

Influence of Monitoring and Evaluation Systems on the Performance of County Flagship Projects in Kwale County, Kenya

SAMUEL TSUMA MKUTA¹, HELLEN KAMWELE², ISABELLA LUTZILI ASAMBA³

¹ Department of Urban Management, School of Planning and Architecture, Maseno University, Kenya

^{2,3} Department of Urban and Regional Planning, School of Planning and Architecture, Maseno University, Kenya

Abstract— Kenyan counties are required by law to operate integrated monitoring and evaluation (M&E) systems, yet about 30% of the flagship projects in Kwale County continue to run late and over budget. This study examined how three features of the county's M&E system, namely data quality, budgetary allocation and the technical capacity of M&E personnel, influence the performance of these projects. A descriptive mixed-methods design with a cross-sectional survey was used. Evidence came from 119 completed questionnaires (a 96.0% response rate), fifteen key informant interviews, three focus group discussions and a review of county planning and budget documents. Quantitative data were analysed with Pearson's correlation and binary logistic regression, and interview material was analysed thematically. All three pillars were rated poor. Mean index scores on a five-point scale were 1.94 for data quality, 1.69 for budgetary allocation and 1.86 for technical capacity, and 62.2% of respondents rated project performance as poor or very poor. Each pillar correlated strongly with performance ($r = 0.712, 0.684$ and 0.731 respectively; $p < 0.001$). In the logistic model all three remained significant predictors, and together they explained 63.3% of the variation in satisfactory performance (Nagelkerke $R^2 = 0.633$). Technical capacity had the largest odds ratio (3.71), followed by data quality (3.27) and budgetary allocation (2.52). Interview evidence suggested that the three weaknesses feed one another. The results indicate that county M&E reform should deal with data systems, protected funding and staff capacity together and not as separate projects.

Index Terms— Budgetary allocation, County flagship projects, Data quality, Monitoring and evaluation, technical capacity

I. INTRODUCTION

Monitoring and evaluation (M&E) systems are widely treated as instruments of good governance. They allow governments to follow progress, enforce

accountability and produce the evidence on which development decisions rest [1]. In many developing countries, and in sub-Saharan Africa in particular, these systems work under tight constraints. Data are often unreliable; budgets for M&E are thin and M&E trained staff are scarce. Each of these weaknesses reduces what a project can achieve.

Kenya's National M&E Policy of 2022 places an effective M&E system on three pillars [2]. The first is data quality, which supports reliable decision-making [3]. The second is an adequate and protected budget. The third is the technical capacity of the staff who design and run monitoring frameworks. The three pillars were not picked arbitrarily. They stand for the informational, financial and human prerequisites of any working M&E system, and they correspond directly to the objectives of this study.

Kenyan counties carry a legal mandate to establish County Integrated Monitoring and Evaluation Systems (CIMES). In practice the record is uneven. Counties invest too little in staff capacity [4], struggle with data management and a shortage of trained personnel [5], and set aside small budgets for the function. Kwale County reflects this national pattern. About 30% of its flagship projects have suffered implementation delays and cost overruns [6], [7]. Several studies have looked at M&E in Kwale [8]; [10], but none has examined the three pillars together.

That omission matters for planning. Single-variable findings from other counties cannot tell a Kwale planner whether data, money or skills should come first when resources are limited. This study therefore tested the three pillars side by side in one model so that

their relative weight could be seen. The main objective was to assess the influence of M&E systems on the performance of flagship projects in Kwale County.

II. PROBLEM STATEMENT

M&E systems are broadly recognised as critical drivers of project performance because they provide the evidence base for tracking implementation, enforcing accountability and guiding resource use. Kenya's National M&E Policy (2022) treats data quality, budgetary allocation and staff technical capacity as the three foundational pillars of a functioning M&E system. Despite this binding policy mandate, roughly 30 percent of Kwale County's flagship projects continue to experience implementation delays or budget overruns, suggesting that the county's M&E systems are not yet operating at the level required to shape project outcomes positively.

A review of the existing literature confirms a clear empirical gap. Okello and Mugambi (2015) examined a single sector descriptively; Ruwa (2016) treated M&E as a secondary concern within a donor-funded livelihoods project; Safari and Kisimbi (2020) concentrated on stakeholder participation without addressing the core attributes of M&E systems; and Otundo Richard and Martin (2024), while linking M&E effectiveness to project performance, focused on donor-funded projects that operate under accountability arrangements distinct from those governing county flagship projects. Studies from other counties, such as Roba (2022) in Marsabit and Murei et al. (2017) in Nakuru, confirm the relevance of M&E budgeting and staff capacity to project outcomes, but none was conducted in Kwale County, none focused on flagship projects, and none examined all three M&E pillars simultaneously. This study responded to that gap by pursuing three interrelated objectives that, together, offer the first integrated empirical account of how M&E systems shape flagship project performance in Kwale County.

III. OBJECTIVES OF THE STUDY

The study was guided by the main objective of assessing the influence of monitoring and evaluation systems on the performance of flagship projects in Kwale County. Three specific objectives were pursued:

- i. To examine the influence of data quality on the performance of M&E systems in tracking the progress of flagship projects in Kwale County.
- ii. To assess the influence of budgetary allocation for M&E activities on the performance of flagship projects in Kwale County.
- iii. To assess the influence of the technical capacity of M&E personnel on the performance of flagship projects in Kwale County.

IV. JUSTIFICATION AND SCOPE OF THE STUDY

The study carries theoretical, methodological and policy value. Theoretically, it advances understanding of M&E effectiveness by examining data quality, budgetary allocation and technical capacity together, rather than in isolation as most prior studies have done, thereby contributing to a more context-sensitive theory of M&E in devolved governance settings. Methodologically, it departs from the largely descriptive approaches used in earlier Kwale County research by applying Pearson correlation and logistic regression alongside qualitative triangulation, setting a more rigorous precedent for future research in comparable counties. At the policy level, the findings offer Kwale County Government evidence-based guidance on which M&E dimension most strongly influences flagship project performance, directly supporting the implementation of the National M&E Policy (2022) and the operationalization of CIMES. The study covered eleven flagship projects spanning infrastructure, health, education, water, tourism and agriculture, implemented across Kwale County's four sub-counties, Matuga, Kinango, Msambweni and Lunga-Lunga, between 2018 and 2025.

V. LITERATURE REVIEW AND THEORETICAL FRAMEWORK

A. Flagship Projects and the National M&E Framework

County flagship projects are strategic, government-led initiatives meant to speed up economic growth and improve service delivery under decentralised governance. In Kenya they are aligned with Kenya Vision 2030 and delivered through County Integrated Development Plans. Their performance is usually judged on effectiveness, efficiency, sustainability and stakeholder satisfaction. Kerzner [11] and the Project Management Institute [12] argue that these outcomes depend on technical capacity, sound governance, active stakeholder management and a working M&E system.

Two traditions inform how M&E effectiveness is understood here. Kusek and Rist place M&E within results-based management [1]. Patton's utilisation-focused evaluation puts the emphasis on findings that intended users can act on, instead of on data gathered for its own sake [13]. The 2022 National M&E Policy builds on an earlier policy of 2012 and on a decade of devolved governance [2]. It names data quality, protected budgets and staff skills as the foundations of an effective system. Whether CIMES succeeds therefore turns on exactly the three variables examined in this study.

B. Evidence on the Three Pillars

Data quality. Data quality covers accuracy, completeness, timeliness, reliability and usefulness. The National M&E Policy treats it as a precondition for evaluation results that anyone can trust [2]. Even so, few county-level studies have treated data quality as a variable in its own right. Okello and Mugambi studied a school health programme in Kwale [8]. Otundo Richard examined monitoring, evaluation, accountability and learning in donor-funded projects in the county but did not isolate data quality [10].

Budgetary allocation. Kusek and Rist link the effectiveness of M&E to financial commitment [1]. In Kenya, Murei *et al.* found that inadequate M&E resourcing weakened data collection and monitoring in the horticulture sector of Nakuru County [4]. Roba reported that water projects receiving more M&E funding performed better in Marsabit County [5]. Neither study dealt with flagship projects in Kwale.

Technical capacity. Okello and Mugambi identified human resource capacity as the most important determinant of M&E effectiveness in the Kwale school health programme, although their measurement was descriptive [8]. Roba reached a different result in Marsabit. Capacity building on its own showed only a weak and statistically insignificant relationship with performance, which she read as a sign that training must be backed by adequate funding and sound data systems [5]. No quantitative study in Kwale has yet tested how the capacity of M&E staff relates to flagship project performance.

C. Theoretical and Conceptual Framework

The study was anchored on the Integrated Monitoring and Evaluation Influence Theory (IMEI Theory), which explains how the three pillars act together on project outcomes. The theory does not belong to a single author. It draws on utilisation-focused evaluation [13] and on theory of change and contribution analysis [14]. IMEI Theory holds that data quality, budgetary allocation and technical capacity are interdependent. Data quality is the foundation, because it lets stakeholders follow progress on accurate and timely information. Budgetary allocation is the enabling mechanism that keeps evaluation, training, analytical tools and reporting running. Technical capacity determines how well the system is executed, since skilled staff are needed to design frameworks, analyse data and turn findings into advice that managers can use.

The theory carries a claim that this study tested directly. If the pillars depend on one another, a weakness in one should limit what the others can achieve. A county with excellent data systems and skilled officers will still fail if no money reaches the M&E function, and a well-funded unit staffed by untrained officers will produce numbers that nobody should trust. The theory has recognised limits. Attribution is hard to establish, the influence of M&E on decisions is difficult to measure, and the theory assumes continuous stakeholder engagement that a resource-constrained county may struggle to sustain. These limits do not displace it as the most suitable framework for a study of three interacting pillars.

Fig. 1 shows the conceptual framework. The three pillars enter as independent variables, with M&E

system effectiveness as the mechanism that converts them into project results. The dependent variable, flagship project performance, was measured through timely delivery, cost efficiency and achievement of set targets, in line with the Kenya Vision 2030 results framework.

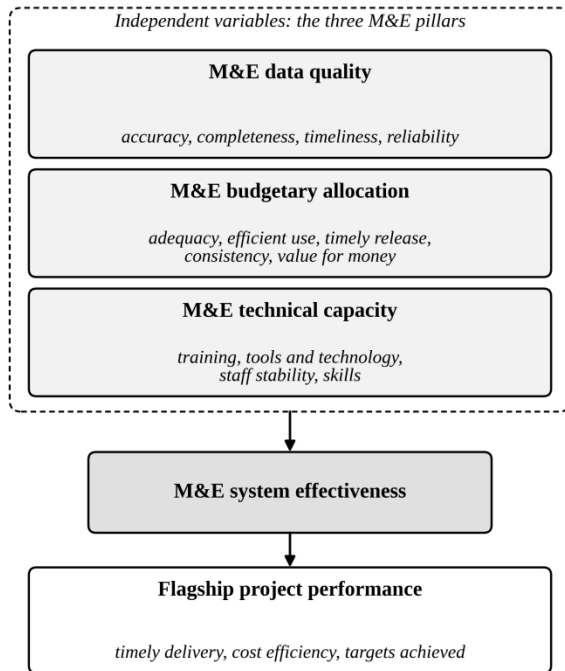


Fig. 1. Conceptual framework.

VI. RESEARCH METHODOLOGY

A. Study Area

Kwale County lies at the southern end of Kenya. It borders Taita-Taveta County to the north-west, Kilifi County to the north and north-east, Mombasa County and the Indian Ocean to the east and south-east, and Tanzania to the south-west. The county covers about 8,270 km², of which 62 km² is water, and has four sub-counties: Matuga, Kinango, Lunga-Lunga and Msambweni. Between FY 2018/19 and FY 2024/25 the county received more than KSh 3.95 billion for development projects [6]. Many of those projects have been delayed, have stalled or are only partly complete. Matuga was the central case. It hosts five flagship projects and took the largest share of funding, about KSh 1.71 billion. Only one project has been completed, and it is not serving its intended purpose. Three are stalled and one has just begun. Kinango has three projects worth KSh 1.13 billion, and the remaining three are in Lunga-Lunga and Msambweni,

which together account for about KSh 1.11 billion. In all, eleven flagship projects were covered.

B. Research Design

The study adopted a descriptive mixed-methods design with a cross-sectional survey strategy [15]. The quantitative strand used structured questionnaires to produce comparable, statistically analysable data from a broad group of stakeholders. The qualitative strand combined key informant interviews, focus group discussions and documentary analysis. These sources added contextual depth and were used to triangulate the survey results [16]. A cross-sectional approach suited the research question, which concerned relationships as they stood at one point in time and not change over a period [17].

C. Population, Sampling and Sample Size

The target population was 180 people directly involved in planning, executing or monitoring flagship projects. It comprised 104 M&E officers, 11 project managers, 10 finance and budget officials, and 55 Project Management Committee (PMC) members. Project beneficiaries were left out. The objectives required technical and institutional knowledge of data quality, M&E budgeting and personnel capacity, which beneficiaries could not reliably supply.

Purposive sampling was used to select key informants among senior officials because of their knowledge of policy and programme implementation [18]. Simple random sampling was applied within the lower-cadre categories to reduce selection bias [19]. The sample size was computed with Yamane's formula [20] at a 5% margin of error, which gave $n = 180 / (1 + 180 \times 0.05^2) = 124$. The sample was then allocated proportionately across the four respondent categories. It represents 68.9% of the population, well above the 10–30% range usually regarded as acceptable in social research [21]. Matuga contributed 100 of the 124 respondents, in line with its share of flagship projects.

D. Data Collection

Fieldwork took place from March to May 2026. The questionnaire had 73 items covering respondent background, data quality, budgetary allocation, technical capacity, overall M&E performance and project outcomes. Most items were binary, and each asked respondents to explain a "true" answer or to

explain and propose a remedy for a “false” one. Other items used five-point scales and percentage bands. The key informant guide had eight thematic sections. Fifteen interviews were held with senior officers, each lasting 60 to 90 minutes and recorded with permission. Three focus group discussions brought together 22 project managers, county budget and planning officials and PMC members for 70 to 90 minutes each. Documentary analysis covered the CIDP, Annual Progress Reports for FY 2018/19 to FY 2024/25, county budget estimates and project completion certificates. Seven project sites were also observed directly.

E. Validity and Reliability

Supervisors reviewed the instruments against the three objectives. Two Chief Officers and three practising M&E officers from Mombasa County then checked content and face validity. Their comments led to nine items being reworded and two duplicates being removed. A pilot study with twelve respondents, who were excluded from the main sample, preceded the survey. Internal consistency was assessed with Cronbach’s alpha [22]. The coefficients were 0.880 for data quality, 0.832 for budgetary allocation, 0.901 for technical capacity, 0.724 for M&E system effectiveness and 0.700 for project performance, with 0.903 for the instrument as a whole. All were at or above the 0.70 threshold [23], so every scale was retained.

F. Measurement and Analysis

Each pillar was scored on six indicators rated from 1 (very poor) to 5 (very good), and the pillar index was the mean of those six scores. Data quality was measured through accuracy, completeness, timeliness and reliability. Budgetary allocation was measured through adequacy, efficient use, timely release, consistency across years and value for money. Technical capacity was measured through training, tools and technology, staff stability and skills. Index scores from 1.00 to 2.49 were read as poor, 2.50 to 3.49 as fair and 3.50 to 5.00 as good. Project performance was measured with five yes/no indicators (objectives achieved, timely completion, completion within budget, quality of deliverables and use of M&E findings) and an overall five-point rating.

Data were analysed in SPSS version 26. Descriptive statistics, Pearson’s correlation, one-way ANOVA and chi-square tests were used, and correlation strength was classified using Cohen’s bands [24]. Binary logistic regression was the principal inferential technique, because the performance outcome was dichotomised. A project was coded as satisfactory where the overall rating was “average” or better and as unsatisfactory where it was “poor” or “very poor”. Ordinary least squares regression was unsuitable for a binary outcome, and logistic regression yields odds ratios that planners can read directly [25]. Interview and focus group material was transcribed, coded and analysed thematically following Braun and Clarke [26], using codes drawn both from the three pillars and from the transcripts.

G. Ethical Considerations

Approval was obtained from the Maseno University Ethics Review Committee. A research permit was issued by the National Commission for Science, Technology and Innovation, and the Kwale County Government gave written authorisation. Every participant received an explanation of the study and gave informed consent. Participation was voluntary, names were optional and no personally identifiable information appears in this paper. Recordings and transcripts were kept on an encrypted drive. Passages that touched on the conduct of individuals were excluded from analysis, since the study concerns systems and not persons.

VII. RESULTS AND DISCUSSION

A. Response Rate and Respondent Profile

Of the 124 questionnaires distributed, 119 were returned complete, a response rate of 96.0%. All fifteen interviews and all three focus group discussions took place, and 23 of the 25 targeted documents (92.0%) were reviewed. Among respondents, 59.7% were men, and the largest age group was 36 to 45 years (39.5%). Degree holders made up 81.5% of the sample, and 37.0% had seven or more years of M&E experience. The respondents were therefore well placed to judge the systems in which they worked.

B. Overall Status of the Three Pillars

Every index fell below the neutral midpoint of 2.50 (Table 1). Respondents judged all three pillars

inadequate, and budgetary allocation was the weakest. Between 92% and 97% of respondents placed each pillar in the poor band, and none placed any pillar in the good band. The project performance index averaged 0.96 out of a possible five affirmed indicators.

Table 1. Descriptive statistics for the study constructs (n = 119)

Construct	Mean	SD	% of Poor band, max. n (%)
Data quality index	1.94	0.33	23.4 110 (92.4)
Budgetary allocation index	1.69	0.33	17.2 115 (96.6)
Technical capacity index	1.86	0.30	21.6 115 (96.6)
Performance index (0–5)	0.96	0.85	19.2 —

Indices run from 1 to 5; % of maximum = $(\text{mean} - 1) \div 4 \times 100$ for the pillars and $\text{mean} \div 5 \times 100$ for performance.

C. Data Quality

The six data quality indicators all scored in the poor range, with item means from 1.76 to 2.14. Consistency of data across sources scored lowest (1.76) and coverage of necessary aspects scored highest (2.14). Only 38.7% of respondents said M&E data were accurate and precise, and only 29.4% said adequate verification mechanisms existed (Table 2). Reporting was slow: 74.8% said data were not collected and reported on schedule, and 37.0% said reports were rarely or never on time. Nearly half (47.0%) rated the data as unreliable or very unreliable, and 73.9% said quality control measures were not effectively applied. One senior planning officer put the problem plainly: “We often receive conflicting project status reports from the field and the project manager’s office. We cannot trust the numbers we are getting.” A focus group participant described three different figures for the completion of one project in the same reporting period, 35%, 52% and 68%, and said the middle one was submitted because it “felt most plausible”. Interviewees added that instruments had not been updated for more than three years, so that the indicators tracked no longer matched project objectives.

These patterns are familiar from other studies. In Kwale, Okello and Mugambi found mismatches between facility records and sub-county reports [8], and Otundo Richard found that project staff doubted their own data because field results did not match official reports [10]. Murei *et al.* described inconsistencies between physical and financial progress reports in Nakuru [4], and Roba found several conflicting versions of the same progress report circulating in Marsabit [5]. Taken together, the studies point to a shared structural weakness in county M&E: reporting formats are not standardised, information systems are not integrated, and verification and quality control are weak.

Data quality was strongly associated with performance ($r = 0.712$, $p < 0.001$; $R^2 = 50.7\%$). The four dimensions correlated at similar strength: accuracy and reliability ($r = 0.774$), completeness ($r = 0.731$), timeliness ($r = 0.725$) and consistency ($r = 0.619$). Such a narrow spread suggests that the dimensions do not stand apart, and that a weakness in one tends to pull down the others. Accuracy ranked first here, while Murei *et al.* ranked completeness first in Nakuru [4]. The difference may reflect the variety of flagship projects and reporting layers in Kwale, which makes the accuracy of a figure more consequential than the volume of figures.

Table 2. Respondents giving the unfavourable answer, by pillar (% , n = 119)

Indicator	%
<i>Data quality</i>	
M&E data are not accurate and precise	61.3
No adequate verification mechanisms	70.6
Data do not cover all necessary aspects	67.2
Data not collected or reported on time	74.8
Data not consistent across sources	72.3
<i>Budgetary allocation</i>	

Indicator	%
M&E budget is inadequate	83.2
Less than 1% of project budget goes to M&E	74.8
M&E funds not used efficiently	69.7
Funds released late	78.2
Funding inconsistent across fiscal years	81.5
<i>Technical capacity</i>	
Available training is inadequate	77.3
Formal training less than once a year, or never	70.6
Staff cannot use analysis software effectively	75.6

D. Budgetary Allocation

Underfunding was severe and widespread. Of the respondents, 74.8% said that less than 1% of project budgets goes to M&E, and a further 19.3% put the share between 1% and 3%. Nobody reported an allocation above 10%. These shares sit well below the 3% to 5% of project budgets commonly cited as a minimum for M&E. Documentary analysis agreed with the survey: in FY 2022/23, M&E-related spending across all eleven flagship projects came to about 0.8% of their combined budgets. Money that was allocated did not always reach its purpose. Funds were released late according to 78.2% of respondents, 44.5% said that less than half of the allocation was absorbed within the fiscal year, and 77.3% said M&E spending did not give value for money. The absence of dedicated M&E budget lines was cited by 71.4%, and 80.7% said the allocation had not improved in three years. The item means ranged from 1.51 for cost-benefit analysis of M&E to 1.88 for value for money. Interviews explained why. A county finance official described M&E as “an afterthought in our budget cycle” and the first line to be cut when budgets tighten. A senior M&E officer reported that CIMES existed on paper but had never been funded, although deploying it would cost about KSh 8 million, less than 0.2% of the annual development budget. Earlier studies describe the same dynamics. Ruwa found that the lack of a dedicated M&E budget line reduced monitoring in a Kinango food security project to an irregular activity [9]. Murei *et al.* named delayed disbursement and missing budget lines as the main obstacles in Nakuru [4], and Roba found that late releases left

monitoring work unfinished when the financial year closed in Marsabit [5].

Budgetary allocation was strongly associated with performance ($r = 0.684$, $p < 0.001$; $R^2 = 46.8\%$). Budget adequacy showed the strongest association among its components ($r = 0.791$) and transparency of the budget process the weakest ($r = 0.585$). Both remained significant. The lesson is not simply that M&E needs more money. Funds that arrive late, lapse unspent or cannot be traced will not strengthen monitoring, so reform has to cover how much is allocated, when it is released and how openly it is managed.

E. Technical Capacity

Training was rated inadequate by 77.3% of respondents and irrelevant to actual M&E duties by 74.8%. More than half (52.1%) said M&E staff receive formal training less than once a year, 18.5% said they never do, and only 10.1% reported training twice a year or more. Turnover compounded the problem. Of the respondents, 73.1% reported high turnover, 64.7% said the average M&E officer stays less than two years, and 81.5% said institutional memory is lost when staff leave. Skills and tools were also lacking. Skills for data collection and analysis were judged inadequate by 68.9%, and only 24.4% said staff could use data analysis software effectively, and 73.9% named the absence of analysis tools as the most critical gap. Overall, 57.1% rated staff competence as poor or very poor, and 16.8% rated it good or excellent.

The interviews gave these figures a human face. The county M&E head reported that only four of the 18 M&E focal points held any formal M&E qualification, and that the others “were assigned to M&E roles because they were available, not because they were qualified”. A project manager in Kinango described a seven-month vacancy during which three active projects had no formal monitoring at all. Okello and Mugambi identified the same reliance on on-the-job learning in Kwale a decade ago [8], and Otundo Richard found that teams with formal training produced better data and made more use of their findings [10]. Murei *et al.* found infrequent and impractical training to be a major obstacle in Nakuru [4].

Technical capacity showed the strongest correlation with performance of the three pillars ($r = 0.731$, $p < 0.001$; $R^2 = 53.4\%$). Its components ranked as follows: adequacy of training ($r = 0.781$), access to professional development ($r = 0.759$), proficiency with tools and software ($r = 0.743$) and alignment between qualifications and roles ($r = 0.642$). As with data quality, the four correlations cluster closely. Officers who lack training also tend to lack development opportunities, software skills and relevant qualifications, and these gaps seem to reinforce one another.

F. Flagship Project Performance

Respondents gave a poor account of project delivery. Only 18.5% said projects achieve their intended objectives, 16.0% said projects finish on time, 20.2% said they finish within budget, 23.5% said deliverables meet expected standards, and 17.6% said M&E findings are used to improve projects. On average, 19.2% of responses were affirmative. In the overall rating, 62.2% called performance poor or very poor and only 11.8% called it good or excellent. This fits the 30% delay and overrun rate recorded in the county's planning documents [6], [7]. Results-based management expects projects to succeed when M&E produces timely, reliable information that managers actually use [1]. The low use of findings (17.6%) points to a break in that chain. Interviewees said that decisions to intervene in a stalling project usually came from a political call or a contractor's complaint, and rarely from an M&E report.

G. Combined Influence of the Three Pillars

Each pillar correlated positively and strongly with performance (Table 3). Correlation treats performance as a continuous score, whereas a project either meets a standard or does not. Binary logistic regression was therefore used to estimate the joint effect. Of the 119 respondents, 45 (37.8%) rated performance as satisfactory and 74 (62.2%) as unsatisfactory. All three pillars were entered together as composite scores.

Table 3. Pearson correlation of each pillar with performance (N = 119)

Pillar	r	R ² (%)	Strongest component
Data quality	0.712	50.7	Accuracy and reliability (0.774)
Budgetary allocation	0.684	46.8	Budget adequacy (0.791)
Technical capacity	0.731	53.4	Training adequacy (0.781)

All coefficients are significant at $p < 0.001$. Cohen's bands [24] classify all three as strong (0.60–0.79).

The model improved significantly on a constant-only baseline, $\chi^2(3, N = 119) = 74.35$, $p < .001$. The Nagelkerke R^2 was 0.633, so the model accounts for about 63% of the variation in whether performance was rated satisfactory. The Cox and Snell R^2 was 0.465. The Hosmer–Lemeshow test was not significant, $\chi^2(8) = 6.42$, $p = .601$, which indicates close agreement between predicted and observed outcomes. The model classified 82.4% of cases correctly. All three predictors were positive and significant (Table 4 and Fig. 2). Holding the other two constant, a one-point gain in technical capacity raised the odds of satisfactory performance by a factor of 3.71, a gain in data quality by 3.27, and a gain in budgetary allocation by 2.52.

Table 4. Binary logistic regression of performance status on the three pillars

Predictor	B (S.E.)	Wald	p	Exp(B)	95% CI
Data quality	1.18 (0.31)	14.40	<.001	3.27	1.77–6.02
Budgetary allocation	0.93 (0.30)	9.65	.002	2.52	1.41–4.53
Technical capacity	1.31 (0.33)	15.43	<.001	3.71	1.93–7.15
Constant	–6.76 (1.42)	22.68	<.001	0.001	—

Outcome coded 1 = satisfactory (n = 45), 0 = unsatisfactory (n = 74). df = 1 for every term.

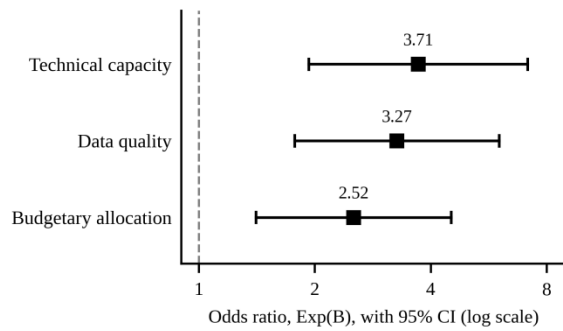


Fig. 2. Adjusted odds ratios for satisfactory performance, with 95% confidence intervals.

The ranking of the pillars should be read with care. The confidence intervals overlap widely, so the data support the conclusion that all three matter but not that one clearly outweighs another. The order does match that of the correlations, and technical capacity stands out in both. This sits uneasily with Roba's finding of a weak effect for capacity building in Marsabit [5], but the two results answer different questions. Roba looked at capacity building by itself. Here capacity was entered alongside data quality and budget, and its effect was estimated with the other two held constant. That is the situation Roba's own interpretation anticipated, in which training bears fruit only where funding and data systems are also present.

H. Qualitative Themes and Interdependence

Four themes ran through the interviews and focus groups. The first was systemic weakness in data practice: tools were not standardised across sub-counties, data collectors were poorly trained and verification protocols did not work. CIMES had been adopted on paper, but most sub-counties still relied on manual systems. The second was chronic underfunding. Participants described a cycle in which small budgets limited training, weak training lowered data quality, and poor data left no persuasive case for more M&E money. The third was loss of human capital. In Lunga-Lunga and Msambweni, PMC members said that projects had been monitored by four different officers over three years, leaving records that were discontinuous and sometimes contradictory. The fourth was a weak link between evidence and decisions. One senior official said, "We produce reports for compliance, not for learning."

Taken together, the themes describe a chain and not three separate faults. Weak budgets were said to cause weak data, because there was no money for verification visits. Unstable staffing was said to explain why available tools went unused. Weak data and unstable staffing together made senior leaders hesitant to act on findings that did exist. A county M&E head summed it up: "All three problems are connected. Solving one in isolation will not work." The statistical results point the same way. Each pillar stays significant when the other two are controlled, and the three together account for about 63% of the variation in outcome. That is the kind of joint influence IMEI Theory predicts, and no earlier Kwale study had tested it.

VIII. CONCLUSIONS AND RECOMMENDATIONS

A. Conclusions

The M&E systems behind Kwale County's flagship projects are weak on all three pillars, and the weakness is closely tied to how the projects perform. The county has put more than KSh 3.95 billion into these projects since FY 2018/19. On this evidence, whether a project delivers or stalls is closely associated with the quality of its data, the protection of its M&E funds and the skills and stability of its monitoring staff. Because those conditions are within the county's control, the findings give the county government something it can act on.

The results also support the central claim of IMEI Theory. The three pillars behave as parts of one system, so reform that treats them separately is likely to disappoint. Extra training money will achieve little where turnover is high, and a protected budget will achieve little where data cannot be trusted. For researchers, the study shows that this claim can be tested with correlation and logistic regression instead of being left as a theoretical assertion. The design can be applied in other devolved settings.

B. Recommendations

The Kwale County Government should create a dedicated M&E budget line within every flagship project budget at the design stage, and not fund M&E from what is left over in implementation budgets. This addresses the lack of dedicated lines cited by 71.4% of

respondents. It should also introduce a reconciliation step so that reports from different sources are compared before they inform decisions, which targets the lowest-scoring data quality indicator. Training for M&E staff should take place at least twice a year, and should be paired with a retention measure such as a defined M&E career track, because training yields little where turnover is as high as respondents reported. A standard cost-benefit template for M&E spending would help the county show the value of the function in later budget cycles.

The Department of Budget and Economic Planning should release M&E funds on the same schedule as core implementation funds, instead of deferring them when resources are tight. Finally, CIMES should be strengthened as a single, sequenced programme that covers data systems, protected budgeting and staff development together. Treating them as competing initiatives would repeat the fragmented approach that this study found to be insufficient.

C. Limitations and Further Research

Four limitations apply. The cross-sectional design cannot establish the direction of causation. The interviews suggest that budget constraints drive data and staffing weaknesses, but a single survey cannot confirm that order. The sample was concentrated in Matuga, so results for the other sub-counties rest on small sub-samples. Most quantitative data are self-reported perceptions and not audited figures, although documents, interviews and focus groups were used to check them. Finally, the study covers county-run flagship projects and should not be extended automatically to donor-funded projects, which operate under different accountability arrangements [10].

Three lines of work follow from these limits. A longitudinal study of the same projects across several budget cycles could test the causal order. A comparison across several counties using the same framework could show whether the prominence of technical capacity is peculiar to Kwale. A direct comparison of county-run and donor-funded projects in one county, using identical instruments, could show whether their different accountability arrangements produce different M&E profiles.

ACKNOWLEDGMENT

The authors thank the Kwale County Government for authorising the study, and the officers, project managers and committee members who gave their time as respondents and interviewees. The authors are also grateful to Maseno University for ethical review and academic support.

REFERENCES

- [1] J. Z. Kusek and R. C. Rist, *Ten Steps to a Results-Based Monitoring and Evaluation System*. Washington, DC, USA: World Bank, 2004.
- [2] Republic of Kenya, *Kenya National Monitoring and Evaluation Policy*. Nairobi, Kenya: The National Treasury and Planning, State Department for Planning, 2022.
- [3] M. Gorgens and J. Z. Kusek, *Making Monitoring and Evaluation Systems Work: A Capacity Development Tool Kit*. Washington, DC, USA: World Bank, 2009.
- [4] L. Murei, H. Kidombo, and C. Gakuu, "Influence of monitoring and evaluation budget on performance of horticulture projects in Nakuru County, Kenya," *Int. J. Econ., Commerce Manage.*, vol. 5, no. 12, pp. 620–633, 2017.
- [5] A. Roba, "Monitoring and evaluation practices on performance of water projects in Marsabit County, Kenya," *Int. J. Social Sci. Manage. Entrepreneurship*, 2022.
- [6] County Government of Kwale, *County Integrated Development Plan (CIDP) 2022–2027*. Kwale, Kenya: County Government of Kwale, 2022.
- [7] County Government of Kwale, *County Annual Progress Reports 2020/21–2024/25*. Kwale, Kenya: County Government of Kwale, n.d.
- [8] L. O. Okello and F. Mugambi, "Determinants of effective monitoring and evaluation system of public health programs: A case study of school-based hand washing program in Kwale County, Kenya," *Int. J. Econ., Finance Manage. Sci.*, vol. 3, no. 3, 2015. [Science Publishing Group](#)
- [9] M. C. Ruwa, "The influence of stakeholder participation on the performance of donor funded projects: A case of Kinango Integrated Food

- Security and Livelihood Project (KIFSLP), Kwale County, Kenya,” M.A. thesis, Univ. Nairobi, Nairobi, Kenya, 2016.
- [10] M. Otundo Richard, “Strategic monitoring, evaluation, accountability, and learning (MEAL) emerging practice and the performance of donor funded development projects in Kwale County, Kenya,” *SSRN working paper*, 2024. [Online]. Available: [SSRN](#)
- [11] H. Kerzner, *Project Management: A Systems Approach to Planning, Scheduling, and Controlling*, 11th ed. Hoboken, NJ, USA: Wiley, 2013.
- [12] Project Management Institute, *A Guide to the Project Management Body of Knowledge (PMBOK Guide)*, 6th ed. Newtown Square, PA, USA: Project Management Institute, 2017.
- [13] M. Q. Patton, *Utilization-Focused Evaluation*, 4th ed. Thousand Oaks, CA, USA: Sage, 2008.
- [14] J. Mayne, “Challenges and lessons in implementing results-based management,” *Evaluation*, vol. 13, no. 1, pp. 87–109, 2007. [SAGE](#)
- [15] J. W. Creswell and V. L. Plano Clark, *Designing and Conducting Mixed Methods Research*, 3rd ed. Thousand Oaks, CA, USA: Sage, 2018.
- [16] A. Tashakkori and C. Teddlie, Eds., *SAGE Handbook of Mixed Methods in Social & Behavioral Research*, 2nd ed. Thousand Oaks, CA, USA: Sage, 2010.
- [17] M. S. Setia, “Methodology series module 3: Cross-sectional studies,” *Indian J. Dermatol.*, vol. 61, no. 3, pp. 261–264, 2016. [PubMed](#)
- [18] M. Q. Patton, *Qualitative Research and Evaluation Methods*, 4th ed. Thousand Oaks, CA, USA: Sage, 2015.
- [19] A. Bryman, *Social Research Methods*, 5th ed. Oxford, U.K.: Oxford Univ. Press, 2016.
- [20] T. Yamane, *Statistics: An Introductory Analysis*, 2nd ed. New York, NY, USA: Harper & Row, 1967.
- [21] O. M. Mugenda and A. G. Mugenda, *Research Methods: Quantitative and Qualitative Approaches*. Nairobi, Kenya: African Centre for Technology Studies, 2003.
- [22] L. J. Cronbach, “Coefficient alpha and the internal structure of tests,” *Psychometrika*, vol. 16, no. 3, pp. 297–334, 1951. [Cambridge](#)
- [23] J. C. Nunnally and I. H. Bernstein, *Psychometric Theory*, 3rd ed. New York, NY, USA: McGraw-Hill, 1994.
- [24] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, NJ, USA: Lawrence Erlbaum Associates, 1988.
- [25] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, 3rd ed. Hoboken, NJ, USA: Wiley, 2013.
- [26] V. Braun and V. Clarke, “Using thematic analysis in psychology,” *Qual. Res. Psychol.*, vol. 3, no. 2, pp. 77–101, 2006. [Taylor & Francis](#)