Welfare Gain from Using Diagnosis Contingent Incentive Contracts to Improve Malaria Treatment*

Maria Dieci † Paul Gertler‡ Jonathan Kolstad§ Carlos Paramo¶

April 13, 2025

Abstract

We examine whether a diagnosis contingent incentive contract structure improves the treatment of malaria, and whether it's best to target those incentives to patients or providers. The contract provides incentives to use rapid tests (RDTs) to diagnose patient malaria status combined with incentives to treat with antimalarial drugs (ACTs) if the patient tests positive but not if test negative. Using data from a cluster randomized field experiment with 140 pharmacies in malaria endemic regions of Kenya, we find that both patient subsidies and provider incentives significantly increased RDT testing uptake. Absent incentives, 87% of suspected malaria patients purchase ACTs, of which as many as 90% are doing so unnecessarily because they do not have malaria. Across all arms, the incentives lead to an increase RDT test use by 25 pp and a 14 pp decline in the purchase of ACTs. The effects are stronger for patient incentives than for provider incentives. Patient incentives are translated into lower prices whereas provider incentives work through information and advice. Using a model of patient choice, we estimate that diagnosis-contingent contracts increase social welfare substantially relative to program costs, with a rate of return of at least 50% across all contract types being tested. The primary gain in welfare comes from a reduction in the use of ACTs from patients who test negative and therefore do not need treatment.

*We thank William Dow, Lia Fernald, Amrita Ahuja, Stefano Bertozzi, Stephanie Bonds, Drew Cameron, Jessica Cohen, Ada Kwan, Eva Lyubich, Edward Miguel, Jacob Odipo, Bernhards Ogutu, and Prashant Yadav for helpful comments. Josh Morrow, Ruth Namai, Veronica Njeri, and Jessica Vernon from Maisha Meds where instrumental to design and implementation. Rita Cuckovich, Andrew Murithii, Salome Omondi, Carolyne Nekesa and Nettah Isavwa, provided excellent research assistance and field management. This study was approved by UC Berkeley Committee for the Protection of Human Subjects, Strathmore University Ethics Review Board in Kenya, and the Kenyan National Commission for Science, Tehcnology and Innovation. The study was registered at the AEA (AEARCTR-0004705). It was funded by grants from BMGF and USAID DIV. The authors declare that they have no material or financial interests in the results in this paper and that all errors are our own.

†mdieci@emory.edu; Emory University Rollins School of Public Health

‡gertler@berkeley.edu; University of California Berkeley, Haas School of Business and NBER

§kolstad@berkeley.edu; University of California Berkeley, Haas School of Business and NBER

¶carlos.paramo@berkeley.edu; University of California Berkeley, Department of Economics

1 Introduction

Central to the provision of medical care, and a defining feature of health care markets, is the separate but interdependent roles of diagnosis and treatment (Arrow 1963). Patients depend on the judgment of more knowledgeable medical care professionals to diagnose their medical conditions and recommend appropriate treatments. Appropriate treatment, in turn, depends on an accurate diagnosis, which uses information from laboratory tests, imaging and other assessments that require provider time and effort and costly technology. That is, diagnosis is a complement to treatment in the production of health care and health.

Despite the importance of diagnosis, most existing theoretical and empirical models of medical care provider behavior do not specify diagnosis and treatment as separate decisions (see, e.g., McGuire 2000 for a review). This leaves a critical gap in understanding the supply and demand for health care because patient and provider beliefs about the need for care and how effectively they can treat a patient depend on the quality of diagnosis (Chandra, Cutler, and Song 2011).[1]

This gap extends to payment and incentive designs that bundle diagnosis and treatment, either explicitly or implicitly, by paying providers in either some form of fee-for-service (FFS) or fixed payment (e.g. diagnostic related groups or DRGs) and patient cost-sharing. These payment mechanisms can only target patients with the average marginal benefit of treatment. These incentive mechanisms can either discourage or encourage care for everyone as opposed to increasing care for high marginal value patients and discouraging care for low marginal value patients. It would be preferable to use incentives (provider reimbursement rates, patient copays, etc.) to encourage treatment for those that have high marginal benefit and discourage treatment for those that have low marginal benefit.

The fact that payment incentives typically only target patients with the average marginal benefit of treatment may help explain why providers often sell large amounts of medically

1. Novel diagnostic technologies that are both more effective at diagnosis and also match specific individuals to specific treatments — so-called precision medicine — make understanding the diagnostic process of particular interest (Stern, Alexander, and Chandra 2017).

unnecessary tests and treatments. In fact, the over-use of medication is ubiquitous worldwide, with consequences for health care spending and patient outcomes (Das and Hammer 2014; Whitehead, Dahlgren, and Evans 2001).[2] There is also substantial evidence of overuse of low-value diagnostics and other aspects of care provision: for example, US providers substantially increased the use of magnetic resonance imaging when they could begin billing for them, with spillover to other costlier care (Baker 2010; Afendulis and Kessler 2007).

One way to better target treatment incentives to high marginal benefit patients is to use information from the diagnosis stage in the structure of treatment incentives. We propose and test a novel incentive structure that separately pays for the diagnostic effort and for treatment contingent on diagnosis for the case of malaria. Specifically, we use a diagnostic-contingent contract structure that provides incentives to first increase the use of rapid diagnostic tests (RDT) to determine if a patient tests positive for malaria.[3] Second, the contract provides additional incentives to treat using front-line antimalarial drugs (artemisinin combination therapies ACTs) only if the patient tests positive for malaria parasites. The contract encourages appropriate treatment both through generating diagnostic information about illness status (i.e. malaria positive or negative) from the RDT and through the diagnosis contingent financial incentives for ACT use. This way the incentive contract both encourages appropriate treatment and discourages unnecessary treatment.

We begin by presenting a model of consumer demand for malaria testing, and introduce diagnosis-contingent contracts to demonstrate the mechanisms driving contract design as well as explore the cost-effectiveness of these contracts from a social-welfare perspective. Evaluating the effectiveness of the contracts through a social welfare lens is important since it is not immediately clear that introducing contracts that aim to increase testing are wel-

2. For example, studies in both the US and China have found high levels of antibiotic overuse, which may affect patient outcomes and contribute to growing rates of drug resistance (Iizuka 2012; Currie, Lin, and Zhang 2011; Currie, Lin, and Meng 2014; Fleming-Dutra et al. 2016; Chen, Gertler, and Yang 2016; Daniels et al. 2019; Sulis et al. 2020; King et al. 2022). Similarly, studies in sub-Saharan Africa have found high levels of antimalarial misuse and overuse in both pharmacy and clinic settings (Ansah et al. 2010; Cohen, Dupas, and Schaner 2015; O'Meara et al. 2016; Cohen et al. 2013).

3. RDTs are highly accurate tests that use a finger prick to confirm the presence or absence of malaria parasites in a patient's blood.

fare enhancing. The effects on social welfare from these policies will depend on the relative costs of treatment and testing, the true positivity rate of patients, as well as the internalities and externalities from unnecessary treatment. Standard cost-effectiveness measures in public health measure the impact of policies on specific program outcomes such as patient health outcomes or changes in patient and provider behavior (demand for medical services or products). The approach that we present in this paper allows to quantify the change in social welfare from implementing these policies relative to the implementation costs for the social planner.

In addition to developing a toolkit to evaluate these contracts, through the model we show that contracts that reduce the price of ACTs conditional on a positive test result are more cost-effective in boosting testing uptake especially when patients overestimate their probability of being malaria-positive. Those patients are more responsive to the level of diagnosis-contingent ACT discount, by helping prevent unnecessary take-up of ACTs by patients who overestimate their probability of being malaria positive. At the same time, because the actual probability of being malaria positive is low, the expected direct costs of the contract are minimal.

Finally, our model illustrates how contracts that target the provider's incentives increase uptake. If providers pass through the discounts the results are equivalent to demand side incentives. This standard result ignores a key channel: information provision. Provider incentives can induce costly effort to inform and advise patients. Testing contingent contract incentivize providers to counsel patients and to do so accurately by encouraging testing followed by treatment only if a test is positive.

Having developed the theory of diagnosis contingent contracts, we use a cluster randomized field experiment (RCT) to estimate the effect of the diagnosis contingent incentives on malaria testing and treatment decisions as well as welfare. We further investigate whether incentives are more effective when they are given to patients through subsidies (demand-side) or to providers through performance incentives (supply-side), or a combination of the

two. We also examine the mechanisms through which the incentives work by leveraging data from an audit study that employs Standardized Patient visits, which allows us to examine behavioral channels of impact without the confounding effects of patient selection. We are, therefore, able to explore whether the effect of the incentives contract on treatment (ACT use) is driven by diagnostic information (i.e. malaria status) versus financial incentives.

Malaria is an important disease to study clinical decision-making because it is a well-understood illness, it has a high disease burden, and nearly all deaths and serious illness are preventable through effective and inexpensive medication (WHO 2021). Despite RDTs being cheap and readily available, less than 10% of patients presenting with malaria symptoms are diagnosed with RDTs or other parasitic tests prior to getting treated.[4] This may explain why large shares of malaria-positive patients go untreated while large shares of malaria-negative patients receive antimalarial medication (Cohen et al. 2013; Cohen, Dupas, and Schaner 2015; O'Meara et al. 2016; Ansah et al. 2010).[5] Low diagnostic testing contributes to a gap between treatment and need; missed diagnoses result in more severe avoidable illness and over-prescription of anti-malaria drugs to malaria-negative patients can lead to heightened drug resistance in the population. In the absence of diagnostic information it is perhaps not surprising that we see large over- and under-treatment for malaria. If providers were aware of their patients' malaria status, treatment could be far better tailored.

We explore these issues in high malaria prevalence counties in Kenya, where over 3.5 million people fall ill with malaria annually. The study population lives near Lake Victoria and on the coast, areas that are most vulnerable to infection (Initiative 2021; Disease Control and Prevention 2018). Over half of malaria patients in Kenya and across East Africa access

4. The literature identifies several potential reasons as to why diagnostic testing is low. Patients may not demand tests because of (i) strong prior beliefs about their malaria status – i.e., a low perceived value of information from testing (Maffioli et al. 2019), (ii) the cost of the test is prohibitive (Cohen, Dupas, and Schaner 2015; O'Meara et al. 2018), and (iii) they do not want to wait for the diagnostic test result – i.e., impatience. Moreover, providers may not prescribe a test prior to treatment because (i) they have established practices of symptom-based diagnosis – i.e., established norms and habits (Mbonye et al. 2013), (ii) they are optimizing perceived patient preferences (Lopez, Sautmann, and Schaner 2022), and (iii) they have profit motives (Currie, Lin, and Meng 2014).

5. Over- and under-treatment are ubiquitous worldwide with implications for both health care cost and health outcomes (Das and Hammer 2014; Whitehead, Dahlgren, and Evans 2001).

treatment via pharmacies, often the preferred access point for primary care given pharmacies' convenience and reliable presence even in areas that are under-served by public health care clinics and hospitals (Musuva et al. 2017; Burton et al. 2011).

We randomized 140 pharmacies into either a status quo control group or one of three treatment groups, each with a two-part incentive: (1) patient subsidies for RDT tests and for the anti-malarial drugs (Artemisinin Combination Therapies – ACTs) conditional on a positive test; (2) pharmacy performance incentives for RDT tests, and for prescribing ACTs conditional on a positive test; and (3) combined incentives (patient subsidies and pharmacy incentives) for RDT and ACTs for confirmed malaria-positive cases. The total value of the incentives was held constant across the three intervention arms. This design allowed us to evaluate the impact of a two-part incentive structure where payouts depend on the full continuum of care (and diagnostic information) as well as to examine the causal effect of targeting that incentive to the patient versus the provider.[6]

We find that both patient subsidies and provider incentives are effective at increasing RDT uptake and at improving targeting treatment to malaria-positive patients. Patient subsidies increase the likelihood that a symptomatic patient takes a RDT test by 27 percentage points over a control group rate of 8 percent.[7] The impact of pharmacy incentives are statistically indistinguishable, increasing the likelihood of RDT uptake by a point estimate of 20 percentage points.

Absent any interventions, 87% of suspected malaria patients purchase ACTs, of which as many as 66% are doing so unnecessarily because they do not have malaria. This represents a high baseline level of medication waste. We find that the incentives lead to an overall decline in ACT usage of 14 percentage points, that incentives increase the likelihood that a

6. Prior literature has studied the impact of demand-side subsidies, but not provider incentives, on malaria care, finding them to be effective at improving testing but not at improving test result adherence (Cohen, Dupas, and Schaner 2015; O'Meara et al. 2016; O'Meara et al. 2018). To our knowledge, at the time of writing, one other ongoing study compares patient and provider incentives in pharmacy-settings (Visser et al. 2024).

7. This result is consistent with what has been found in prior literature on consumer subsidies for for RDTs and other health goods.(Dupas 2014; Cohen, Dupas, and Schaner 2015)

patient purchases ACTs combined with a diagnostic test by 7 percentage points, and that incentives lead to large declines (16-22 percentage points) in the likelihood that patients purchase malaria treatment without a diagnostic test. The increase in ACT purchase along with a test is consistent with the positivity rate of patients seeking malaria care in our sample. Therefore, contracts are effective at decreasing unnecessary treatments by malaria negative patients while continuing to offer access to appropriate care for malaria positive patients.

We find that patient subsidies result in significantly lower RDT prices (43%) but none of the provider incentives were passed through to clients in terms of lower prices. Instead, provider incentives were associated with pharmacists giving more explanation of RDT results and counseling on treatment based on the test results. Both contracts have the same impact on ultimate demand for both testing and ACT treatments but financial incentives seemed to work through an information and advice pathway when targeted to providers whereas demand subsidies induced more patients to purchase RDTs that provided accurate illness status information leading to more appropriate use of RDTs and ACTs. Both patient and provider incentives led to fewer instances of providers telling patients who chose not to test that they were malaria positive and should purchase ACTs.

The diagnosis contingent incentive contracts are very cost-effective from a social welfare perspective. Our approach allows us to translate standard cost-effectiveness measures from the health economics literature such as impact of demand per dollar spent into implied effects on social welfare. For each dollar that the policy maker allocates into diagnosis contingent contracts, at least 0.59 additional patients shift demand towards RDT tests. On the other hand, patient and pharmacy welfare increases by at least $1.50 for each dollar that the policy maker allocates to the contracts, implying a rate of return of at least 50% in terms of social welfare. The estimated welfare gains are mostly the result of reductions in overall health expenditures by patients. These estimates are lower bounds since they exclude any potential externalities and internalities from unnecessary treatment.

This paper relates to a number of different literatures. We directly contribute to work on performance-based financing mechanisms that reward providers for both quantity and quality of health services delivered by paying for key outputs (Basinga et al. 2011; Gertler, Giovagnoli, and Martinez 2014; Ahmed et al. 2023; Yip et al. 2014; Peabody et al. 2014; Witter et al. 2012; Miller and Babiarz 2013). These studies suggest that properly incentivizing providers can lead to improvements in health care utilization and key health outcomes, but the evidence has been limited to a relatively narrow set of indicators and outcomes. Additionally, the literature on performance incentives focuses on the price effects, but ignores mechanisms through which incentives operate. Our study provides evidence on behavioral channels through which provider performance incentives may impact quality.

Second, we contribute to these two strands of literature by innovating in how health financing contracts are structured. Conditional cash transfers for preventive health visits, for example, incentivize health care utilization by lowering the cost of care to patients. In the US, insurance products that have modest copays or deductibles operate in the same way - by lowering the price patients pay. These examples, as well as other demand-side incentives for healthcare, highlight how these financing models typically operate - they reimburse a flat rate for services used through lower prices/copays (Arrow 1963; Pauly 1980; McGuire 2000; Cutler and Zeckhauser 2000).

Performance pay models that reward providers either directly through bonuses tied to services provided (see e.g.Basinga et al. 2011; Gertler, Giovagnoli, and Martinez 2014; Ahmed et al. 2023; Yip et al. 2014; Peabody et al. 2014; Witter et al. 2012; Miller and Babiarz 2013) for examples) or indirectly through capitation have a similar structure. Payments are typically made based on services provided, and do not explicitly take into account diagnostic information when setting copays for various services.[8] This fits in with other literatures on performance pay based on outcomes, including applications for medical doctors and teachers (Campbell et al. 2009; Prendergast 1999; Podgursky and Springer 2007).Financial incen-

8. Though, of course, we acknowledge that patients with different medical needs will pay different amounts for their care because services needed will be diagnosis-dependent.

tives are well-established tools used around the world to promote a wide range of health behaviors. Typically, demand and supply side incentives are studied independently, in this paper we compare them directly and put the two literatures in conversation with each other. Demand-side incentives all operate based on the assumption that either price itself is a barrier to adopting a health behavior, or an incentive can nudge people to overcome other non-pecuniary barriers. Price experiments for health treatments have shown that people do not respond uniformly to prices, and instead the nature of the health decision and timing of the benefits affect demand elasticity (Dupas and Miguel 2017; Dow, White, and Bertozzi 2016; Dupas 2014; Cohen, Dupas, and Schaner 2015; O'Meara et al. 2016).

Third, it adds to the literature on how personalized health information and financial incentives can be combined to change health behavior. Information combined with financial incentives has shown more promise in encouraging health behavior adoption (Meredith et al. 2013; Ma et al. 2014; Dupas 2011). But, the quality of the information matters: general health information tends to be less effective in changing individual behaviors than individually tailored messages targeted at the key decision-makers (Gong 2015). Studies that have examined whether information provided by a malaria diagnostic test changes treatment-seeking behavior have found mixed results – information is effective in steering some patients towards appropriate treatment options, especially when coupled with an incentive, but many elect to ignore test results when making treatment decisions (Cohen, Dupas, and Schaner 2015; O'Meara et al. 2018). This study leverages the two steps of the testing and treatment decision by providing a financial incentive for treatment conditional on the personalized health information provided by the test. We test the extent to which appropriate malaria treatment use is a result of information (RDT result) or an added financial incentive (for ACT), contributing to the long literature on the role of information and information asymmetries in health decision-making (Arrow 1963).

Finally, we contribute to the broader literature studying how incentives targeted at the demand-side or the supply-side can affect prices and demand (e.g. Busse, Silva-Risso, and

Zettelmeyer 2006). We also contribute to this conversation in the health economics literature.

2 Model

In this section we develop a simple model to elucidate the rationale for our diagnosis-contingent contract design. We begin with a model of patient demand for RDT testing. Using the derived demand we characterize our diagnosis contingent contract and explore the impact on demand and outcomes. We demonstrate the efficiency of a diagnosis contingent contract relative to direct subsides (e.g. for RDTs alone). We then turn to the joint decision process between the patient and the provider and explore the role of diagnosis contingent contracts in information provision and treatment decisions.

2.1 Patient demand for RDTs

We allow for two sequential choicesFirst, the patient decides whether or not to to test for malaria. Second, the patient must decide whether to purchase an antimalarial medication. Patient demand depends on malaria status $M \in \{m, m'\}$ with probability of malaria $P(m)$, or some other health condition causing the symptoms with probability $P(m')$, such that $P(m) = 1 - P(m')$. The value of treatment (receiving an ACT) depends on true health status.

The patient's realized utility depends on their true malaria status $M \in \{m, m'\}$ and whether they consume an antimalarial. If the patient is malaria positive and left untreated, they receive a disutility $-d_m$. Similarly, if the patient is malaria negative, the patient has disutility $d_{m'}$ from the non-malaria condition. Receiving unnecessary malaria treatment leads to a disutility represented by $-d_w$, in addition to the unnecessary expenditure on treatment. For instance, side effects from antimalarials, the true underlying condition staying untreated for longer, and malaria resistance concerns can affect the value of d_w.

Patients choose whether to buy an ACT and pay price p_a to receive malaria treatment.

We write expected utility with and without ACT purchase as:

$$\mathbb{E}(U) = \begin{cases} \text{ACT} & -p_a - P(m')(d_{m'} + d_w) \\ \text{No ACT} & -P(m')d_{m'} - P(m)d_m \end{cases} \tag{1}$$

Without the purchase of an RDT, true malaria status M is uncertain. If the patient buys an ACT, they will avoid the dis-utility of having untreated malaria, but they will have a probability $P(m')$ of incurring a dis-utility for inappropriate treatment. If the patient does not buy the ACT, the patient has probability $P(m)$ of incurring disutility d_m from having untreated malaria

We first assume that $\mathbb{E}(U(\text{ACT})) > \mathbb{E}(U(\text{No ACT}))$. That is, for the observed patients, consuming ACT will always be optimal under uncertainty. This assumption implies that the disutility d_m is sufficiently large that a patient with malaria-presenting symptoms will always chose to receive an anti-malarial treatment when their status M is uncertain.[9] Therefore, the patient's value of not testing is given by:

$$\mathbb{E}(U(\text{No RDT})) = -p_a - P(m')(d_{m'} + d_w) \tag{2}$$

Patients, on the other hand, have the option to purchase an RDT test for price p_r which will inform them on their status M, that is, $P(m|\text{RDT positive}) = 1$ and $P(m|\text{RDT negative}) = 0$. We assume that buying an ACT and an RDT is feasible for the patients (i.e. $p_a + p_r <$ B for the patient's healthcare budget B).

We further assume that if a patient knows they are malaria positive after incurring cost p_r, then the patient will purchase an ACT. This follows from the assumptions that patient will purchase an ACT under uncertainty and that they can afford both a RDT and ACT. Both assumptions are consistent with what we observed in the data.[10]. Moreover, since RDTs are assumed to remove the patient's uncertainty on their malaria status, if an RDT

9. This assumption holds in our sample, since we only observe patients that make a Malaria purchase – that is, if a patient does not purchase an RDT, they purchase an ACT.

10. The probability of buying an ACT if the patient tested positive is $> 95\%$

is negative, the patient should not purchase an antimalarial.[11]

We can now write the patient's value of purchasing an RDT as:

$$\mathbb{E}(U(\text{RDT})) = -p_r - P(m)p_a - P(m')d_{m'} \tag{3}$$

Because patients purchase an ACT after buying an RDT if they test positive, the patient's value of RDTs depends on the price of both products. RDTs do, however, guarantee that the patient will not incur disutility d_w, and allows for the possibility of avoiding unnecessary expenditures on antimalarial treatments.

Combining 2 and 3 together, the patient buys an RDT if and only if:

$$\begin{aligned} &\mathbb{E}(U(\text{RDT})) > \mathbb{E}(U(\text{No RDT})) \\ \iff &-p_r - P(m)p_a - P(m')d_{m'} > -p_a - P(m')(d_{m'} + d_w) \\ \iff &\boldsymbol{P(m')(p_a + d_w) > p_r} \end{aligned} \tag{4}$$

Equation 4 describes the patient's optimality condition for the purchase of RDTs. The decision to purchase an RDT depends on 4 values key values: the patient's beliefs about their malaria status, the patient's perceived disutility of incorrectly receiving malaria treatment, and the prices of both RDTs and ACTs. We express demand for RDTs as:

$$D(p_a, p_r) = P(P(m'|i)(p_a + d_{w,i}) - p_r > 0) \tag{5}$$

where $P(m'|i)(p_a + d_{w,i})$ is a random variable reflecting individual $i's$ beliefs about malaria risk and disutility from receiving unnecessary ACT treatment when they are malaria negative. In our notation, we allow for the possibility that patient's beliefs $\{P(m'|i), d_{w,i}\}$ about $P(m')$ and d_w to be incorrect. Moreover, since the demand is a cumulative density function,

11. Again, this is consistent with our data since the probability of buying an ACT after testing negative is 5%.

this expression yields simple comparative statistics. Demand is increasing on the price of the ACT and the patient's beliefs about $P(m')$ and d_w, and decreasing on the price of the RDT:

$$\frac{\partial D(p_a, p_r)}{\partial p_a} > 0$$
$$\frac{\partial D(p_a, p_r)}{\partial p_r} < 0$$
$$\frac{\partial D(p_a, p_r)}{\partial P(m'|i)} > 0$$
$$\frac{\partial D(p_a, p_r)}{\partial d_{w,i}} > 0$$

2.2 Patient diagnosis contingent contracts

Our model of demand for diagnosis in hand we turn to contract design. A diagnosis contingent contract simply reduces the price of the ACT that the patient pays conditional on the patient testing and the outcome of that test. We express an ACT with discounted price as $p^*_{a|r} = (1 - \delta_a)p_a < p_a$ available only if the patient purchases an RDT through the contract's program. Parameter δ_a describes the relative size of the discount. The patient can continue to purchase the ACT without an RDT at market price p_a. We extend this by reducing the cost of the RDT test by δ percent. The discounted price for RDTs is given by $p^*_r = (1 - \delta_r)p_r < p_r$ for discount rate δ_r and market price p_r. Under a diagnosis contingent contract, the patient's optimality condition becomes:

$$
\begin{aligned}
& p_a - P(m)p^*_{a|r} + P(m')d_w > p^*_r \\
\iff & p_a - P(m)(1-\delta_a)p_a + P(m')d_w > p^*_r \\
\iff & (1 - P(m)(1-\delta_a))p_a + P(m')d_w > (1-\delta_r)p_r
\end{aligned} \tag{6}
$$

When $\delta_r = \delta_a = 0$, this condition is identical to equation 4. When $\delta_r \in [0,1]$ and $\delta_a \in [0,1]$, the comparative statistics implied by equation 4 continue to hold. However, patient demand for RDTs is now increasing on the discounts δ_r and δ_a:

$$
\begin{aligned}
\frac{\partial D(p_a, p_r, \delta_a, \delta_r)}{\partial \delta_r} > 0 \\
\frac{\partial D(p_a, p_r, \delta_a, \delta_r)}{\partial \delta_a} > 0
\end{aligned}
$$

Diagnostic contingent contract structures that target prices paid by patients for both testing and treatment based on the outcome of the test, increase demand for malaria testing and prevent unnecessary treatments. When deciding whether to test, patients not only care about the price of the RDT, but also about the cost of treatment. Our proposed contract increases demand for testing through changes in the price of the both products. A lower price for RDTs increases demand for testing. Second, a conditional discount on ACTs indirectly increasing the value that testing by making the expected cost of testing positive lower.

Note that demand for ACTs is fully determined by the demand for RDTs in this model. Diagnosis contingent contracts reduce demand for malaria treatment on expectation since patients who purchase tests as a result of these interventions will only purchase treatment if they are actually positive. In particular, demand for ACTs is given by:

$$D_{ACT}(p_a, p_r, \delta_a, \delta_r) = 1 - P(m')D(p_a, p_r, \delta_a, \delta_r) \tag{7}$$

Whether this reduction in treatment at the expense of more testing is welfare enhancing from the social planner's perspective is not immediately clear. On one hand, patients might fail to internalize externalities and internalities from unnecessary treatment (akin to the "internalities" discussed in (Baicker, Mullainathan, and Schwartzstein 2015)), or overestimate their probability of being malaria positive, in which case interventions that increase demand for testing could lead to increases in patient and social welfare. On the other hand, increasing demand for testing could prove distortionary if it leads to an increase in medical expenditures that is beyond the value from avoiding unnecessary treatments. Section 2.4 discusses the welfare implications of diagnosis contingent contracts in more detail.

2.3 Provider diagnosis contingent contracts

Providers counsel malaria suspect patients on the value of testing. For simplicity, assume that providers can signal the value of testing to the patient through $\theta \in \{0, 1\}$, whether the provider recommends to be tested or not. Patients update their beliefs about the true value of $P(m)$ and d_w based on the provider's counseling, with functions given by $d_{w,i}(\theta)$ and $P(m|\theta, i)$.

If a provider recommends a test $\theta = 1$, it is likely that the patient will interpret this as a signal that either $P(m')$ or d_w are high, or equivalently, that the value of testing is high. If this is the case, then the provider's recommendation to test increases the demand for RDTs:

$$D(p_a, p_r|\theta = 1) - D(p_a, p_r|\theta = 0) > 0 \tag{8}$$

Provider's motivations to recommend testing are potentially twofold. On one hand,

providers care about the patient's welfare and the potential for increasing malaria resistance with unecessary treatments. On the other hand, providers might care about their financial incentives. For concreteness, let the provider's decision to recommend testing be given by:

$$\theta = 1\{W(d_w, d_{sw}, P(m), p_r, p_a, \delta_r, \delta_a) + \lambda f(\boldsymbol{\pi}, t) > e_d\} \tag{9}$$

Such that $W(d_w, d_{sw}, P(m), p_r, p_a, \delta_r, \delta_a)$ is a function that represents the provider's internalized patient's welfare and concerns about malaria resistance in their community, and $f(\boldsymbol{\pi}, t)$ represents the provider's financial incentives to recommend testing. Parameter d_{sw} denotes the marginal externality of a patient receiving unnecessary treatment while d_w denotes internalities to the patient from unnecessary treatment as discussed above. The provider's financial incentives are a function of the vector of markups $\boldsymbol{\pi}$ for all the malaria products sold in the pharmacy and a vector of any incentives included in the diagnosis contingent contracts (t). In particular:

$$\begin{aligned} f(\boldsymbol{\pi}, t) &= \mathbb{E}[\Pi|\theta = 1] - \mathbb{E}[\Pi|\theta = 0] \\ &= \Sigma_k(\pi_k + t_k)(P[k|\theta = 1] - P[k|\theta = 0]) \end{aligned} \tag{10}$$

Such that Π are the provider's expected profits from a malaria counseling interaction with the patient. These profits depend on the probability of patients deciding to buy product $k \in \{r, a, a|r\}$ conditional on their advice to the patient. The provider's incentives include direct transfers to the provider from the diagnosis contingent contract, and the markups of the pharmacy for the sale of the distinct available products.

Provider diagnosis contingent contracts change the financial incentive structure from the sale of the malaria products, encouraging providers to recommend testing. In particular, these contracts increase $\pi_k + t_k$ for $k \in \{r, a|r\}$. Since θ is likely to be positively correlated with the patient's beliefs about the value of testing, the probabilities that patients buy an

RDT (r) or an ACT conditional on an RDT sale ($a|r$) should both be increasing on θ. In other words, $P[r|\theta = 1] - P[r|\theta = 0] > 0$ and $P[\{a|r\}|\theta = 1] - P[\{a|r\}|\theta = 0] > 0$. Therefore, provider diagnostic contingent contracts increase the provider's incentives to recommend testing $f(\boldsymbol{\pi}, t)$, and thus, potentially increase demand for testing by the patient.

2.4 Patient Welfare

We now turn to the analysis of the expected welfare effects of diagnosis contingent contracts, beginning with the effects of the contracts on patient welfare. Absent diagnosis-contingent contracts, expected patient welfare is given by:

$$\begin{aligned}\mathbb{E}(Welfare) &= P(RDT) * \mathbb{E}[U(RDT)] + P(RDT') * \mathbb{E}[U(NoRDT)] \\ &= -p_a[(1 - D(0)) + D(0)P(m)] \\ &-p_r D(0) - d_{m'} P(m') - d_w P(m')(1 - D(0))\end{aligned} \tag{11}$$

Equation 11 describes the welfare from a patient seeking malaria care implied by the model. First, define $\delta = \{\delta_a, \delta_r, t_r, t_{a|r}\}$ to be the vector of incentives from the contract including patient discounts (δ_a, δ_r) and provider transfers $(t_{a|r}, t_r)$. Then the demand for an RDT in the absence of a contract is given by $P(RDT) = D(\delta = 0) = D(0)$.

On the other hand, when diagnosis contingent contracts are available with incentives $\delta = \hat{\delta}$, patient welfare becomes:

$$\begin{aligned}\mathbb{E}(Welfare) &= P(RDT) * \mathbb{E}[U(RDT)] + P(RDT') * \mathbb{E}[U(NoRDT)] \\ &= -p_a[(1 - D(\hat{\delta})) + D(\hat{\delta})P(m)(1 - \delta_a))] \\ &-p_r D(\hat{\delta})(1 - \delta_r) - d_{m'} P(m') - d_w P(m')(1 - D(\hat{\delta}))\end{aligned} \tag{12}$$

In order to get the welfare effect from the introduction of a contract with incentives $\hat{\delta}$, we take the difference between equations 12 and 11. For simplicity, define $\Delta_D := D(\hat{\delta}) - D(0)$ as the contract's effect on the demand for RDTs. Welfare gains are thus:

$$\underbrace{p_a[\Delta_D P(m') + P(m)D(\hat{\delta})\delta_a]}_{\text{Unnecessary ACT expenditure}} - \underbrace{p_r[\Delta_D - D(\delta)\delta_r]}_{\text{RDT expenditure}} + \underbrace{d_w P(m')\Delta_D}_{\substack{\text{Internalities} \\ \text{from unnecessary ACT}}} \tag{13}$$

Equation 13 is composed of three terms. The first term shows the effect on patient welfare due to changes in ACT expenditures. Note that since contracts have a positive effect on demand for RDTs ($\Delta_D > 0$), the first term is positive. In other words, patient welfare increases due to a reduction on ACT expenditures. The second term corresponds to the effect of contracts on patient welfare due to changes in RDT expenditures. The sign of this term is unclear, and will depend on whether the decrease in testing costs due to the contract outweigh increased demand for testing. Finally, the third term relates to an increase in patient welfare from avoiding unnecessary treatment. Contracts decrease unnecessary ACT uptake due to increased testing uptake. This term should be a lower bound for patient welfare, given that there are likely internalities associated with appropriate malaria careNote that the overall welfare effect to the patient is ambiguous, given that patient expenditures for testing could increase above their optimal level since patients can have incorrect beliefs about $P(m)$ and their private values for d_w.

2.5 Social welfare

To study the social welfare effect of this contract, one needs to consider the costs of implementing the contract and the pharmacy welfare in addition to the patient welfare effects discussed above. First, the policy maker introduces incentives δ which are paid whenever the patient tests or purchases an ACT conditional on a positive test result. The expected costs to the social planner from implementing the contract net any externalities generated

by the contract are given by:

$$D(\hat{\delta})[\delta_r p_r + t_r] + D(\hat{\delta})P(m)[\delta_a p_a + t_{a|r}] - d_{sw}P(m')\Delta_D \tag{14}$$

Note that while the social planner pays incentives δ_r and t_r for every patient who buys an RDT, the incentives associated with treatment (δ_a and $t_{a|r}$) are only paid for malaria positive patients. Moreover, the social planner internalizes any externalities generated by the contract.

Second, the profits to the pharmacy or clinic offering malaria care to the patient can change as a result of the introduction of these contracts. Taking into account the pharmacy welfare is both important for studying social welfare effects and the sustainability of the policy. The effect of the contracts on pharmacy profits are ambiguous. This is because pharmacy welfare will depend on the profit margins for treatment and testing, in addition to the effects on patient choice probabilities. The effect on the pharmacy profits from these contracts is given by:

$$E[\Pi|\delta = \hat{\delta}] - E[\Pi|\delta = 0] = \Delta_D[\pi_r - \pi_a P(m')] + D(\hat{\delta})[t_r + P(m)t_{a|r}] \tag{15}$$

Where Π is the expected profits from a malaria-suspect patient encounter, $\pi_a = p_a - c_a$ is the profit margin for treatment, and $\pi_r = p_r - c_r$ is the profit margin for testing[12]. The marginal costs of testing and treatment are given by c_a and c_r. The first term in equation 15 relates to the change in profits from the introduction of the contract, whereas the second term relates to transfers from the policy maker to the pharmacy as part of the contract. In the results section, we find evidence that profits to pharmacies either stay unchanged or increase in our setting.

With the effects of these policies on patient welfare, pharmacy profits, and program costs in hand, we can now derive an equation for total social welfare effects from diagnosis-

12. Without loss of generality, assume that $\pi_a = \pi_{a|r}$; where differences in profitability across the a and $a|r$ pseudo-products are captured by transfer $t_{a|r}$.

contingent contracts by adding up the distinct components. Therefore, the effect on social welfare is given by:

$$\begin{aligned} &\Delta_D[P(m')p_a - p_r + d_w P(m') + \pi_r - \pi_a P(m') + d_{sw} P(m')] \\ &= \Delta_D[P(m')c_a - c_r + (d_w + d_{sw})P(m')] \end{aligned} \tag{16}$$

Note that the sign of the effect on social welfare of diagnosis-contingent contracts is ambiguous. Since the effect on demand for testing is positive, social welfare effects will depend on the relative magnitudes of marginal costs of testing and treatment — c_a and c_r — as well as the positivity rate of malaria-suspect patients — $P(m)$. In fact, equation 16 implies the following result:

Theorem 1 *Let* $P(m') > \frac{c_r}{d_w + d_{sw} + c_a}$*, then a diagnosis-contingent contract is welfare-improving.*

Condition in proposition 1 has an intuitive interpretation. First, if the probability of patients having malaria is very high ($P(m') \simeq 0$), these contracts are unlikely to increase welfare since most patients will end up consuming antimalarial regardless of whether they purchase a test or not. Second, the higher the cost of testing relative to treatment ($c_r >> c_a$), the less likely that a contract that steers demand towards testing will increase social welfare. Third, if unnecessary treatments lead to costly externalities and internalities (d_{sw}, d_w), these contracts are likely to increase social welfare. Fourth, the higher the cost of treatment (c_a), the more likely that contracts that avoid unnecessary health expenditures will be welfare-improving[13].

13. While this condition relates to whether these contracts are welfare-improving if enforced, it does not guarantee that pharmacies will be willing to participate in the program. If these contracts lead to social gains, incentives to the pharmacy ($t_{a|r}$ and t_r) rather than the patient (δ_a and δ_r) can be a useful tool to align incentives and increase participation.

2.6 Observed lower bound on social welfare effects

Note that the value of d_w, which relates to internalities and externalities from over-treatment, is difficult to measure and unlikely to be directly observed. However, one can bound the social welfare effects of these contracts presented in equation 16 by focusing on market characteristics that are more commonly observable such as marginal costs and positivity rates. This is because unnecessary treatments do not carry any benefits ex-post so the term that includes $d_w + d_{sw}$ is either null or positive. In particular, a bound on social welfare effects from these contracts is given by:

$$\Delta_D[P(m')c_a - c_r] \tag{17}$$

This lower bound on social welfare includes the change in patient expenditures, profits for the pharmacy, and costs of introducing such contracts for the policy maker.

If marginal costs are unobserved and only data on product prices is available, an alternative lower bound can be constructed. This is because by revealed choice, the change in profits for pharmacies that accept these contracts is either positive or null, so one can focus solely on the components that relate to changes in patient expenditures, pharmacy transfers from the contract, and contract costs, leading to the following lower bound:

$$\Delta_D[P(m')p_a - p_r] \tag{18}$$

2.7 Cost-effectiveness of a diagnosis contingent contract

We now discuss which contract characteristics are more desirable for a policy maker interested in increasing social welfare. In order to evaluate different policy designs from the policy maker's perspective we need to introduce an additional metric. Given a fixed budget, a policy maker is not just interested in increasing welfare, but rather doing so in a way that maximizes the impact of each dollar spent on implementing the policy. In other words,

a policy maker is concerned about the cost-effectiveness of these contracts. Following our conceptual framework, the cost effectiveness of contract $\hat{\delta}$ that aims to increase social welfare is given by the ratio of equations 16 and 14:

$$\underbrace{\frac{\Delta_D}{D(\hat{\delta})[\delta_r p_r + t_r] + D(\hat{\delta})P(m)[\delta_a p_a + t_{a|r}]}}_{\text{RDT uptake cost-effectiveness}} \underbrace{[P(m')c_a - c_r + (d_w + d_{sw})P(m')]}_{\text{Social welfare weight}} \tag{19}$$

Note that the cost-effectiveness of these contracts given by equation 19 can be decomposed into two parts. The first component is a standard cost-effectiveness measure for a policy that aims to influence the uptake of a product: the per-dollar impact of the policy on the demand for testing. The second component is a social welfare weight, which relates to whether increasing demand for testing is socially desirable.

Our modified cost-effectiveness measure (relative to the standard measure in the health policy literature) highlights the fact that a policy maker aiming to increase the uptake of testing could be overestimating or underestimating the cost-effectiveness of a program from a social welfare perspective. For example, the policy could inadvertently reduce social welfare if testing is too costly relative to its social value. Alternatively, the policy could be generating social gains far and beyond the standard measure due to the prevention of unnecessary expenditures and the reduction of externalities from unnecessary treatments. In what follows we assume that the condition in proposition 1 holds, implying that increasing the demand for testing is socially desirable.

With our cost-effectiveness measure in hand, we now turn to the question of which contract characteristics increase cost-effectiveness. In particular, should the policy maker target incentives to test directly or should they target incentives to treat for patients with a positive test result? Note that since contract characteristics δ only appear in the first component of our cost-effectiveness measure, we can focus our attention on which contract characteristics increase demand for testing the most per dollar amount.

We find that diagnosis-contingent ACT price-reductions can be more cost-effective than

RDT discounts under some circumstances. To see why this is the case, consider the marginal effect on patient demand from changes to discounted prices for ACTs and RDTs, holding the market price constant. For exposition purposes, assume that all patients have the same perceived probability of being malaria positive $P(m|i) = A$. This leads to:

$$\frac{\partial D(\delta)}{\partial p_r^*} = -D'(\delta) \tag{20}$$

$$\frac{\partial D(\delta)}{\partial p_{a|r}^*} = -P(m|i)D'(\delta) = -A * D'(\delta) \tag{21}$$

As observed in equations 20 and 21, patients are more responsive to changes in the price of the RDT than changes in the price of the ACT by a factor of $1/A \in [1, \text{inf})$. However, the real cost of the program is influenced by the true probability that a given patient is malaria positive — $P(m)$. For an RDT discount, the per-patient program cost will be on expectation $(p_r - p_r^*)D(p_a, p_r, p_r^*, p_{a|r}^*)$, while for a conditional ACT discount the expected cost per patient will be $P(m)(p_a - p_{a|r}^*)D(p_a, p_r, p_r^*, p_{a|r}^*)$. This is because the program only pays for the ACT subsidy when a patient is malaria positive.

Hence, if the true probability of being malaria positive is sufficiently low relative to the patients' beliefs, conditional ACT subsidies can be more cost-effective. To see why, note that as the belief $A \to 1$, the marginal effect of a conditional ACT discount approaches the marginal effect of RDT discounts. However, as $P(m) \to 0$, the expected per patient cost of the conditional ACT discount will approach zero.

3 Experimental Design

The study randomized 140 pharmacies into 4 groups – 3 intervention groups and a control group. The three treatment arms are (Appendix Table 1):

1. Patient subsidy group (T1): Clients who seek care for suspected malaria cases at these

pharmacies pay a subsidized price for RDTs (90% subsidy, a 10 Kes copay) and a subsidized price for ACTs (80% subsidy, a 30 Kes copay) conditional on a confirmed positive malaria diagnosis. The prices are advertised in large posters in prominent spots in the pharmacy.

2. Pharmacy incentive group (T2): Pharmacy owners receive an incentive to sell RDTs (90 Kes), and an additional incentive to prescribe ACTs to malaria-positive patients (80 Kes). Pharmacy attendants receive a 30 Kes incentive for recording transaction information in the malaria case management platform and completing the sale of incentivized products. Pharmacies are free to set prices charged to patients.

3. Combined group (T3): Clients are eligible for discounted rapid tests (60% subsidy, a 40 Kes copay) and discounted ACTs conditional on a positive test result (60% subsidy, a 60 Kes copay). Pharmacy owners receive an incentive to sell rapid tests (15 Kes), and an additional incentive to prescribe ACTs to malaria-positive patients (15 Kes). Pharmacy attendants receive a 30 Kes incentive for recording transaction information in the malaria case management platform and completing the sale of incentivized products. Pharmacies are free to set prices charged to patients.

The total value of the incentive was held fixed at 200 Kes (~$2 USD in 2021 exchange rates) across all treatment arms.[14]

The pharmacies participating in the study are existing users of Maisha Meds's digital sales management platform. Maisha Meds is a Kisumu-based healthcare social enterprise that provides sales and inventory management support to small pharmacies and clinics throughout Kenya. The platform records all pharmacy transactions and product stock. The incentive interventions were integrated into Maisha Meds's digital platform and managed centrally by the Maisha team. Subsidy and incentive amounts were automatically calculated based on

14. The incentive amount is consistent with prior literature, was determined after a pilot phase, and was calibrated to ensure pharmacy profitability would not be adversely affected, compared to the status quo.

the products that are being bought/sold and verified by implementation staff independent of the pharmacies prior to disbursement to ensure implementation fidelity.

Pharmacy staff received training on the importance of diagnostic testing (all arms), proper RDT administration, and use of the malaria case management tool. Stocks of RDTs and ACTs were provided on consignment through Maisha Meds in the intervention arms, while in the control group they managed their own stock.

3.1 Sample Enrollment

The sample consists of for-profit pharmacies and the clients that present with malaria symptoms located in the thirteen counties in the malaria endemic and epidemic areas of Kenya's western regions. These pharmacies manage their own stock and sales of diagnostic tests and medications. They set their own prices and sell at market prices.

To be eligible to participate in the study, pharmacies needed to be part of the Maisha Meds network and active users of the Maisha Meds digital sales and inventory management platform. Additionally, they had to be licensed pharmacies that were registered with Kenya's Pharmacy and Poisons Board. They also had to be willing to be randomized to one of the study arms, manage their sales through the digital tool, and to offer incentives (either supply- or demand-side) for malaria testing and treatment if assigned to one of the intervention arms.

All eligible pharmacies were mapped. Those located at least 0.5 km from other potential study participants were invited to participate.[15] Using these criteria 175 pharmacies were identify as eligible and were invited to participate in the study, of which 140 accepted.[16] These 140 pharmacies were randomly assigned to one of the four arms in waves, stratified on average monthly malaria product sales volumes (above/below median), urban/rural, and location of pharmacy in lake endemic county. Figure B1 shows the geographic span of the experiment across the target regions in Kenya and the final selection of pharmacies.

15. The average distance between study sites is 6.24 km (range of 0.5 km to 46.2 km).

16. Appendix Table B4 reports balance on baseline variables between pharmacies accepted (in sample) and those that declined (refusals). Facilities that declined to participate had been using the digital sales platform for longer than facilities in the sample frame. No other meaningful imbalances were found.

3.2 Data

See Appendix Table B1 for study timeline and a description of the primary sources of data. The study was initially planned to begin in June 2020, but was delayed due to COVID-19. The pharmacy onboarding, patient exit survey, standardized patient visits, and control group testing activities were all done in person following appropriate COVID-19 precautions.[17] The pharmacy baseline surveys were conducted over the phone.

We use the following data sources for analysis:

1. *Baseline data:*

    (a) Pharmacy owner survey: survey about number of staff, pharmacy business operations, patient volumes, pharmacy characteristics, costs and revenues, and knowledge of malaria case management.

    (b) Pharmacy staff survey: survey about malaria case management knowledge, worker motivation, and use of the digital platform used to manage sales and inventory.

2. *Administrative data:*

    (a) Sales data: continuously collected transaction data including prices and quantities of products purchased, location, date, and time of sale, and pharmacy staff who made the sale for over 50,000 malaria-related patient encounters between June 2021 - February 2022.[18]

    (b) Malaria case management data: continuously collected transaction data on all rapid test and treatment purchases made through incentive program, including information on age/gender of patient, rapid test result, prices and quantities of

17. The research and implementation teams followed Kenyan and UC Berkeley CPHS guidelines for conducting research while keeping study staff, implementation staff, and study subjects safe from COVID-19. All personnel and pharmacy staff were required to wear masks, maintain 1 meter distance from each other, and sanitize hands frequently. The research and implementation teams provided adequate PPE and hand sanitizer for all study and implementation personnel. Pharmacies were required by the Kenyan government to have all staff wearing masks, and have hand washing stations for staff and pharmacy clients, and pharmacies in our sample were compliant with these requirements during the study period.

18. Prices observed in the data are retail prices set by pharmacists in the digital tool.

medications purchased, location, date, and time of sale. Over 8,000 malaria transactions logged between June 2021 - February 2022.

3. *Standardized Patient Survey:* We employ standardized mystery patients (SP) to measure the appropriateness of the care delivered using the same clinical case scenario. We trained individuals (SPs) to present an identical standardized illness case scenario as real walk-in clients to providers. During encounters with providers, SPs portrayed real patients presenting a standardized, pre-scripted acute adult malaria case. The SPs were confirmed to be malaria-negative based on malaria microscopy tests administered by a reliable, high-quality laboratory before and after the month of field work. SPs and field work supervisors also monitored any potential symptoms throughout field work; all were otherwise healthy. By using trained SPs portraying the same illness case to generate the care data, we avoid bias from selection on patient illness type and severity that is inherent in care data collected using other common methods such as patient exit interviews, direct clinical observation, or health records (Peabody et al. 2000).

4. *Patient exit survey data:* survey with a random sample of 1654 eligible adult pharmacy clients across all study sites (12.6 clients/site).[19] This survey includes information on quality of care, symptoms, prices and quantities of medications and diagnostic tests purchased, beliefs about their illness status, malaria test result if applicable, and basic demographics.

5. *Testing subsample data:* data on test positivity from testing of random subset of 230 pharmacy clients at control group sites to obtain test positivity rate in a sample unaffected by the interventions (8.5 clients tested/site, 28 sites participated). Additional test positivity data from administrative records from 10 control group pharmacies that

19. In order to be eligible, clients must have sought care for malaria symptoms for themselves or a family member present at the pharmacy with them. Trained research staff visited each study pharmacy during an unannounced 5 day period, and screened all patients who exhibited malaria-related symptoms or purchased malaria products for eligibility. There were 1674 possible respondents screened, and 1654 respondents who completed the survey.

kept records of tests conducted (N=2547) on-site between January-February 2022.

6. *Endline data:*

   (a) Pharmacy owner survey: survey about number of staff, pharmacy business operations, patient volumes, pharmacy characteristics, costs and revenues, and altruistic tendencies.

   (b) Pharmacy staff survey: survey on malaria case management knowledge, worker motivation, use/familiarity with the digital platform used to manage sales and inventory and manage malaria cases, and altruistic tendencies.

3.3 Estimation Methods

All primary analyses are conducted at the patient level.[20] For all binary outcomes, we report marginal treatment effects from adjusted logistic regression models using the following specification:

$$Pr(Y_{ip}) = expit(\beta_0 + \beta_1 T_{1ip} + \beta_2 T_{2ip} + \beta_3 T_{3ip} + \lambda_s + \boldsymbol{X}_p + \epsilon_{ip}) \tag{22}$$

where Y_{ip} is a malaria testing or treatment outcome, T_{jip} are treatment assignment indicators for each intervention j for individual i seeking care at pharmacy p, with the control group as the reference category, λ_s are strata fixed effects, and ϵ_{ip} is the error term. We include variables that had significant imbalance with the control group at the $\leq 10\%$ level at baseline (Tables B2 and B3) as covariates in this adjusted model ($\boldsymbol{X_p}$), as specified in the pre-analysis plan. The β terms represent the log-odds of the treatment effect of each intervention relative to the control group, as percentage point changes. However, we report all results in terms of marginal treatment effects and p-values from Wald tests comparing the marginal treatment effect coefficients of the interventions to each other. Results of unadjusted models (excluding

20. The analyses specified in this section were pre-registered in a pre-analysis plan (AEARCTR-0004705). We discuss any deviations from the pre-analysis plan where relevant.

$\boldsymbol{X_p}$) are consistent with findings from the adjusted models, and can be made available upon request.

4 Sample Balance Across Study Arms

Tables B2 and B3 report the experimental balance checks at baseline, and shows that randomization was balanced across a large set of pre-specified covariates. Out of 84 tests conducted, 5 are significant at the $\leq$ 10 percent level. When we conduct a joint test for orthogonality using a multinomial logit model with treatment assignment as the categorical outcome, we find that the χ^2-test produces a p-value of 0.46. This suggests that these covariates are not jointly predictive of group assignment. In the adjusted models, we control for covariates that were unbalanced at baseline from comparisons with the control group.

5 Results

5.1 RDT Use

Table 2 (Columns 1 & 2) report the estimated effects of the incentive interventions on RDT use. As a reference point, only 8% of patients who sought care for malaria-related symptoms in control group pharmacies purchased a rapid diagnostic test prior to obtaining treatment, which is consistent with trends found across the full pharmacy sample prior to the start of the experiment (Appendix Figure B2) as well as with other existing research on rapid diagnostic test use in pharmacy settings across East Africa.[21]

The contracts increased RDT use substantially, consistent with the comparative statistics derived in sections 2.2 and 2.3. Patients who sought care in treatment pharmacies across all intervention arms were 25 percentage points more likely to purchase a diagnostic test (column 1, Table 2). Looking at each incentive arm separately in column 2 we find

21. For example, in Cohen et al. 2013; Cohen, Dupas, and Schaner 2015; O'Meara et al. 2016; Ansah et al. 2010.

large and statistically significant effects in all three arms; patient discounts resulted in a 27 percentage point increase in RDT uptake, while both the pharmacy incentives and the combination of patient discounts and pharmacy incentives resulted in a 20 percentage point increase. However, the differences across arms were not statistically significant from each other. Pharmacy-administered incentives to either patients or providers lead to more people being tested for malaria prior to receiving treatment.

5.2 ACT Use

Table 2 columns 3 & 4 present results on the impact of incentives on overall ACT use. Despite the very limited diagnostic tests sold in the control group, the vast majority of control group patients who sought care for suspected malaria purchased ACTs (87%) or another form of treatment (11%).[22] Based on an underlying malaria prevalence rate of 10% derived from the random testing exercise done in the control group, 90% of these individuals purchase ACTs unnecessarily, suggesting potentially large levels of medication waste.

In general, we find that the incentives caused a statistically significant decrease of 14 percentage points in ACT purchase (column 3), and between 9-15 percentage point decline when looking at each incentive intervention separately (column 4). Again, the three arms are not statistically distinguishable from each other. This result is consistent with the model implications. Since contracts are effective at increasing testing rates, more patients learn their true malaria status. Given that malaria suspect patients default to purchasing antimalarials absent a test (94% of users)[23], this leads to an overall decrease in demand for treatment. This is consistent with the demand function for treatment presented in section 2.2, and follows from the empirical observation that patients tend to follow the test results and true positivity rates are low conditional on seeking treatment.

22. Appendix Table B6 shows (zero) impact on non-ACT antimalarial medication sales.
23. Control group mean from Appendix Table B6, columns 5 & 6

5.3 Mechanisms

As exemplified in the model, there are two potential mechanisms through which diagnosis contingent contracts could impact demand for testing: changes in prices to the patient, or changes in the counseling of the provider. We investigate mechanisms using data from an audit study that uses standardized patients (SP) to measure the content of the care visit using the same clinical case scenario, with results in Table 3.[24] SPs have an advantage over client exit surveys or administrative transaction data in our setting as they avoid bias from selection on patient characteristics and malaria status (real or perceived). The SP data, by capturing data on an unselected patient sample - where the only variation is by the experimental design - allows us for cleaner identification of mechanisms.

We trained individuals (SPs) to present an identical standardized illness case scenario as real walk-in clients to providers. SPs followed a uniform script for how to present a suspected malaria case in a pharmacy setting: SPs were instructed to complain of fever, headache and joint pains in their opening statement and then provided additional information about their illness episode and health history if the pharmacist followed up with additional questions. SPs conducted a total of 411 visits across 137 facilities in the study sample, with three different SPs visiting each facility. SP visits provided a unique opportunity to assess the implementation fidelity and quality of care of the patient-provider interaction at study pharmacies.

First, we report intervention effects on patient prices, as reported in SP exit surveys (Table 3, columns 1-2). We find a 18% price reduction for RDT, in all intervention arms (column 1). In column 2, we see that this price pass-through is driven by the patient discount arm, but is not reflected in the supply-side incentive arms. The discount was passed through as a 43% price reduction for patients when the incentive was administered as

24. SPs have been used to measure quality of care extensively. For example see: Peabody et al. 2000; Das et al. 2012; Das et al. 2016; Mohanan et al. 2015; Kwan et al. 2018; Kwan et al. 2019; Kwan 2022; Das et al. 2022; Boone et al. 2023.

a consumer subsidy (which implies a price elasticity of demand of 7.86).[25] This suggests that in the patient discount arm, the increase in testing uptake and improvements in treatment targeting can be explained by reduced patient prices on rapid diagnostic tests and conditional discounts on treatment.

But this price mechanism does not appear to explain why we find similar effects on testing and treatment decisions in the two supply-side incentive arms. We test an alternative mechanism: provider counseling. Columns 3-5 of Table 3 present results on pharmacist advice and counseling behavior, using data from SP exit surveys. Specifically, we analyze SP reports of whether the pharmacist comprehensively explained their test result and treatment regimens, a measure of quality of counseling. We find that only 31% of the SPs at the control group sites report receiving comprehensive information on tests and / or treatment options, and the incentives significantly improve this (11 percentage points, on average, from column 3). When looking at patient and provider incentives separately, we see that the improvements in counseling are driven entirely by the provider-side incentives (columns 4 and 5). When pharmacists are incentivized directly, they are 15 percentage points more likely than control group pharmacists to clearly explain treatment options to SPs (pooling the provider arms, as in column 4, and looking at them separately, as in column 5). This suggests that when incentivized directly, pharmacists change their behavior and provide more comprehensive counseling on testing and treatment options to suspected malaria patients.

Taken together, these results suggest that price-pass through is likely to explain the demand-side treatment effects, and the information/counseling channel is likely to explain the supply-side treatment effects we find in Table 2.

6 Welfare

In this section, we estimate the welfare impact of introducing diagnosis contingent contracts using the framework presented at the beginning of the paper. The measure derived in equa-

25. Using 338% change in quantity, calculated from point estimates in Table 2

tion 16 quantifies the impact on social welfare per unit of cost for the social planner. Before proceeding to the welfare analysis, we will present evidence that validates key assumptions made in our conceptual framework.

6.1 Validation of model assumptions

In this section we discuss the assumptions made in the model in order to derive our measure of social welfare cost-effectiveness. We first assume that under uncertainty a malaria-suspect patient's value of being treated is higher than forgoing treatment. That is, $\mathbb{E}(U(\text{ACT})) > \mathbb{E}(U(\text{No ACT}))$, and therefore untested patients will choose to purchase treatment. This assumption is motivated by the fact that observed patients in the study pharmacies self-select into seeking treatment. A deviation from this assumption would be if we found evidence that the contracts increase the volume of patients seeking malaria care. As seen in column 1 of table B6, the share of patients seeking malaria care in the pharmacies remains unchanged in treatment groups relative to the control group. Therefore, the main decision being made by patients seeking malaria care in study pharmacies is whether to test or not.

We also assume that patients who decide to take a test will follow the test result when deciding whether to treat. This assumption holds in our data, where more than 95% of patients who test positive for malaria purchase an ACT, and only 5% of patients who test negative decide to purchase a treatment anyways. This finding is in contrast to the results from Cohen, Dupas, and Schaner 2015, where over half of patients testing negative elect to receive treatment. There are a couple of differences between our studies that might explain these different patterns. First, our study introduces conditional ACT incentives that tie subsidy access to the test result, minimizing incentives for moral hazard from patients offered heavily discounted treatment even if negative. Second, our study took place a decade after their study, so patient trust for test results might have been influenced through learning and the widespread availability of RDTs after the COVID-19 pandemic. Third, our study's population are patients that self-selected into seeking malaria care in pharmacies, whereas

patients in the Cohen, Dupas, and Schaner 2015 study were recruited in their homes, which generates higher overall demand for treatment relative to their control group.

6.2 Are diagnosis contingent contracts welfare enhancing?

Within the model's framework, whether preventing unnecessary treatment by encouraging higher testing is socially optimal depends on the relative costs of testing and treatment, the probability of being malaria positive, and the internalities and externalities from unnecessary treatment (as given by the condition from theorem 1). Diagnosis contingent incentive contracts have the potential of increasing total social welfare because it is possible for patients to make suboptimal decisions (from either a social or individual perspective) if they fail to internalize the externalities and internalities of unnecessary treatment, or if they hold incorrect beliefs about their malaria status.

We find evidence that the condition in proposition 16 holds in our setting, suggesting that diagnosis-contingent contracts are expected to increase social welfare. While we do not observe d_w and d_{sw}, we do observe proxies for costs of testing and treatment, as well as the positivity rate P(m). For a given patient in our data, the cost of treatment c_a is on average 160KES. On the other hand, the cost of testing is 60KES. The probability of a patient being malaria negative is $P(m') = 0.90$. Therefore, regardless of the value of $d_w > 0$ and $d_{sw} > 0$, these contracts are expected to increase welfare in our setting ($P(m') > 0.37$). Moreover, even if the costs of the product are unobserved and only the prices to the patient are directly observed, the equation in 17 implies a stronger test for whether shifting demand for testing is welfare increasing[26]. One can use the price to the patient of each product category instead of the costs. The average expenditures on treatment by patients is roughly 265KES, while the mean expenditure on tests is 120KES. Therefore, even with this stronger test one would expect total welfare to increase if contracts that

26. This condition assumes that by revealed choice, pharmacy profits stay constant or increase from contracts. We find evidence of this being the case, where using administrative transaction data pharmacy profits appear to increase in all treated arms relative to the control group.

encourage testing are introduced ($P(m') > 0.45$).

6.3 Social welfare effects

In this section we estimate the per-unit-of-cost effect of these interventions on social welfare using the measure derived in equation 19. Table B7 presents the inputs used for estimation. To estimate average per patient costs of the program, we assume that marginal operational costs of these incentive contracts are negligible, allowing us to focus on the direct costs of implementing the price reductions and incentives. The contracts are implemented through a digital system, so the main operational cost are stock deliveries to these pharmacies. With economies of scale, the marginal cost of delivery at the patient level is likely negligible relative to the cost of the product itself when the contracts take advantage of existing supply-chain infrastructure, such as in our context. Moreover, we use administrative data on marginal costs of treatment and testing for pharmacies and the positivity rate from the random testing to estimate the welfare weight used to transform the cost-effectiveness measure into social welfare effects.

For ease of interpretation, we report the increase in patient and provider welfare per unit of cost for the social planner, where a value above 1 implies a positive rate of return from the intervention. Table 4 displays the results. Column 3 shows the lower bound on the social welfare weight implied by the average cost (of holding stock) and the mean unconditional positivity rates of patients in the sample. For each patient that shifts demand towards testing, we expect a social welfare increase of \$0.85 on expectation (\$1.85 for patients and pharmacies accounting for transfers). Column 4 shows the cost-effectiveness of each treatment arm in terms of RDT demand increase. With patient subsidies, 0.96 additional patients purchase an RDT per dollar cost for the policy maker. Pharmacy incentives are slightly less cost-effective, with only a 0.68 patient increase per dollar. Finally, the combined treatment leads to a 0.59 increase in patients purchasing RDTs per dollar.

Column 5 shows the cost-effectiveness in terms of patient and provider welfare gains. The

patient subsidy arm results in a $1.81 increase on the welfare of patients and pharmacies per dollar spent by the policy maker, implying a rate of return of 84% . The other two arms are also cost-effective, with effects of $1.58 and $1.50 for the pharmacy incentive and the combined treatment arms respectively. The increase in welfare is the result of the prevention of unnecessary treatment costs due to increased testing. Note that it could have been possible for these interventions to reduce social welfare if the increase in testing had increased overall malaria treatment expenditures. Therefore, it is reassuring that in this context the social welfare weight is positive and that in this context shifting demand towards testing is socially desirable.

While the social welfare results presented here offer a lower bound on the effect of these interventions, it is likely that this lower bound is close to its true value. The marginal externality of a patient receiving unnecessary treatment is likely to be small in our setting given that the interventions were implemented in a small subset of the overall population. Similarly, internalities from unnecessary treatment such as patient's concerns about malaria resistance in their communities and side effects from unnecessary treatment are unlikely to be large relative to the changes in health expenditures.

7 Optimal contract design

Our framework, combined with the random variation from the experiment, allows us to study how contract design impacts its effectiveness in increasing social welfare. There are two margins that we consider: whether to target the patient or the provider's incentives, and whether to give direct incentives to test or diagnosis-contingent discounts on treatment.

In order to evaluate the relative effectiveness of each contract component we need to estimate how the demand for testing responds to them. Note that equation 6 along with the assumption that patients have an additive taste shock for testing implies that demand for RDTs can be represented as follows:

$$RDT_i = 1\{C(\theta) + \alpha(p_r - p_a) + \alpha_{a|r,\theta} p_{a|r} > \epsilon_i\} \tag{23}$$

Where $C(\theta) - \epsilon_i = \alpha P(m'|\theta, i) d_{w,i}(\theta)$, and $\alpha_{r|a,\theta}/\alpha = P(m|\theta, i)$. This representation highlights that patients care about the cost of testing relative to treatment $(p_r - p_a)$[27]. Parameters α and $\alpha_{a|r,\theta}$ relate to the demand response to direct RDT discounts and diagnosis-contingent discounts.

To answer the question of whether to target the patient or the provider's incentives, we need to recover the marginal effect of the provider incentives on test demand. Note that in the model this effect operates through changes in advice and information θ. Therefore, to recover the marginal effect of provider incentives one needs to use a proxy for the provider's incentives to recommend. As discussed in appendix 8.1, function $g(\pi, t) = (\pi_r + t_r) + P(m)(\pi_{a|r} + t_{a|r})$ is an appropriate proxy for the provider's financial incentives[28].

Therefore, for a patient-provider unit, one can recover estimates of $\frac{\partial D(.)}{\partial p_r}$, $\frac{\partial D(.)}{\partial p_{a|r}}$, and $\frac{\partial D(.)}{\partial g}$ through the following linear-probability equation:

$$P(RDT_i) = \beta_0 + \beta_1 g(\pi, t) + \alpha(p_r - p_a) + \alpha_{a|r} p_{a|r} + e_i \tag{24}$$

To identify causal estimates of the parameters of interest we need exogenous variation on $g(\pi, t)$, $p_r - p_a$ and $p_{a|r}$. Note that the experiment generates exogenous variation on provider's incentives for tests $g(\pi, t)$, testing price p_r, and diagnosis conditional ACT price $p_{a|r}$ by randomizing access to patient and provider incentives. However, the experiment does not generate independent variation between p_r and $p_{a|r}$ since all treatment arms include direct RDT incentives and diagnosis contingent incentives for ACTs. Nevertheless, in the

27. This follows from the model and not the functional form assumption for the error term.

28. One can show that $g(\pi, t)$ proxies for $f(p, c, t)$ when θ is unobserved and one uses the experimental variation as instruments. See appendix 8.1.

administrative data we observe the marginal costs to the pharmacy for treatment outside of the loyalty program, allowing us to instrument for p_a. Therefore, by instrumenting for $g(\pi, t)$, $(p_r - p_a)$, and $p_{a|r}$ through treatment assignment indicators and marginal costs for malaria treatment, one can recover causal estimates of the three marginal effects of interest.

7.1 Empirical model results

Table 5 shows the estimates from equation 24. Column 1 shows the main estimates, using treatment assignment indicators and marginal costs of treatment as instruments for $(p_r - p_a)$, $p_{a|r}$, and $g(\pi, t)$. Demand for tests responds to the three contract characteristics. First, strikingly, demand for tests has a very similar response to both direct RDT discounts and diagnosis contingent ACT discounts, implying that patients have very upwardly biased beliefs about their malaria status. We find that $\alpha = -0.086$ and $\alpha_{a|r} = -0.085$, both significant at the 99% level. This suggests that the average patient's beliefs that they are malaria positive are $\mathbb{E}[P(m|i)] = 0.085/0.086 = 98\%$, highly overestimating their probability of having the disease relative to the true probability (10%). If patients had accurate beliefs about their malaria status, demand for testing should only respond to diagnosis contingent incentives by 90% of the price elasticity[29]. This suggests that diagnosis contingent incentives are likely to be very cost-effective since program costs are a function of the true probability, rather than the patient's biased beliefs.

Demand for tests also responds to the financial incentives of the provider, conditional on price. The coefficient for $g(\pi, t)$ is equal to 0.22, suggesting that on the margin, demand responds strongly to the financial incentives of the provider. This result is consistent with the reduced form analysis in table 2, where demand for testing in treatment 2 increased absent any discounts to the patient. As observed in the mechanisms results in table 3, this seems to be driven by changes in the information shared by the provider to the patient.

29. Note that in our stylized model we do not include patient risk-aversion. Risk aversion could be an alternative channel contributing to the strong response to diagnosis-contingent contracts, although patients would need to be extremely risk-averse to explain this pattern in the absence of unbiased beliefs.

7.2 Counterfactual contract design

Now we turn to the question of which contract design is more effective at increasing social welfare. Table 6 shows the implied welfare effects from counterfactual contracts using the parameters estimated in Table 5. In our counterfactual contracts we fix the maximum payment per patient from the social planner to $2USD, but vary whether the incentives are loaded on direct RDT discounts (incentives), diagnosis-contingent ACT discounts (incentives), patient discounts, or provider incentives. We use the fitted values from the linear probability model estimated in Table 5 to obtain the predicted treatment effects on demand from these contracts.

First, in Panel A we compare three patient contracts: a contract that loads the incentive on a RDT discount, a contract with a diagnosis-contingent ACT discount, and the experimental contract from T1. Column 1 shows that the three contracts have similar expected effects on testing demand of 17 p.p., within the confidence interval of the estimates in table 2. This is because we estimated α and $\alpha_{a|r}$ to have almost the same value. However, expected costs per patient vary widely between the three types of contracts, ranging from $0.75 to $0.08 per patient on expectation. This is because RDT discounts are paid every time a patient decides to purchase a test, significantly increasing costs relative to the experimental design. On the other hand, contracts with only diagnosis-contingent discounts incur a cost only when the patient purchases a test and is malaria positive.

We find that diagnosis-contingent contracts are very cost effective relative to direct RDT discounts. Columns 3 and 5 show the cost-effectiveness and welfare estimates. In terms of the traditional cost-effectiveness measure, for each dollar of program cost the RDT discounts increase testing by 0.23 additional patients, while diagnosis-contingent ACT discounts increases demand by 2.27 patients per dollar. The experimental contract, which combines both design components, has a cost-effectiveness of 0.45 additional patients who test per dollar. These patterns translate to similar welfare gains, where diagnosis contingent contracts increase patient and pharmacy welfare by $2.93 for each dollar spent by the policy maker on

the contracts. Direct RDT subsidies and the experimental contract are significantly less effective, with a welfare effect of $1.20 and $1.39 respectively.

Analogously, in Panel B we compare three contract designs: one that loads provider incentives on RDT sales, one that loads incentives on diagnosis -contingent ACT sales, and the experimental provider contract from T2. A key difference from the patient's contract counterfactuals is that we assume that providers have unbiased beliefs about the probability that the patient has malaria. Nevertheless, it is still ambiguous whether diagnosis-contingent contracts are more cost-effective than direct RDT discounts since this will be determined by the patient's response to the provider's recommendation (β_1) relative to the contract costs.

Patient demand for testing varies significantly depending on the provider's incentive design. RDT incentives increase demand by 44 p.p., while diagnosis-contingent contracts only increase demand for testing by 4 p.p. – since providers understand the low probability that these incentives materialize. The experimental contract has an effect of 28 p.p., well within the confidence interval of the effect estimated directly in Table 2. The program costs also vary meaningfully across contract designs. RDT incentives are very costly on expectation relative to alternative designs ($1.28 per malaria-suspect patient case) since they are paid whenever a patient tests and the effect on demand from these incentives is large. Diagnosis-contingent contracts, on the other hand, are low-cost on expectation ($0.05 per malaria-suspect patient case) since the demand response is small and the incentives are only paid when a malaria positive patient tests. We find that the diagnosis-contingent provider contract is more cost-effective than the RDT incentive contract (0.90 versus 0.34 additional tests sold per dollar spent by the policy-maker). Similarly, diagnosis contingent contracts have a larger effect on social welfare, increasing patient and pharmacy welfare by $1.76 per unit of cost relative to $1.29 for the RDT discount.

Therefore, regardless of whether the policy targets the patient or the provider, diagnosis-contingent contracts are more cost-effective. Overall, the most cost-effective contract design are diagnosis-contingent discounts for the patient, since patients hold very biased beliefs

about their malaria status.

8 Discussion

This paper examines the effects of a novel diagnosis-contingent contract structure to improve malaria case management in a cluster-randomized control trial in Kenya. The experimental treatments provided financial incentives to patients, pharmacists, or both for RDTs and ACTs conditional on testing positive for malaria and were implemented in private sector pharmacies in thirteen malaria-prone counties. By tying financial incentives for treatment to diagnostic outcome, we propose a flexible innovation in how payment contracts for health services could be structured to emphasize quality of care rather than service volume.

This paper contributes to the literature on performance-based financing mechanisms by examining the behavioral channels through which provider incentives impact healthcare quality. It also innovates in health financing contracts, proposing differential payment structures based on diagnostic information. Additionally, it explores how combination of personalized health information and financial incentives influence health behavior. And finally, it adds to our understanding of how incentives targeted at the demand-side or the supply-side can affect decision-making.

We find encouraging results of the demand- and supply-side incentives on both testing and treatment targeting. Overall, the incentives interventions increased RDT use substantially in a setting with very low baseline testing levels. On average, patients who sought care in treatment pharmacies were 25 percentage points more likely to receive a formal malaria diagnosis prior to purchasing treatment for suspected malaria. This represents a more than 300% increase over the control group. Incentive interventions were also effective encouraging appropriate use of antimalarials. We find an overall 14 percentage point decrease in the use of ACTs as a result of the treatment, and this is due to malaria negative patients opting out of purchasing unnecessary antimalarials. For patients who test positive, we find that

they are appropriately nudged to take ACTs, consistent with their diagnostic test result. Interestingly, we find statistically indistinguishable effects of the demand-side and supply-side treatment arms, suggesting that incentives yield similar outcomes whether they are provided directly to patients or they are provided to pharmacists.

We explore mechanisms through which the incentive interventions worked in order to contextualize the main findings. We find that the patient subsidies for RDTs resulted in significantly lower prices being paid by patients (43% reduction in price). However, we find no evidence of pass-through of the RDT incentive in either of the two supply-side arms, and no evidence of price pass-through on ACT prices in any of the three treatment arms. Instead, we find evidence that in the supply-side incentive arms, pharmacists explained diagnosis and treatment options more comprehensively to their patients. Improved, individualized health information appears to be the channel through which the supply-side incentives resulted in the overall changes in RDT and ACT use seen in the main results. In sum, the demand subsidies induced more patients to purchase RDTs that provided accurate illness status information, which led to more appropriate use of ACTs. And, the supply-side incentives led pharmacists to provide more detailed diagnosis counseling and treatment recommendations, yielding similar overall effects on malaria case management.

Finally, we find that the diagnosis contingent incentive contracts are cost-effective, largely due to the fact that they led to large reductions in malaria-negative patients taking unnecessary antimalarial drugs. Taken together, our results imply that diagnosis-contingent contracts may have the potential to reduce medical waste and curb spending while better targeting health care resources to areas of proven need.

References

Afendulis, Christopher C, and Daniel P Kessler. 2007. "Tradeoffs from integrating diagnosis and treatment in markets for health care." *American Economic Review* 97 (3): 1013–1020.

Ahmed, Tashrik, Aneesa Arur, Damien De Walque, and Gil Shapira. 2023. "Incentivizing quantity and quality of care: evidence from an impact evaluation of performance-based financing in the health sector in Tajikistan." *Economic Development and Cultural Change* 71 (2): 000–000.

Ansah, Evelyn K, Solomon Narh-Bana, Michael Epokor, Samson Akanpigbiam, Alberta Amu Quartey, John Gyapong, and Christopher JM Whitty. 2010. "Rapid testing for malaria in settings where microscopy is available and peripheral clinics where only presumptive treatment is available: a randomised controlled trial in Ghana." *Bmj* 340.

Arrow, Kenneth J. 1963. "Uncertainty and the welfare economics of medical care." *The American Economic Review* 53 (5): 941–973.

Baicker, Katherine, Sendhil Mullainathan, and Joshua Schwartzstein. 2015. "Behavioral hazard in health insurance." *The Quarterly Journal of Economics* 130 (4): 1623–1667.

Baker, Laurence C. 2010. "Acquisition of MRI equipment by doctors drives up imaging use and spending." *Health Affairs* 29 (12): 2252–2259.

Basinga, Paulin, Paul J Gertler, Agnes Binagwaho, Agnes LB Soucat, Jennifer Sturdy, and Christel MJ Vermeersch. 2011. "Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation" [in en]. *The Lancet* 377, no. 9775 (April): 1421–1428. ISSN: 0140-6736, accessed January 26, 2020. https://doi.org/10.1016/S0140-6736(11)60177-3. http://www.sciencedirect.com/science/article/pii/S0140673611601773.

Boone, Claire E, Paul J Gertler, Grace Makana Barasa, Joshua Gruber, and Ada Kwan. 2023. "Can a private sector engagement intervention that prioritizes pro-poor strategies improve healthcare access and quality? A randomized field experiment in Kenya." *Health Policy and Planning* 38 (9): 1006–1016.

Burton, Deron C, Brendan Flannery, Bernard Onyango, Charles Larson, Jane Alaii, Xingyou Zhang, Mary J Hamel, Robert F Breiman, and Daniel R Feikin. 2011. "Healthcare-seeking behaviour for common infectious disease-related illnesses in rural Kenya: a community-based house-to-house survey." *Journal of health, population, and nutrition* 29 (1): 61.

Busse, Meghan, Jorge Silva-Risso, and Florian Zettelmeyer. 2006. "$1,000 cash back: The pass-through of auto manufacturer promotions." *American Economic Review* 96 (4): 1253–1270.

Campbell, Stephen M, David Reeves, Evangelos Kontopantelis, Bonnie Sibbald, and Martin Roland. 2009. "Effects of pay for performance on the quality of primary care in England." *New England Journal of Medicine* 361 (4): 368–378.

Chandra, Amitabh, David Cutler, and Zirui Song. 2011. "Who ordered that? The economics of treatment choices in medical care." *Handbook of health economics* 2:397–432.

Chen, Brian K, Paul J Gertler, and Chun-Yuh Yang. 2016. "Physician ownership of complementary medical services." *Journal of Public Economics* 144:27–39.

Cohen, Jessica, Pascaline Dupas, and Simone Schaner. 2015. "Price Subsidies, Diagnostic Tests, and Targeting of Malaria Treatment: Evidence from a Randomized Controlled Trial" [in en]. *American Economic Review* 105, no. 2 (February): 609–645. ISSN: 0002-8282, accessed November 21, 2019. https://doi.org/10.1257/aer.20130267. https://www.aeaweb.org/articles?id=10.1257/aer.20130267.

Cohen, Jessica L, Prashant Yadav, Corrina Moucheraud, Sarah Alphs, Peter S Larson, Jean Arkedis, Julius Massaga, and Oliver Sabot. 2013. "Do price subsidies on artemisinin combination therapy for malaria increase household use?: evidence from a repeated cross-sectional study in remote regions of Tanzania." *PloS one* 8 (7).

Currie, Janet, Wanchuan Lin, and Juanjuan Meng. 2014. "Addressing antibiotic abuse in China: an experimental audit study." *Journal of development economics* 110:39–51. ISSN: 0304-3878.

Currie, Janet, Wanchuan Lin, and Wei Zhang. 2011. "Patient knowledge and antibiotic abuse: Evidence from an audit study in China." *Journal of health economics* 30 (5): 933–949.

Cutler, David M, and Richard J Zeckhauser. 2000. "The anatomy of health insurance." In *Handbook of health economics,* 1:563–643. Elsevier.

Daniels, Benjamin, Ada Kwan, Madhukar Pai, and Jishnu Das. 2019. "Lessons on the quality of tuberculosis diagnosis from standardized patients in China, India, Kenya, and South Africa." *Journal of clinical tuberculosis and other mycobacterial diseases* 16:100109.

Das, Jishnu, and Jeffrey Hammer. 2014. "Quality of primary care in low-income countries: facts and economics." *Annu. Rev. Econ.* 6 (1): 525–553.

Das, Jishnu, Alaka Holla, Veena Das, Manoj Mohanan, Diana Tabak, and Brian Chan. 2012. "In urban and rural India, a standardized patient study showed low levels of provider training and huge quality gaps." *Health affairs* 31 (12): 2774–2784.

Das, Jishnu, Alaka Holla, Aakash Mohpal, and Karthik Muralidharan. 2016. "Quality and accountability in health care delivery: audit-study evidence from primary care in India." *American Economic Review* 106 (12): 3765–99.

Das, Veena, Benjamin Daniels, Ada Kwan, Vaibhav Saria, Ranendra Das, Madhukar Pai, and Jishnu Das. 2022. "Simulated patients and their reality: an inquiry into theory and method." *Social Science & Medicine* 300:114571.

Disease Control, Centers for, and Prevention. 2018. *CDC Malaria Support in Kenya.* https://www.cdc.gov/malaria/malaria_worldwide/cdc_activities/kenya.html.

Dow, William H, Justin S White, and Stefano Bertozzi. 2016. "Incentivizing health care utilization and health outcomes." In *World Scientific Handbook of Global Health Economics and Public Policy: Volume 1: The Economics of Health and Health Systems,* 421–451. World Scientific.

Dupas, P., and E. Miguel. 2017. "Chapter 1 - Impacts and Determinants of Health Levels in Low-Income Countries" [in en]. In *Handbook of Economic Field Experiments,* edited by Abhijit Vinayak Banerjee and Esther Duflo, 2:3–93. Handbook of Economic Field Experiments. North-Holland, January. Accessed January 26, 2020. https://doi.org/10.1016/bs.hefe.2016.09.003. http://www.sciencedirect.com/science/article/pii/S2214658X16300113.

Dupas, Pascaline. 2011. "Health behavior in developing countries." *Annu. Rev. Econ.* 3 (1): 425–449. ISSN: 1941-1383.

———. 2014. "Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment." Publisher: Wiley Online Library, *Econometrica* 82 (1): 197–228. ISSN: 0012-9682.

Fleming-Dutra, Katherine E, Adam L Hersh, Daniel J Shapiro, Monina Bartoces, Eva A Enns, Thomas M File, Jonathan A Finkelstein, Jeffrey S Gerber, David Y Hyun, Jeffrey A Linder, et al. 2016. "Prevalence of inappropriate antibiotic prescriptions among US ambulatory care visits, 2010-2011." *Jama* 315 (17): 1864–1873.

Gertler, Paul, Paula Giovagnoli, and Sebastian Martinez. 2014. *Rewarding Provider Performance to Enable a Healthy Start to Life: Evidence from Argentina's Plan Nacer.* Accessed November 25, 2019. https://openknowledge.worldbank.org/bitstream/handle/10986/18801/WPS6884.txt.

Gong, Erick. 2015. "HIV testing and risky sexual behaviour." *The Economic Journal* 125 (582): 32–60.

Iizuka, Toshiaki. 2012. "Physician agency and adoption of generic pharmaceuticals." *American Economic Review* 102 (6): 2826–58.

Initiative, U.S. Presidents Malaria. 2021. *U.S. Presidents Malaria Initiative Kenya Malaria Operational Plan FY 2020.*

King, Jessica, Timothy Powell-Jackson, James Hargreaves, Christina Makungu, and Catherine Goodman. 2022. "Pushy Patients Or Pushy Providers? Effect Of Patient Knowledge On Antibiotic Prescribing In Tanzania: Study examines the effect of patient knowledge on antibiotic prescribing in Tanzania." *Health Affairs* 41 (6): 911–920.

Kwan, Ada. 2022. *Implementing quality of care measures: lessons from a standardized patient study in seven provinces of China,* 6.

Kwan, Ada, Benjamin Daniels, Sofi Bergkvist, Veena Das, Madhukar Pai, and Jishnu Das. 2019. "Use of standardised patients for healthcare quality research in low- and middle-income countries." *BMJ global health* 4 (5): e001669.

Kwan, Ada, Benjamin Daniels, Vaibhav Saria, Srinath Satyanarayana, Ramnath Subbaraman, Andrew McDowell, Sofi Bergkvist, Ranendra K Das, Veena Das, Jishnu Das, et al. 2018. "Variations in the quality of tuberculosis care in urban India: a cross-sectional, standardized patient study in two cities." *PLoS medicine* 15 (9): e1002653.

Lopez, Carolina, Anja Sautmann, and Simone Schaner. 2022. "Does patient demand contribute to the overuse of prescription drugs?" *American Economic Journal: Applied Economics* 14 (1): 225–260.

Ma, Xiaochen, Zhongqiang Zhou, Hongmei Yi, Xiaopeng Pang, Yaojiang Shi, Qianyun Chen, Mirjam E Meltzer, Saskia Le Cessie, Mingguang He, and Scott Rozelle. 2014. "Effect of providing free glasses on childrens educational outcomes in China: cluster randomized controlled trial." *Bmj* 349:g5740. ISSN: 1756-1833.

Maffioli, Elisa M, Wendy OMeara, Elizabeth Turner, and Manoj Mohanan. 2019. "Can Individuals Beliefs Help Us Understand Non-Adherence to Malaria Test Results - Evidence from Rural Kenya." *Economic Research Initiatives at Duke (ERID) Working Paper,* no. 243.

Mbonye, Anthony K, Sham Lal, Bonnie Cundill, Kristian Schultz Hansen, Siân Clarke, and Pascal Magnussen. 2013. "Treatment of fevers prior to introducing rapid diagnostic tests for malaria in registered drug shops in Uganda." *Malaria journal* 12:1–10.

McGuire, Thomas G. 2000. "Physician agency." *Handbook of health economics* 1:461–536.

Meredith, Jennifer, Jonathan Robinson, Sarah Walker, and Bruce Wydick. 2013. "Keeping the doctor away: experimental evidence on investment in preventative health products." *Journal of Development Economics* 105:196–210. ISSN: 0304-3878.

Miller, Grant, and Kimberly Singer Babiarz. 2013. "Pay-for-performance incentives in low-and middle-income country health programs."

Mohanan, Manoj, Marcos Vera-Hernández, Veena Das, Soledad Giardili, Jeremy D Goldhaber-Fiebert, Tracy L Rabin, Sunil S Raj, Jeremy I Schwartz, and Aparna Seth. 2015. "The know-do gap in quality of health care for childhood diarrhea and pneumonia in rural India." *JAMA pediatrics* 169 (4): 349–357.

Musuva, Anne, Waqo Ejersa, Rebecca Kiptui, Dorothy Memusi, and Edward Abwao. 2017. "The malaria testing and treatment landscape in Kenya: results from a nationally representative survey among the public and private sector in 2016." *Malaria journal* 16 (1): 1–13.

O'Meara, Wendy Prudhomme, Diana Menya, Jeremiah Laktabai, Alyssa Platt, Indrani Saran, Elisa Maffioli, Joseph Kipkoech, Manoj Mohanan, and Elizabeth L. Turner. 2018. "Improving rational use of ACTs through diagnosis-dependent subsidies: Evidence from a cluster-randomized controlled trial in western Kenya" [in en]. *PLOS Medicine* 15, no. 7 (July): e1002607. ISSN: 1549-1676, accessed November 21, 2019. https://doi.org/10.1371/journal.pmed.1002607. https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002607.

O'Meara, Wendy Prudhomme, Manoj Mohanan, Jeremiah Laktabai, Adriane Lesser, Alyssa Platt, Elisa Maffioli, Elizabeth L. Turner, and Diana Menya. 2016. "Assessing the independent and combined effects of subsidies for antimalarials and rapid diagnostic testing on fever management decisions in the retail sector: results from a factorial randomised trial in western Kenya" [in en]. *BMJ Global Health* 1, no. 2 (September). ISSN: 2059-7908, accessed November 21, 2019. https://doi.org/10.1136/bmjgh-2016-000101. https://gh.bmj.com/content/1/2/e000101.

Pauly, Mark. 1980. "Doctors and Their Workshops: Economic Models of Physician Behavior." In *Doctors and Their Workshops: Economic Models of Physician Behavior,* 119–122. University of Chicago Press.

Peabody, John W, Jeff Luck, Peter Glassman, Timothy R Dresselhaus, and Martin Lee. 2000. "Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality." *Jama* 283 (13): 1715–1722.

Peabody, John W, Riti Shimkhada, Stella Quimbo, Orville Solon, Xylee Javier, and Charles McCulloch. 2014. "The impact of performance incentives on child health outcomes: results from a cluster randomized controlled trial in the Philippines." *Health Policy and Planning* 29 (5): 615–621.

Podgursky, Michael J, and Matthew G Springer. 2007. "Teacher performance pay: A review." *Journal of policy analysis and management* 26 (4): 909–949.

Prendergast, Canice. 1999. "The provision of incentives in firms." *Journal of economic literature* 37 (1): 7–63.

Stern, Ariel D, Brian M Alexander, and Amitabh Chandra. 2017. "How economics can shape precision medicines." *Science* 355 (6330): 1131–1133.

Sulis, Giorgia, Pierrick Adam, Vaidehi Nafade, Genevieve Gore, Benjamin Daniels, Amrita Daftary, Jishnu Das, Sumanth Gandra, and Madhukar Pai. 2020. "Antibiotic prescription practices in primary care in low-and middle-income countries: a systematic review and meta-analysis." *PLoS medicine* 17 (6): e1003139.

Visser, Theodoor, Jeremiah Laktabai, Emmah Kimachas, Joseph Kipkoech, Diana Menya, David Arthur, Yunji Zhou, Tabitha Chepkwony, Lucy Abel, Emily Robie, et al. 2024. "A cluster-randomized trial of client and provider directed financial interventions to align incentives with appropriate case management in private medicine retailers: results of the TESTsmART Trial in Lagos, Nigeria." *PLOS Global Public Health* 4 (7): e0002938.

Whitehead, Margaret, Göran Dahlgren, and Timothy Evans. 2001. "Equity and health sector reforms: can low-income countries escape the medical poverty trap?" *The Lancet* 358 (9284): 833–836.

WHO. 2021. *WHO Guidelines for malaria.* Licence: CC BY-NC-SA 3.0 IGO.

Witter, Sophie, Atle Fretheim, Flora L Kessy, and Anne Karin Lindahl. 2012. "Paying for performance to improve the delivery of health interventions in low-and middle-income countries." *Cochrane database of systematic reviews,* no. 2.

Yip, Winnie, Timothy Powell-Jackson, Wen Chen, Min Hu, Eduardo Fe, Mu Hu, Weiyan Jian, Ming Lu, Wei Han, and William C Hsiao. 2014. "Capitation combined with pay-for-performance improves antibiotic prescribing practices in rural China." *Health affairs* 33 (3): 502–510.

Tables

Table 1: Incentive amount details, by treatment arm (Back: 3)

Subsidy and incentive amounts

	Control (C)	**Patient discount (T1)**	**Pharmacy incentive (T2)**	**Both (T3)**
Patient discounts? (USD)				
Rapid test	-	$0.90	-	$0.60
ACT (malaria +)	-	$1.10	-	$0.80
ACT (malaria -)	-	$0.00	-	$0.00
Provider incentives (USD)				
Rapid test	-	-	$0.90	$0.20
ACT (malaria +)	-	-	$0.80	$0.10
ACT (malaria -)	-	-	$0.00	$0.00
Transaction completion	-	-	$0.30	$0.30
Total incentive amount (USD)	$0.00	$2.00	$2.00	$2.00

Table 2: Impact on rapid test and ACT uptake (Back: 5.1)

	Rapid test uptake		ACT uptake	
	(1)	(2)	(3)	(4)
Pooled treatment	.25**		-.139**	
	(0.051)		(0.049)	
Patient discount		.267*		-.145*
(γ_{T1})		(0.106)		(0.069)
Pharmacy incentive		.194**		-.0892^{+}
(γ_{T2})		(0.065)		(0.050)
Patient discount and		.201**		-.136**
pharmacy incentive (γ_{T3})		(0.054)		(0.047)
Control mean	0.081	0.081	0.867	0.867
Wald test p-val ($\gamma_{T1} \neq \gamma_{T2} \neq \gamma_{T3}$)		0.827		.602
Wald test p-val ($\gamma_{T1} \neq \gamma_{T2}$)		0.540		0.433
Wald test p-val ($\gamma_{T1} \neq \gamma_{T3}$)		0.606		0.904
Wald test p-val ($\gamma_{T2} \neq \gamma_{T3}$)		0.940		0.394
N	51441	51441	51486	51486

Standard errors are clustered at the facility level
Controls: months active on platform, baseline malaria sales, female owner, strata and calendar month FE
Wald test comparisons of difference in marginal effects (γ) between treatment arms
Denominator is all patients that purchased malaria product during study period
Col 1 & 2: 45 obs dropped b/c no transactions logged during study period
$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Table 3: Mechanisms analysis, price pass-through and provider counseling (Back: 5.3)

	Log Price of rapid test		Provider Counseled Patient on Treatment		
	(1)	(2)	(3)	(4)	(5)
Pooled treatment	-.181*		.111*		
	(0.088)		(0.056)		
T1		-.427*		.0353	.0354
		(0.174)		(0.069)	(0.069)
T2		-.0273		.162*	
		(0.094)		(0.069)	
T3		-.0895		.135+	
		(0.095)		(0.069)	
Pooled T2 & T3					.149*
					(0.059)
Control group mean	2.880	48.952	0.314	0.314	0.314
Test p-val ($\gamma_{T1} \neq \gamma_{T2} \neq \gamma_{T3}$)		0.09		0.314	
Test p-val ($\gamma_{T1} \neq \gamma_{T2}$)		0.029		0.169	
Test p-val ($\gamma_{T1} \neq \gamma_{T3}$)		0.066		0.075	
Test p-val ($\gamma_{T2} \neq \gamma_{T3}$)		0.548		0.157	
Test p-val ($\gamma_{T1} \neq \gamma_{T2\&T3}$)					0.064
N	137	137	411	411	411

Standard errors in parentheses

Columns 1 & 2: Facility-level analysis; Columns 3-5: SP-visit-level analysis

Wald test comparisons of difference in marginal effects (γ) between treatment arms

Column 1 & 2 covariates: facility-level indicator for whether any SP was offered an RDT, indicator for missing price information

Column 3-5 covariates: SP fixed effects

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Table 4: Patient and pharmacy welfare growth per unit of program cost

Comparison	Δ_D	Cost	Weight	C.E.	Welfare
	(1)	(2)	(3)	(4)	(5)
T1 vs. C	0.27	0.28	0.85	0.96	1.81
T2 vs. C	0.19	0.28	0.85	0.68	1.58
T3 vs. C	0.20	0.34	0.85	0.59	1.50

Note: This table shows the increase in patient and pharmacy welfare per unit of program cost. Column 1 is the effect of each arm on demand for RDTs. Column 2 is the per-patient average cost of the program across all malaria suspect patient encounters in a given study group. Column 3 is the social welfare weight following the formula presented in 19. Column 4 is the cost effectiveness of each arm in terms of RDT demand increase. Column 5 is the welfare gains for patients and providers per dollar cost for the social planner (equation 19 plus 1 to account for patient and provider transfers for ease of interpretation).

Table 5: Marginal effect of contract characteristics on patient demand for testing

	IV	OLS	IV
α $(p_r - p_a)$	-0.08641***	-0.08956***	-0.05987***
	(0.01854)	(0.01683)	(0.02133)
$\alpha_{a\|r}$ $(p_{a\|r})$	-0.08490***	-0.02015	-0.05283*
	(0.03005)	(0.03113)	(0.03158)
Provider incentives $(g(\pi, t))$	0.21913**		
	(0.10756)		
Implied patient's $P(m\|i)$	98%	22.5%	88%

Note: This table presents the results from equation 24. Variables $p_r - p_a$, $p_{a|r}$, and $g(\pi, t)$ are in USD. The first column shows the estimates from equation 24 using treatment assignment indicators and marginal costs of treatment as instruments. First stage F-statistics are 56, 68, and 63 respectively. The second column presents OLS estimates of a model that excludes controls for the provider's financial incentives. Column 3 shows IV estimates of a model that excludes the financial incentives of the provider. The last row shows the implied patient's belief about their probability of having malaria using the following formula: $\alpha_{a|r}/\alpha$.

Table 6: Counterfactual contracts

	Δ_D (1)	*Cost* (2)	C.E. (3)	Weight (4)	Welfare (5)
Panel A: Patient contracts					
RDT discounts	0.17	0.75	0.23	0.85	1.20
Diagnosis-contingent ACT discounts	0.17	0.08	2.27	0.85	2.93
Experimental contract	0.17	0.38	0.45	0.85	1.39
Panel B: Provider contracts					
RDT incentives	0.44	1.28	0.34	0.85	1.29
Diagnosis-contingent ACT incentives	0.04	0.05	0.90	0.85	1.76
Experimental contract	0.28	0.48	0.59	0.85	1.50

Notes: this table compares the welfare effects of different counterfactual contract designs. Counterfactual effects are estimated using the fitted results from the linear probability model estimated in Table 5 using the average input values in the control group as baseline.

Figures

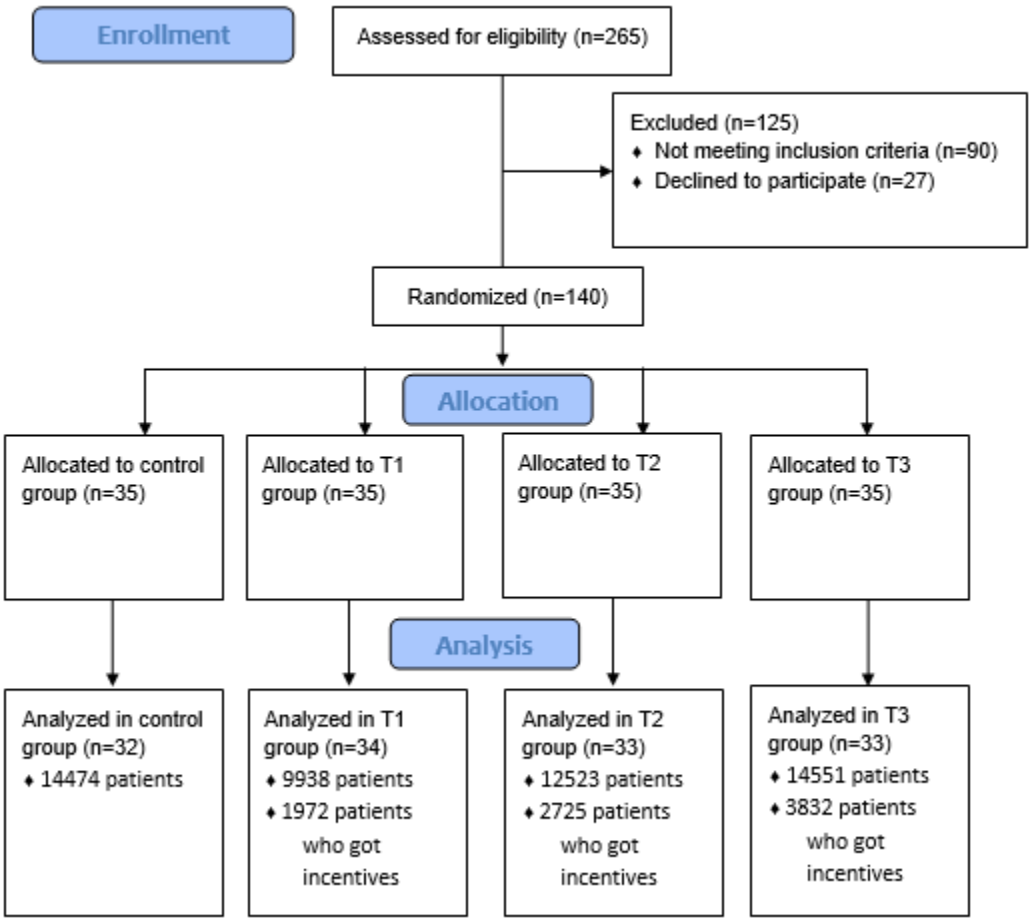

Figure 1: Study flow diagram (Back: 3)

Appendix A. Site-level test positivity rates detail

In the intervention arms (T1, T2, T3), test positivity for the tested sample is observed directly from transaction records for patients that tested for malaria using the incentivized rapid tests. In the control group, we do not observe test positivity for individual patients. In the transaction data for all sites, we do observe whether clients purchased a rapid test and what their treatment choice was. From administrative aggregate testing records provided by a subset of control group sites that keep records on malaria positivity rates, we know that 24% of tests came back positive between January - February 2022. We use this test positivity rate, combined with the test positivity rate obtained from an independent random testing exercise of a subset of patients seeking care in control group sites, to impute site-level test positivity rates absent any incentives for the control group. We follow a parallel process to obtain site-level test negativity rates (with the third, omitted, group being the untested sample).

Appendix B. Supplementary Tables and Figures

8.1 Appendix Tables and Figures

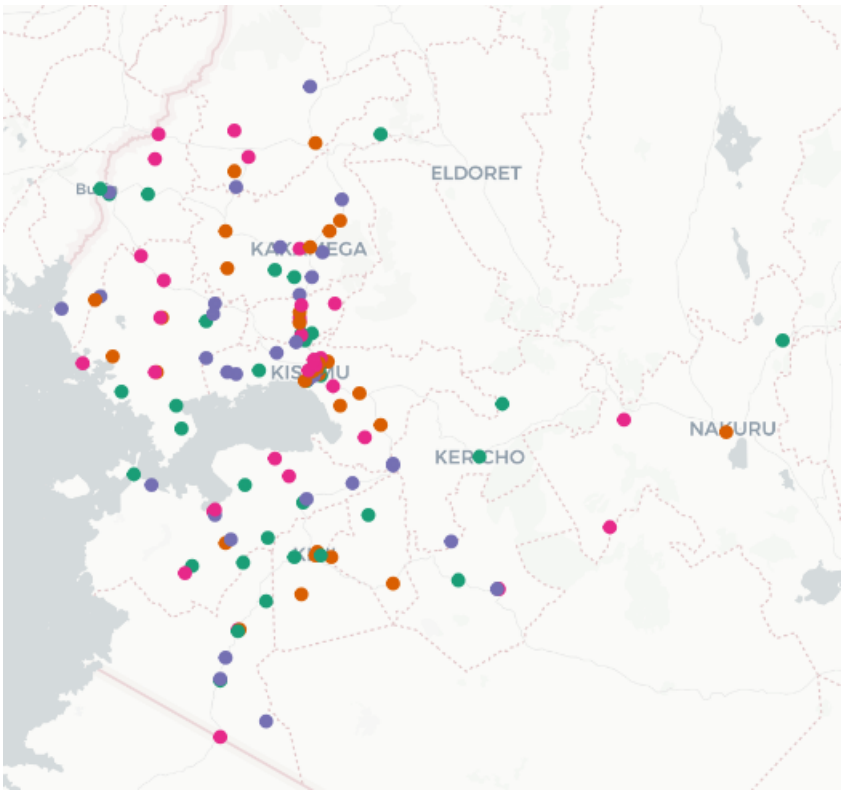

Figure B1: Map of study sites (Back: 3.1)

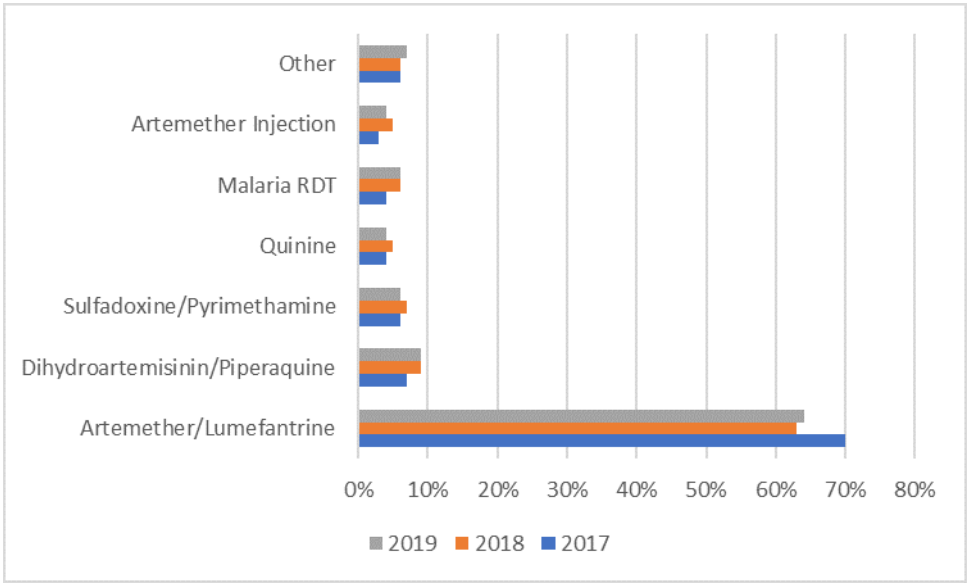

Figure B2: Malaria product sales volumes by type, 2017-2019 (Back: 5.1)

Table B1: Study timeline (Back: 3.2)

Jun-Dec '21	**Experiment launch**: baseline pharmacy survey with 233 pharmacy owners and staff from all 140 sites; staggered onboarding of 140 pharmacies to intervention and study
Aug '21-Feb '22	**Monitoring**: implementation team monitors intervention implementation through regular outreach calls and random site visits; ongoing administrative data collection through digital platform
Oct '21-Jan '22	**Patient exit survey**: survey of random sample of 1654 adult clients who seek care for malaria-like symptoms
Dec '21-Feb '22	**Standardized patient visits**: 412 mystery shopper visits by enumerators presenting as suspected malaria patients, to obtain data on patient-pharmacist interaction, implementation fidelity, and quality of care
Jan-Feb '22	**Control group testing**: testing of random subset of 230 pharmacy clients at control group sites to obtain test positivity rate
Mar '22	**Pharmacy endline survey**: survey of all pharmacy staff and owners at conclusion of the data collection period

Table B2: Baseline balance table, stratification variables (Back: 4)

	Below median malaria sales	Participated in earlier study phase	Urban	Malaria endemic location
Patient discounts	.0857	.0286	.0857	-.114
	(0.119)	(0.088)	(0.110)	(0.087)
Provider incentives	-2.59e-16	5.75e-17	.2^{+}	-.0857
	(0.119)	(0.088)	(0.110)	(0.087)
Hybrid	.0857	-.0857	.143	.0286
	(0.119)	(0.088)	(0.110)	(0.087)
Control mean	0.371	0.171	0.200	0.886
N	140	140	140	140

Multinomial logit test for joint orthogonality produces p-value from χ^2 test of 0.46
Binary Linear regression and linear probability models
$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Table B3: Baseline balance table, other baseline characteristics (Back: 4)

	Total months active	Average monthly malaria sales	Average monthly ACT sales	Average monthly RDT sales	Pharmacy only no clinic capability	Pct. female staff	Pharmacy owner age	Average staff age	Female owner	Number of staff
Patient discounts	-1.45	-6.29	-1.97	2.68	.105	-.0927	-1.51	.342	-.203*	-.0347
	(1.981)	(11.367)	(9.981)	(1.788)	(0.094)	(0.102)	(1.791)	(1.294)	(0.098)	(0.122)
Provider incentives	-4.15*	-9.81	-3.55	1.1	.059	.0475	-.356	.0757	-.214*	-.0433
	(2.004)	(11.501)	(10.098)	(1.809)	(0.095)	(0.103)	(1.812)	(1.310)	(0.099)	(0.123)
Hybrid	-.743	-12.8	-7.67	-.12	.116	-.00763	-1.42	.198	-.175+	-.115
	(1.977)	(11.342)	(9.959)	(1.784)	(0.094)	(0.102)	(1.787)	(1.291)	(0.097)	(0.121)
Control mean	10.143	64.797	52.845	4.304	0.714	0.443	37.429	29.371	0.357	1.543
N	140	140	140	140	140	140	140	140	140	140

Multinomial logit test for joint orthogonality produces p-value from χ^2 test of 0.46
Linear regression and linear probability models with strata fixed effects
$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Table B4: Baseline balance between facilities in sample and refusals (Back: 4)

Variable	(1) In sample	(2) Declined	(3) (2)-(1)
Number of months active on digital sales management tool	12.04	16.81	4.76**
	(9.43)	(8.52)	(0.01)
Average monthly malaria sales, 2019-2020	63.39	66.47	3.08
	(63.56)	(75.45)	(0.83)
Average monthly quality treatment sales, 2019-2020	54.41	61.25	6.84
	(54.26)	(72.61)	(0.59)
Average monthly rapid test sales, 2019-2020	6.48	5.39	-1.08
	(9.93)	(11.26)	(0.62)
Site was in earlier pilot study phase	0.16	0.23	0.07
	(0.37)	(0.43)	(0.32)
Site is in an urban area	0.31	0.34	0.04
	(0.46)	(0.48)	(0.69)
Site is in a malaria endemic county	0.84	0.86	0.01
	(0.37)	(0.36)	(0.84)
Site is a pharmacy	0.56	0.56	-0.00
	(0.50)	(0.51)	(1.00)
Observations	140	35	175

In sample facilities include those that were randomized to one of the study arms and were on-boarded successfully.

Table B5: Primary outcomes regressed on baseline characteristics (Back: 4)

	(1)	(2)	(3)
	Rapid test uptake	ACT uptake with test	ACT uptake without test
Months on sales management tool	.00143	.00197^{+}	-.00104
	(0.002)	(0.001)	(0.002)
Below median baseline malaria sales	.194**	.0369	-.155*
	(0.066)	(0.034)	(0.064)
Average monthly malaria sales, 2019-2020	-.000374	-.000687	-.00552*
	(0.001)	(0.000)	(0.002)
Average monthly ACT sales, 2019-2020	-.00211^{+}	-.000573	.00812**
	(0.001)	(0.001)	(0.003)
Average monthly rapid test sales, 2019-2020	.0157**	.0095**	-.0119**
	(0.003)	(0.002)	(0.003)
Site was in earlier pilot study phase	-.00984	-.00811	.0372
	(0.052)	(0.036)	(0.055)
Site is in an urban area	.0183	.0105	.0195
	(0.054)	(0.026)	(0.054)
Site is in a malaria endemic county	.0729	.0648**	-.105
	(0.073)	(0.024)	(0.073)
Site does not have clinical capabilities	.673**	.224**	-.652**
	(0.050)	(0.071)	(0.046)
% of staff who are female	.147^{+}	.0561	-.154*
	(0.078)	(0.039)	(0.075)
Age of pharmacy owner	.00767**	.00261*	-.00754*
	(0.003)	(0.001)	(0.003)
Average age of pharmacy staff	.000397	.000689	-.00217
	(0.006)	(0.002)	(0.005)
Female owner	-.202**	-.0851*	.182**
	(0.063)	(0.033)	(0.063)
Number of staff	.053	.0383	-.0728
	(0.060)	(0.029)	(0.059)
N	51486	51486	51486

Linear probability models for primary outcomes on baseline characteristics
Standard errors are clustered at the facility level
$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Table B6: Impact on antimalarial uptake, ACT uptake with & without test (Back: 5.2)

	Antimalarial sales overall		Non-ACT sales overall		ACT uptake with test		ACT uptake without test	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pooled treatment	.000427		-.0173		.0748*		-.197**	
	(0.023)		(0.022)		(0.034)		(0.060)	
Patient discount		.0259		-.0334		.072		-.218*
(γ_{T1})		(0.025)		(0.029)		(0.050)		(0.110)
Pharmacy incentive		-.00873		-.0327		.0769^{+}		-.161*
(γ_{T2})		(0.032)		(0.029)		(0.045)		(0.075)
Patient discount and		-.00655		.00262		.0511^{+}		-.183**
pharmacy incentive (γ_{T3})		(0.027)		(0.031)		(0.029)		(0.068)
Control mean	0.197	0.197	0.109	0.109	0.057	0.057	0.809	0.809
Wald test p-val ($\gamma_{T1} \neq \gamma_{T2} \neq \gamma_{T3}$)		0.231		0.582		0.839		0.881
Wald test p-val ($\gamma_{T1} \neq \gamma_{T2}$)		0.156		0.976		0.938		0.629
Wald test p-val ($\gamma_{T1} \neq \gamma_{T3}$)		0.202		0.333		0.710		0.782
Wald test p-val ($\gamma_{T2} \neq \gamma_{T3}$)		0.943		0.330		0.587		0.802
N	265610	265610	51181	51181	51486	51486	51441	51441

Standard errors are clustered at the facility level

Controls: months active on platform, baseline malaria sales, female owner, strata and calendar month FE

Wald test comparisons of difference in marginal effects (γ) between treatment arms

Denominator is all patients that purchased malaria product during study period

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Table B7: Inputs for social welfare effect calculations

	T1	T2	T3
(A) Average incentive cost inputs			
Share or transactions with diagnosis contingent ACT incentive	0.072	0.088	0.095
Share or transactions with RDT incentive	0.072	0.088	0.095
Diagnosis contingent ACT incentive cost	\$0.90	\$1.20	\$1.10
RDT incentive cost	\$1.10	\$0.80	\$0.90
(B) Social welfare weight inputs			
$P(m')$	0.90	0.90	0.90
c_a	\$1.58	\$1.58	\$1.58
c_r	\$0.57	\$0.57	\$0.57

Notes: This table presents the inputs used to estimate the welfare effects in table 4. Panel A presents the inputs used to estimate the average per patient cost of the incentives using administrative data on transactions and program implementation costs. Panel B presents the inputs for the social welfare weight lower-bound, using administrative data on marginal costs of treatment and testing, and estimated positivity rates from the random testing activity.

Appendix C. Proxy for provider's incentives

In this appendix we discuss the proxy used for the financial incentives of the provider to recommend $f(p, c, t)$. Given that we do not use a measure of θ for estimation, we need to proxy for $f(p, c, t)$. We first introduce some notation. Let $\Delta P(k) = P(k|\theta = 1) - P(k|\theta = 0)$ for product k. Then we can express $f(p, c, t)$ as:

$$f(p, c, t) = (\pi_r + t_r)\Delta P(r) + (\pi_{a|r} + t_{a|r})\Delta P(a|r) + (\pi_a + t_a)\Delta P(a)$$

Note that $P(a|r) = P(m)P(r)$, since only malaria positive patients who test purchase incentivized treatment. Therefore, $\Delta P(a|r) = P(m)\Delta P(r)$. Hence:

$$\begin{aligned} f(p, c, t) &= \Delta P(r)[(\pi_r + t_r) + P(m)(\pi_{a|r} + t_{a|r})] + (\pi_a + t_a)\Delta P(a) \\ &= \Delta P(r)g(p, c, t) + (\pi_a + t_a)\Delta P(a) \end{aligned}$$

We argue that $g(p, c, t) = (\pi_r + t_r) + P(m)(\pi_{a|r} + t_{a|r})$ can be used as a proxy for $f(p, c)$ for the purposes of our counterfactuals when using treatment indicators from the RCT as instruments in equation 24. First, note that $\Delta P(r)$ is a constant, so the coefficient for the financial incentives when using $g(p, c, t)$ will be scaled by this constant. Second, the experiment generates exogenous variation in the price of RDTs and diagnosis contingent incentives. If we assume that pharmacy prices for ACTs do not respond to the interventions, then the treatment assignment is orthogonal to $(\pi_a + t_a)\Delta P(a)$. This assumption is reasonable and consistent with the reduced form results on prices since, as seen in table 3, there is no passtrhough of pharmacy incentives on prices for RDTs and ACTs.