evaluation did not find a statistically significant difference in the total DALYs averted between the control and intervention health facilities. Hence, although the PBF program improved healthcare service quality and significantly increased utilization, in

DALYs averted, of ANC1, it did not have a significant impact on overall service utilizationas well as utilization of other key selected services such as family planning long-term and short-term, ANC4, PNC1, skilled deliveries, OPD Under-5. The evaluation found that the intervention had a significant positive impact on ANC1 service utilization in all intervention zones. Review of relevant literature also shows mixed results about the impact of PBF on service utilization. While some studies have reported a positive impact of PBF on health service utilization, others did not. For example, a study by Meessen et al. (2011)[24] that evaluated the impact of a PBF program on healthcare service utilization in the Democratic Republic of Congo found that the PBF program did not have a statistically significant impact on service utilization. Similarly, a study by Paul et al. (2014) evaluated the impact of a PBF program on maternal and child health services in Nigeria found that the PBF program did not have a statistically significant impact on service utilization[25]. This finding may suggest that the PBF program did not address the underlying factors that influence health service utilization, such as access to health facilities, quality of care, cultural and social factors, or cost. It may also indicate that the program design or implementation was inadequate in targeting the specific needs and challenges faced by the population.

vi. No statistically significant differences were found in mean cost per DALY averted between control and intervention facilities, indicating that the PBF program did not result in a significant improvement in the efficiency of health services. Similarly, no statistically significant differences was also found in cost per DALY averted between pastoral and agrarian areas.

vii. The evaluation also found significant disparities in cost per capita between health centers and hospitals. Health centers have a higher cost per capita compared to hospitals, and the difference was statistically significant.

viii. The PBF program has led to improvements in data quality, although overall data quality remains low. The findings demonstrate a positive impact of the PBF program on health information system, data-based decision making, and transparency. The program has changed health workers' attitudes, improved their knowledge and skills in data management, and contributed to better registration and documentation of health-related data.

ix. The PBF program had positive impact on motivation, compensation and benefits, and overall job satisfaction of health workers. Health workers in intervention facilities reported higher satisfaction with their compensation and benefits, work load, and higher overall job satisfaction compared to health workers in control health facilities. But the evaluation did not find statistically significant differences between health workers in intervention and control facilities in most other domains of job satisfaction such as nature of job and responsibilities, organizational practices and functioning, working environment, career development and job security, performance, recognition, overall well-being. . It was also found that health workers in intervention facilities reported statistically significantly higher overall job satisfaction than those in control facilities, despite not having statistically significant differences in most other domains of job satisfaction except in the area of compensation and benefits. This result may also suggest that the PBF program addressed the underlying issues that affect job satisfaction, which is compensation and benefits,

[24] Meessen, B., Soucat, A., Sekabaraga, C. (2011). Performance-based financing: just a donor fad or a catalyst towards comprehensive health-care reform? Bulletin of the World Health Organization, 89(2), 153–156. doi: 10.2471/BLT.10.077339
[25] Paul, E., Albert, L., Bisala, B. N., et al. (2014). Performance-based financing in low-income and middle-income countries: Isn't it time for a rethink? BMJ Global Health, 3(1), e000664. doi: 10.1136/bmjgh-2017-000664

x. The findings of the end of program evaluation found that the program led to an increase in patient satisfaction with healthcare services. The subsidies and technical support provided by the program allowed health facilities to make important health services and medicine available to their communities. Facilities were also able to improve their infrastructure and make their facilities attractive and comfortable for patients. This is consistent with the findings of several studies that found positive impact of PBF programs on patient satisfaction. For example, a study conducted in Burundi found that PBF programs led to an increase in patient satisfaction with healthcare services (Bertone et al., 2018)[26]. Similarly, a study conducted in Cameroon found that PBF programs improved patient satisfaction with healthcare services in both urban and rural areas (Witter et al., 2013)[27].

xi. PBF program appears to have increased patient waiting time. The findings suggest that the PBF program has increased patient flow and service utilization (though not significantly compared to control facilities). However, intervention health facilities do not appear to increase their workforce to keep up with the increased demand for health services.

xii. The program has also been effective in improving regulatory capacity of WorHOs and ZHDs to provide regular and supportive supervision and constructive feedback. .

xiii. The findings suggest that the continuation and sustainability of program outcomes and benefits after end the program is uncertain. It is clear that the program requires strong support from the government and other donors in order to maintain the progress and capitalize on it. Furthermore, the lack of institutionalization of the program in the existing government health system further complicate the issue of sustainability. Without alternative revenue generating mechanisms for health facilities, it is unlikely that the program benefits will be sustained. Government's commitment to scale up the program and replicate it across the country also remains uncertain. Even if the government decides to scale up the program, it may not have strong technical capacity to provide regular monitoring and supportive supervision, provide incentives based on verified performance, and do other components of PBF adequately.

6.2. Recommendations

Based on the conclusions drawn above, the following general and stakeholder-specific recommendations are made.

Government

i. The PBF program achieved mixed results. While it increases quality of care, it faces challenges related to increasing volume of service utilization, sustainability, efficiency, and institutionalization. By focusing on these areas and building on the positive impacts of the program, it is possible to continue to improve the quality of care in Ethiopia. Additionally, further piloting and evaluation with more complete data is important before scaling up the PBF program in Ethiopia. Moreover, it's important to use public fund to pay for performances and to integrate PBF into the regular provider payment mechanism (either via MoH/RHB) as well as the routine M&E system to enhance sustainability and efficiency.

[2626] Basinga, P., Gertler, P. J., Binagwaho, A., Soucat, A. L., Sturdy, J., & Vermeersch, C. M. J. (2011). Paying primary health care centers for performance in Rwanda. Washington, DC: The World Bank.

[27] Witter, S., Fretheim, A., & Kessy, F. L. (2013). Performance-based financing in low-income and middle-income countries: Isn't it time for a rethink? BMJ Global Health, 3(1), e000664.

ii. The evaluation demonstrates that the PBF program can be successfully implemented to increase quality of health care in public health facilities that often have capacity constraints. But the program was managed by a non-governmental organization that has strong capacity and program management experience. There is a need for more pilot in a context where the public sector fully takes over and manages implementation of such a program from end to end and without significant outside technical support.

iii. The PBF program had positive impacts on health providers in terms of higher motivation, and satisfaction with their compensation and benefits and motivation. Therefore, it is important to build on these positive impacts by continuing to provide incentives based on verified performance and improving overall job satisfaction.

iv. Strengthen health management information system to ensure that quality data is available for measuring effectiveness, impact and efficiency of PBF program, and support evidence-based decision making.

v. Strengthen the relationship between the PBF program and CBHI scheme to further enhance the quality and accessibility of care for CBHI beneficiaries.

Donors/Development Partners

vi. Support the government in piloting the PBF program at a relatively larger scale and enhancing efficiency as well as its impact on volume of health service utilization.

vii. Support government in strengthening health management information system.

viii. Provide technical assistance and share best practices from other PBF programs implemented in similar contexts, with a focus on improving volume of service utilization, efficiency and sustainability of such programs.

ix. Encourage and fund additional research on PBF programs, focusing on the factors that influence utilization, efficiency and sustainability to inform future policy decisions and program design.

x. Conduct research to explore whether separating the fund holder and the performance verification activities would enhance transparency and accountability or not.

xi. Extend the program's duration particularly in areas with shorter implementation periods, and support further pilot of the program.

Cordaid/ Project Implementers

xii. Conduct a comprehensive review of the program design and implementation to identify the potential reasons for the lack of impact on utilization of all health services except ANC1. The review may involve an assessment of the program components, adequacy of financial incentives, the mechanisms for disbursing funds, the accessibility of health services, and the cultural and social factors that influence health-seeking behavior. Based on the findings of the review, the PBF program may need to be adapted to better align with the specific needs and challenges faced by the target population. Moreover, it may be necessary to consider complementary interventions to address the underlying factors that influence health service utilization, such as improving infrastructure and transportation, strengthen provision of health education and awareness-raising activities, and addressing gender and social inequalities. As data quality issues may also be a factor behind this, it

may be useful to address missing values and other data quality issues surrounding the data used for this evaluation, and re-evaluate the impact of the program on health service utilization.

xiii. As the program did not show a significant difference in cost-effectiveness, there is a need to reconsider the design and implementation of the PBF program to identify potential areas for improvement; or consider alternative approaches that may be more cost-effective or conduct further research to understand the factors that may have influenced the cost-effectiveness of the PBF program

xiv. Support government efforts in enhancing the utilization, efficiency and sustainability of the PBF program by providing technical assistance and sharing best practices.

xv. Share the PBF program evaluation findings and lessons learned with government bodies at all levels and other stakeholders.

xvi. Ensure that future programming gives due attention to involving government bodies at all levels to promote ownership the program that is needed to pave the way for scaling up and sustainability.

xvii. Carefully handover the program to the government by designing and implementing exit/sustainability strategies. Among others, the sustainability plan may consider the need to strengthen committee members through training, experience sharing, involvement of the local, regional and federal government bodies, create alternative revenue-generating mechanisms for health facilities to maintain program benefits, provision of training to new health workers and refresher training to existing workers until health facilities and particularly new ones are able improve their institutional capacity.

xviii. Encourage and facilitate the sharing of best practices and lessons learned among health facilities and other stakeholders to promote continuous improvement in data quality and transparency.

xix. Consider conducting a comprehensive economic evaluation of the program involving larger number of health facilities in the sample to understand the cost-effectiveness of the program and determine areas where improvements can be made to optimize the program's impact.

Health Facilities

xx. The PBF program had positive impacts on health providers in terms of higher motivation, and satisfaction with their compensation and benefits and motivation. Therefore, it is important to build on these positive impacts by continuing to provide incentives based on verified performance and improving overall job satisfaction.

xxi. Given the positive impact of the PBF program on patient satisfaction with healthcare services, it is important to continue and scale up the program with a focus on enhancing patient satisfaction by making health facilities attractive and comfortable for patients, improving the attitude of health providers, and reducing patient waiting time.

xxii. Share best practices and lessons learned from successful health facilities to improve the efficiency and cost-effectiveness of the PBF program across different contexts.

xxiii. Develop alternative revenue-generating mechanisms to ensure financial sustainability and enhance the sustainability of results made possible through the PBF program support.

General recommendations

xxiv. Future PBF programs in Ethiopia need to give due attention to health worker compensation and benefits to enhance their job satisfaction, while doing further analysis to identify the other domains that affect job satisfaction among health workers in the country.

xxv. Continue monitoring the PBF program and its impact on utilization and cost-effectiveness in the Ethiopian health sector, with a focus on identifying opportunities for further optimization.

xxvi. Given the positive impact of the PBF program on quality service delivery and patient satisfaction with healthcare services, it may be worthwhile to continue piloting the program on a relatively larger scale and in diverse contexts while monitoring and evaluating its cost-effectiveness, impact on the volume of health service utilization, and sustainability.

xxvii. The PBF program had positive impacts on health providers in terms of higher motivation, and satisfaction with their compensation and benefits, and overall satisfaction with their job. Therefore, it is important to build on these positive impacts by continuing to provide incentives based on verified performance and improving overall job satisfaction.

xxviii. Future PBF programs need to balance demand and supply side by ensuring staffing levels at health facilities grow with the growth of demand for health service utilization caused by the program.

Appendix: 1. Demographic characteristics of health workers that participated in the survey

Health Facility and Health Worker Background information		Total		Jimma		Borena		West Gojjam	
		Intervention	Control	Intervention	Control	Intervention	Control	Intervention	Control
		x̄ /%	x̄ /%	x̄ /%	x̄ /%	x̄ /%	x̄ /%	x̄ /%	x̄ /%
Type of Facility	Primary Hospital	14	12	11	0	14	25	33	50
	General Hospital	7	4	11	6	0	0	0	0
	Health Centre	79	84	78	94	86	75	67	50
	Other	0	0	0	0	0	0	0	0
Sex of the health worker	Male	82	84	81	87	86	83	78	67
	Female	18	16	19	13	14	17	22	33
Age (in years)		32.19	37.73 *	32.88	38.25 *	30.61	39.24 *	31.7	32.06
Marital status	Single	42*	24	48	32	29	8	33	17
	Married/living together	57*	76	50	68	71	92	67	83
	Widowed	1	0	2	0	0	0	0	0
Who are you employed by?	Federal government	6	8	2	3	19	17	0	17
	Regional Health Bureau	55	59	54	58	71	75	22	33
	Zonal Health Office	36	27	44	39	10	0	44	17
	Other (specify)	4	6	0	0	0	8	33	33
Type of employment	Permanent and pensionable	100	96	100	94	100	100	100	100
	Fixed term	0	4	0	6	0	0	0	0
Level of education	Bsc	58	55	56	55	67	67	56	33
	Medical Doctor	6	4	6	0	0	8	22	17
	Msc	1	2	2	0	0	8	0	0
	Diploma	33	39	35	45	33	17	22	50
	Others specify	1	0	2	0	0	0	0	0
Highest number of education years completed?		8.36	8.53	4.37	4.39	15.29	16	16.1	15
Years worked as a health worker at this facility?		2.54	4.53*	2.54	3.55	2.33	7.67*	3	3.33
Months worked as a health worker at this facility?		5.02	4.35	5.11	3.97	3.52	4.67	8	5.67
Position of the health worker as designated by federal ministry of health	CEO	2	2	0	0	5	8	11	0
	Medical Officer	7	2	11	0	0	8	0	0
	Health Officer (HO)	21	18	22	19	19	25	22	0
	Nurse (B.Sc.)	15	8	19	6	10	17	11	0
	Midwives	18	24	15	23	29	17	11	50
	Nurse (Diploma)	20	22	26	26	5	8	22	33
	BSc anaesthetist/Nurse anaesthetist	0	2	0	3	0	0	0	0
	Radiographer	0	2	0	3	0	0	0	0
	Physiotherapist	0	2	0	3	0	0	0	0

		Laboratory technologist	1	0	2	0	0	0	0	0
		Laboratory technician	1	4	0	6	5	0	0	0
		Pharmacist	1	0	0	0	5	0	0	0
		Pharmacy technician (Druggist)	5	4	4	6	10	0	0	0
		Other	7	8	2	3	14	17	22	17
Which of the following services have you provided within the past 3 months?		Supervise CHW / Village Health Worker (VHW)	18	18	9	23	38	17	22	0
		Supervise TBA (Traditional Birth Attendant)	15	10	17	16	10	0	22	0
		Consultation for children	50	35	39	29	71	50	67	33
		Consultation for adults	52	41	46	42	67	42	56	33
		Family planning	40	57	35	52	62	58	22	83*
		ANC (Antenatal care)	43	57	37	52	62	58	33	83
		PNC (Postnatal Care)	42	47	35	45	62	42	33	67
		Deliveries in facility	46	49	41	45	67	42	33	83
		Tuberculosis (TB) treatment/diagnosis	31	29	22	19	48	50	44	33
		Vaccinations	36	43	28	35	48	50	56	67
		Malaria treatment	35	43	22	32	57	58	56	67
		Nutrition (feeding, counseling, growth monitoring, management of malnutrition)	39	43	28	26	62	58	56	100
		CHW training	11	8	9	10	19	8	0	0
		HIV/AIDS treatment and care	25	27	11	13	52	42	44	67
		Other (Specify)	12	24	13	29	10	17	11	17
When was the most recent time you were trained in the following after undergoi	IMCI (Intergrated Management of Childhood Illnesses)	Less than 1 yr	13*	2	15	0	14	8	0	0
		> 1 yr ago	21	27	28	29	10	33	11	0
		Never	65	71	57	71	76	58	89	100
	Malaria	Less than 1 yr	2	6	0	6	10	8	0	0
		> 1 yr ago	27	16	31	13	24	33	11	0
		Never	70	78	69	81	67	58	89	100
		Less than 1 yr	8	4	11	0	5	17	0	0

ng professional education?	TB diagnosis and treatment	> 1 yr ago	21	18	26	13	14	33	11	17
		Never	70	78	63	87*	81	50	89	83
	FP (Family Planning) methods	Less than 1 yr	4	10	4	16*	0	0	11	0
		> 1 yr ago	31	33	39	29	19	42	11	33
		Never	65	57	57	55	81	58	78	67
	Labor and delivery	Less than 1 yr	10	0	13	0	5	0	0	0
		> 1 yr ago	20	20	20	19	24	25	11	17
		Never	70	80	67	81	71	75	89	83
	Mental Health	Less than 1 yr	1	0	2	0	0	0	0	0
		> 1 yr ago	19*	2	24	0	14	8	0	0
		Never	80	98*	74	100	86	92	100	100
	Management training	Less than 1 yr	10	2	11	0	10	8	0	0
		> 1 yr ago	25*	8	30	13	19	0	11	0
		Never	65	90*	59	87*	71	92	89	100
	Community Health	Less than 1 yr	5	4	2	3	14	8	0	0
		> 1 yr ago	21*	6	31*	6	5	8	0	0
		Never	74	90*	67	90*	81	83	100	100
	Pre/post-natal care	Less than 1 yr	4	0	6	0	0	0	0	0
		> 1 yr ago	25	14	28	13	19	25	22	0
		Never	71	86	67	87*	81	75	78	100
	HIV/AIDS care and management	Less than 1 yr	7	0	6	0	10	0	11	0
		> 1 yr ago	21	16	28	13	5	8	22	50
		Never	71	84	67	87*	86	92	67	50
	Hypertension	Less than 1 yr	4	0	6	0	0	0	0	0
		> 1 yr ago	19*	6	24*	6	14	8	0	0
		Never	77	94*	70	94*	86	92	100	100
	Diabetes	Less than 1 yr	5	0	7	0	0	0	0	0
		> 1 yr ago	20*	4	26*	6	14	0	0	0
		Never	75	96*	67	94*	86	100	100	100
	EmOnc (Emergency Obstetric and Neonatal Care)	Less than 1 yr	6	0	9	0	0	0	0	0
		> 1 yr ago	25	12	30	13	19	17	11	0
		Never	69	88*	61	87*	81	83	89	100
	HBB (Help Baby Breath)	Less than 1 yr	2	0	4	0	0	0	0	0
		> 1 yr ago	21	16	30	16	10	8	0	33
		Never	76	84	67	84	90	92	100	67
	LSS (….)	Less than 1 yr	0	2	0	3	0	0	0	0
		> 1 yr ago	15	0	24	0	0	0	0	0
		Never	85	98*	76	97*	100	100	100	100
	EPI (Expanded Programme of Immunization)	Less than 1 yr	6	2	6	0	10	8	0	0
		> 1 yr ago	24	22	33	29	10	17	0	0
		Never	70	76	61	71	81	75	100	100
	Infection Control	Less than 1 yr	8	10	9	6	10	25	0	0
		> 1 yr ago	30	18	37	23	19	17	11	0
		Never	62	71	54	71	71	58	89	100

*p<0.05

Appendix: 2. Declared and verified data, and data reliability for all indicators in all intervention health facilities by zone and quarter

Intervention Zone	Phase of Entry into the PBF program	Reporting Quarter in Gregorian Calendar	Type of Data		data reliability
			Declared	Verified	
Borena	Phase I	2015_Q2	4625	2021	44%
Borena	Phase I	2015_Q3	2776	219	8%
Borena	Phase I	2015_Q4	5513	383	7%
Borena	Phase I	2016_Q2	3998	1498	37%
Borena	Phase I	2016_Q3	5333	5256	99%
Borena	Phase I	2016_Q4	1997	2820	141%
Borena	Phase I	2017_Q1	1778	959	54%
Borena	Phase I	2017_Q2	1281	1078	84%
Borena	Phase I	2017_Q3	1816	870	48%
Borena	Phase I	2017_Q4	1684	1750	104%
Borena	Phase I	2018_Q1	3705	3524	95%
Borena	Phase I	2018_Q2	3282	3577	109%
Borena	Phase I	2018_Q3	2280	1979	87%
Borena	Phase I	2018_Q4	835	790	95%
Borena	Phase I	2019_Q1	1004	880	88%
Borena	Phase I	2019_Q2	740	272	37%
Borena	Phase I	2019_Q3	1430	455	32%
Borena	Phase I	2019_Q4	1114	501	45%
Borena	Phase I	2020_Q1	1578	1150	73%
Borena	Phase I	2020_Q2	1737	1258	72%
Borena	Phase I	2020_Q3	2297	2070	90%
Borena	Phase I	2020_Q4	1408	1052	75%
Borena	Phase I	2021_Q1	1679	1280	76%
Borena	Phase I	2021_Q2	2349	2299	98%
Borena	Phase I	2021_Q3	2653	2001	75%
Borena	Phase I	2021_Q4	2104	1819	86%
Borena	Phase I	2022_Q1	2101	1880	89%
Borena	Phase I	2022_Q2	1486	1313	88%
Borena	Phase I	2022_Q3	2084	1409	68%
Borena	Phase II	2018_Q3	644	572	89%
Borena	Phase II	2018_Q4	802	1485	185%
Borena	Phase II	2019_Q1	1804	1205	67%
Borena	Phase II	2019_Q2	1953	1587	81%
Borena	Phase II	2019_Q3	2652	1462	55%
Borena	Phase II	2019_Q4	2586	2050	79%
Borena	Phase II	2020_Q1	1996	1642	82%
Borena	Phase II	2020_Q2	1886	1719	91%
Borena	Phase II	2020_Q3	2409	2011	83%
Borena	Phase II	2020_Q4	3601	2924	81%
Borena	Phase II	2021_Q1	3835	2433	63%
Borena	Phase II	2021_Q2	2844	2274	80%
Borena	Phase II	2021_Q3	3126	2754	88%

Intervention Zone	Phase of Entry into the PBF program	Reporting Quarter in Gregorian Calendar	Type of Data		data reliability
			Declared	Verified	
Borena	Phase II	2021_Q4	3478	2621	75%
Borena	Phase II	2022_Q1	2640	2472	94%
Borena	Phase II	2022_Q2	2997	2251	75%
Borena	Phase II	2022_Q3	2411	1602	66%
Borena	Phase III	2021_Q3	3219	1182	37%
Borena	Phase III	2021_Q4	6691	3964	59%
Borena	Phase III	2022_Q1	6595	5694	86%
Borena	Phase III	2022_Q2	8367	6871	82%
Borena	Phase III	2022_Q3	5428	3777	70%
Total Borena			**138631**	**100915**	**73%**
Jimma	Phase I	2019_Q3	7820	274	4%
Jimma	Phase I	2019_Q4	20094	10846	54%
Jimma	Phase I	2020_Q1	23255	17153	74%
Jimma	Phase I	2020_Q2	16859	14171	84%
Jimma	Phase I	2020_Q3	21521	18553	86%
Jimma	Phase I	2020_Q4	24390	20880	86%
Jimma	Phase I	2021_Q1	28823	26299	91%
Jimma	Phase I	2021_Q2	26579	21773	82%
Jimma	Phase I	2021_Q3	25041	21189	85%
Jimma	Phase I	2021_Q4	24934	22605	91%
Jimma	Phase I	2022_Q1	27000	23397	87%
Jimma	Phase I	2022_Q2	27500	22763	83%
Jimma	Phase I	2022_Q3	26541	20419	77%
Jimma	Phase II	2020_Q4	3589	10	0%
Jimma	Phase II	2021_Q1	14990	6848	46%
Jimma	Phase II	2021_Q2	13596	11633	86%
Jimma	Phase II	2021_Q3	15867	12509	79%
Jimma	Phase II	2021_Q4	15429	13733	89%
Jimma	Phase II	2022_Q1	18832	16888	90%
Jimma	Phase II	2022_Q2	17691	15479	87%
Jimma	Phase II	2022_Q3	18989	12798	67%
Total Jimma			**419340**	**330220**	**79%**
West Gojjam	Phase I	2021_Q4	3468	49	1%
West Gojjam	Phase I	2022_Q2	10110	3348	33%
West Gojjam	Phase I	2022_Q3	6888	5650	82%
Total West Gojjam			**20466**	**9047**	**44%**

Appendix: 3. List of Organizations Consulted and Number of People Interviewed

S.N.	Name of Organization	# of people actually interviewed
1	Federal Ministry of Health (FMoH)	1

2	The Embassy of the Kingdom of The Netherlands in Ethiopia	1
3	Oromia Regional Health Bureau	4
4	Amhara Regional Health Bureau	2
5	Borena Zone Health Department	1
6	Jimma Zone Health Department	1
7	West Gojjam Zone Health Department	1
8	Jimma Zonal Finance and Economic Development Department	1
9	Borena Zone Finance and Economic Development Department	1
10	West Gojjam Finance and Economic Development Department	1
11	Woreda Health Offices in intervention woredas	10
12	Woreda Finance and Economic Development Offices	10
13	Hospitals in intervention woredas	6
14	Health Centers in intervention woredas	22
15	Cordaid Global Office	2
16	Cordaid Country Office program staff	2
17	Jimma Zone- Cordaid Field Office Staff/ PPA	1
18	Borena Zone- Cordaid Field Office Staff/ PPA	2
19	West Gojjam Zone- Cordaid Field Office Staff/ PPA	2
Total		71

Appendix: 4. List of Health Facilities Assessed for End of Program Evaluation

S.N.	Region	Zone	Woreda	Name of Health Facility	Facility Category
1.	Oromia	Borena	Yabelo Town	Dikale PHCU	Intervention
2.	Oromia	Borena	Wachile	Moyale General Hospital	Intervention
3.	Oromia	Borena	Wachile	Webi PHCU	Intervention
4.	Oromia	Borena	Moyale	Wacille PHCU	Intervention
5.	Oromia	Borena	Dire	Magado PHCU	Intervention
6.	Oromia	Borena	Dire	Mega PHCU	Intervention
7.	Oromia	Borena	Dire	Soda PHCU	Intervention
8.	Oromia	Guji	Abaya	Guangua health center	Control
9.	Oromia	Guji	Bule Hora	Bulehora Health Center	Control
10.	Oromia	Guji	Galana	Tore Health Center	Control
11.	Oromia	Guji	Galaana	Hospitaala Jalqaba Galaanaa	Control
12.	Oromia	Jimma	Dedo	Meteso PHCU	Intervention
13.	Oromia	Jimma	Dedo	Lalo PHCU	Intervention
14.	Oromia	Jimma	Dedo	Sheki PHCU	Intervention
15.	Oromia	Jimma	Dedo	Korjo PHCU	Intervention
16.	Oromia	Jimma	Dedo	Dedo Primary Hospital	Intervention
17.	Oromia	Jimma	Jimma	Shenan Gibe General Hospital	Intervention
18.	Oromia	Jimma	Mencho	Bilu Harsu PHCU	Intervention
19.	Oromia	Jimma	Mencho	Darge Bortolo PHCU	Intervention
20.	Oromia	Jimma	Mencho	Mole PHCU	Intervention
21.	Oromia	Jimma	Agaro	Agaro General Hospital	Intervention
22.	Oromia	Jimma	Setema	Sentema Kecha PHCU	Intervention

S.N.	Region	Zone	Woreda	Name of Health Facility	Facility Category
23.	Oromia	Jimma	Setema	Gatira Health Center	Intervention
24.	Oromia	Jimma	Setema	Setema Primary Hospital	Intervention
25.	Oromia	Jimma	Setema	Gesecha PHCU	Intervention
26.	Oromia	Jimma	Omo Beyem	Dakano Elke PHCU	Intervention
27.	Oromia	Jimma	Omo Beyem	Yela sasach Health center	Intervention
28.	Oromia	Jimma	Sokoru	Gebjiro PHCU	Intervention
29.	Oromia	Jimma	Sokoru	Deneba PHCU	Intervention
30.	Oromia	Bedele	Bedele	Haro Kera HC	Control
31.	Oromia	Bedele	Bedele	Haro Kamise HC	Control
32.	Oromia	Bedele	Bedele	Gamada HC	Control
33.	Oromia	Bedele	Chora	Gefo HC	Control
34.	Oromia	Bedele	Chora	Kumbabe HC	Control
35.	Oromia	Bedele	Chora	Abdella HC	Control
36.	Oromia	Bedele	Chora	Kiltu Shibo HC	Control
37.	Oromia	Bedele	Chora	Ababora HC	Control
38.	Oromia	Bedele	Diddesa	Chalo Health Center	Control
39.	Oromia	Bedele	Gechi	Hurufa Health Center	Control
40.	Oromia	Bedele	Bedele	Bedele General Hospital	Control
41.	Oromia	West Gojam	North/ Semen Achefer	Yismala Health Center	Intervention
42.	Amhara	West Gojam	North/ Semen Achefer	Kunzila Health Center	Intervention
43.	Amhara	West Gojam	North/ Semen Achefer	Liben Primary Hospital	Intervention
44.	Amhara	West Gojam	South Achefer	Kat HC	Control
45.	Amhara	West Gojam	South Achefer	Durbetie Primary Hospital	Control

Appendix: 5. Effectiveness per Unit Delivered (DALYs)

Effectiveness per Unit Delivered (DALYs)

S/N	Health Services	Health Center	Hospital	Source-HC	Source-Hospital
1	Home visit	0 - 0.115	N/A	Jha, Bangoura et al. (1998)	
2	Outpatient consultation	0.333	0.901	Jha, Bangoura et al. (1998)	Jha, Bangoura et al. (1998)
3	Inpatient day	0.236	0.236	Jha, Bangoura et al. (1998)	Creese, Floyd et al. (2002)
4	Minor surgery	1	N/A	World Bank Estimation 2020	
5	Major surgery	N/A	9.318		World Bank 2020 calculations based on Shillcutt et al, 2010
6	Hospital referral	0	N/A	-	
7	Center counter-referral	N/A	0		
8	Family planning	0.069	N/A	World Bank Estimation 2020	
9	FP: Sterilization	N/A	0.763		World Bank 2020 calculations
10	FP: Contraception	N/A	0.069		World Bank 2020 calculations
11	Prenatal consultation	0.136	N/A	Terris-Prestholt, Watson-Jones et al. (2003)	
12	ANC: Malaria treatment	0.571	N/A	Sicuri, Bardaji et al. (2010)	
13	ANC: Tetanus vaccine	0.001		Broughton (2016)	
14	Institutional delivery	0.055	0.055	World Bank Estimation (2020)	World Bank Estimation (2020)
15	Complicated delivery	N/A	40		Jamison, Breman et al. (2006)
16	Postnatal consultation	0.011	N/A	Authors' calculations based on Jamison, Breman et al. (2006)	
17	Fully immunized child	0.69	N/A	Jha, Bangoura et al. (1998)	
18	Growth monitoring	0.029	N/A	Jha, Bangoura et al. (1998)	
19	Child treated for severe malnutrition	0.021	N/A	Jha, Bangoura et al. (1998)	
20	HIV VCT	1.55	1.55	Creese, Floyd et al. (2002)	Creese et al., 2002
21	HIV ARV treatment	0.367	0.367	Yazdanpanah, Losina et al. (2005)	Yazdanpanah et al., 2005
22	PMTCT HIV screening	1.55	N/A	Creese, Floyd et al. (2002)	

23	PMTCT HIV ARV treatment	N/A	5.155		Kuznik, Lamorde et al. (2012)
24	PMTCT HIV newborn	N/A	5.155		Kuznik, Lamorde et al. (2012)
25	TBC screening	0	0		-
26	TBC treatment	4.715	5.388	Jha, Bangoura et al. (1998)	Jha, Bangoura et al. (1998)

Source: The World Bank. https://documents1.worldbank.org/curated/en/238291593572868686/text/Cost-Effectiveness-Analysis-of-Performance-Based-Financing-for-the-Delivery-of-a-Health-Benefits-Package-in-The-Republic-of-Congo-HRBF-Impact-Evaluation.txt