Published in final edited form as:

Rev Econ Stud. 2014 ; 81(4): 1331–1365. doi:10.1093/restud/rdu020.

Learning, Misallocation, and Technology Adoption: Evidence from New Malaria Therapy in Tanzania

Achyuta Adhvaryu*

*University of Michigan Ross School of Business, adhvayu@umich.edu, www.achadhvayu.com

Abstract

I study how the misallocation of new technology to individuals who have low *ex post* returns to its use affects learning and adoption behavior. I focus on antimalarial treatment, which is frequently over-prescribed in many low-income country contexts where diagnostic tests are inaccessible. I show that misdiagnosis reduces average therapeutic effectiveness, because only a fraction of adopters actually have malaria, and slows the rate of social learning due to increased noise. I use data on adoption choices, the timing and duration of fever episodes, and individual blood slide confirmations of malarial status from a pilot study for a new malaria therapy in Tanzania to show that individuals whose reference groups experienced fewer misdiagnoses exhibited stronger learning effects and were more likely to adopt.

Keywords

technology adoption; social learning; diagnosis; malaria; ACT; Tanzania

1 Introduction

How do individuals make decisions about the use of new technologies when the returns to adoption are uncertain, and may depend on the unobserved characteristics of adopters? This question has received considerable attention in economics, particularly in the literature on technology adoption in the developing world. Learning, either from one's own past usage or from the experiences of others, has been demonstrated to be a key driver of adoption in these settings. Much of this work has focused on agricultural innovations (Bandiera and Rasul, 2006; Conley and Udry, 2010; Foster and Rosenzweig, 1995; Munshi, 2004). A few recent studies examine the link between learning and adoption in the context of health technologies (Björkman-Nyqvist et al., 2013; Dupas, 2013; Kremer and Miguel, 2007).

In general, the literature mentioned above seeks to understand the evolution of beliefs and behavior when the returns to adoption vary based on individuals' fixed characteristics. For example, the expected yield for a new crop may differ based on soil quality or local access to fertilizer; how quickly do individuals learn whether and to what extent to cultivate the crop given this heterogeneity?

In this paper, I study a related but distinct learning process, applicable in contexts in which the returns to adoption exhibit heterogeneity based on unobserved factors that vary stochastically within individuals over time. This type of learning framework is particularly relevant in the case of therapeutic innovations, for which the returns to adoption rely crucially on the presence or absence of disease. A new therapy may be extremely beneficial for those who actually have the illness it is designed to treat, but may have low or even negative returns for those who do not.

Promoting the therapy without regard to this heterogeneity—and thus misallocating treatment to some fraction of individuals—would clearly be sub-optimal: at a minimum, resources would be spent inefficiently, and those who were treated inappropriately may be worse off than if they had not adopted. I show that in addition to these immediate negative effects, misallocation can stifle efficient information transfer and thus hamper adoption by rendering the treatment unappealing even for those who would reap high returns from its use.

In particular, I document the negative impacts of malaria misdiagnoses on the rates of learning and adoption for artemisinin-based combination therapy (ACT), a new and effective type of antimalarial treatment in Tanzania. I focus on this context for the following two reasons. First, malaria has large negative economic consequences (Bleakley, 2010; Cutler et al., 2010; Hong, 2007; Lucas, 2010; Sachs and Malaney, 2002), and access to and acceptance of effective malaria treatments is a key component of the global malaria control strategy. Yet, in many endemic malarial environments, parasitic resistance has rendered most existing malaria therapies ineffective (Baird, 2005).

ACT, the new therapy I study, is currently the most effective treatment for *P. falciparum*, the prevalent form of the malaria parasite in much of sub-Saharan Africa and Southeast Asia (Arrow et al., 2004). The unsubsidized price of ACT is generally prohibitively high, with a median retail price of above 6 USD per regimen. At this price, ACTs would be largely inaccessible; thus, a large global subsidy called the Affordable Medicines Facility for malaria (AMFm) is currently being piloted in 7 African countries and Cambodia. AMFm aims to reduce the retail price of ACT enough to make it comparable to other available therapies—that is, less than 1 USD per regimen (Tougher et al., 2012).

Second, misallocation of treatment—in this context, the over-treatment of fever—is overwhelmingly common (Amexo et al., 2004). Due to lack of access to proper diagnostic technology (optic microscopy or rapid diagnostic tests), most diagnoses are made based on the presence of fever, the primary symptom of malaria. Presumptively treating fever as malaria leads to vast over-treatment; for example, in one study in Tanzania, blood slide data showed that more than half the individuals receiving treatment for malaria at government hospitals were not actually infected (Reyburn et al., 2007). Studies examining the rate of misdiagnosis among individuals purchasing antimalarial therapies from informal private-sector providers find similar magnitudes (Cohen et al., 2012).

Misdiagnosis has grave consequences in both the short and long run for the health and economic well-being of affected populations. In the short run, inappropriate treatment induces losses in economic productivity, since patients are slower to recover when they are

not treated for the underlying cause of their symptoms (Amexo et al., 2004). Misallocation of treatment also results in less effective usage of public and private health care resources. Estimates of the expenditure outlays per fever episode—medicine prices, time and capital costs—imply that for each appropriately treated case of malaria, 3 to 5 dollars in public funds are wasted on an inappropriately treated case (as well as additional costs of delayed recovery and subsequent health care for misdiagnosed patients) (Chanda et al., 2011). In the long run, the over-prescription and overuse of proven therapies engendered by inadequate access to diagnostic technology speeds up the spread of parasitic resistance (Arrow et al., 2004).¹

I show in this study that there is an additional information-based externality cost to misdiagnosis: it can slow down the learning process, making the new therapy unappealing even for those who would reap high returns from adoption. I introduce misdiagnosis into an otherwise standard social learning model, in which individuals learn over time about the effectiveness of a new therapy from the outcomes of past adopters. I show that misdiagnosis affects learning and adoption behavior in two ways. First, misdiagnosis scales down the expected benefits of adoption, since even if the new therapy were fully effective, individuals would only realize its benefits if they really had the disease. Second, misdiagnosis, as it generates noise, makes it more difficult for individuals to extract information about the new therapy's effectiveness from past adopters' outcomes.

I test the model's predictions using household survey data from a pilot program in Tanzania, through which ACT was distributed at formal-sector health facilities and dispensaries. I exploit the random sampling and plausibly exogenous timing of survey enumeration to construct reference groups for learning based on the geographic and temporal proximity of self-reported acute illnesses.² I then use data on individual blood slide tests for malaria to calculate the rate of misdiagnosis faced by each individual's reference group. The main advantage of this strategy is that the misdiagnosis rate varies at the individual level, allowing me to restrict attention to variation within very narrowly defined village-by-time fixed effects.

Since the association between current adoption decisions and past adopters' outcomes could be due to factors other than learning, I difference across the treatment district and neighboring comparison districts, and before and after the new therapy's introduction. In line with the theory's predictions, my results show that 1) the introduction of ACT increased health facility usage for individuals with fever; 2) there is a strong learning-based influence of past adopters' outcomes on future adoption probabilities; but 3) both of these effects are significantly smaller for individuals whose reference groups happened to be more misdiagnosed.

This study makes three main contributions. First, it contributes to our inchoate understanding of the diffusion of health technologies, particularly for treatment innovations

¹It is worth noting that I only deal with one form of malaria misdiagnosis here: individuals receiving ACT when they are negative for malaria. Of course, in endemic malarial areas, the converse error—not seeking care when asymptomatic malaria is present—also commonly occurs. This type of parasitemia is important, particularly because it regulates the rate of malaria transmission (Lindblade et al., 2013). But as it does not directly affect learning after ACT usage, since the error occurs exactly when ACT is *not* used, I do not address it in this study.

²By “acute illness” I mean febrile illnesses starting within 14 days of the date of survey.

whose returns depend heavily on the presence of a specific disease. Encouraging efficient learning may be especially important in this context because of counterfeiting: as Björkman-Nyqvist et al. (2013) point out, learning about ACT effectiveness is complicated not only by the uncertainty of malaria, but also by high prevalence of lesser-quality counterfeit drugs in the informal sector. Relatedly, while empirical investigations of learning are increasingly common in the economics literature, few studies have sought to examine the determinants of the rate of learning. Understanding what drives variation in the magnitude of the learning effect across contexts is an important endeavor as policymakers seek new ways to encourage the rapid take-up of effective technologies.

Second, this study contributes to the literature on the sustainability of development aid in low-income countries. A commonly heard (and contested) argument in this debate is that the endogenous evolution of beliefs and adoption rates through learning may generate sustained take-up of effective innovations (Dupas, 2013; Kremer and Miguel, 2007). Price subsidies and effectiveness beliefs are, in effect, substitutes: the higher belief about product quality, the less subsidized it needs to be to get above the adoption threshold. However, the results of my study suggest that misdiagnosis can substantially hamper the upward evolution of beliefs through learning, and thus an initial subsidy for ACTs through AMFm may not be enough to generate sustained adoption of ACTs if the problem of misdiagnosis is not adequately addressed.

Finally, this study contributes to the literature on public subsidies for goods with positive externalities (Ashraf et al., 2010; Cohen and Dupas, 2010; Tarozzi et al., 2011). The results suggest that misdiagnosis generates a negative *information externality*, in addition to the potential epidemiological externality of abetting parasitic resistance. Thus, the introduction of diagnostic technology would have both a private and a public value. The optimal Pigouvian subsidy should take into account the relative magnitudes of these values: the higher the information externality via learning, the stronger the argument for a public subsidy.

The remainder of the paper is organized as follows. Section 2 develops the model. Section 3 describes the pilot program, survey data, and definitions of key variables. Section 4 shows some descriptive evidence on ACT's effectiveness, and adoption and learning behavior. Section 5 develops an empirical strategy to test the model's predictions. Section 6 reports the main results and summarizes robustness checks. Section 7 concludes.

2 Model

In this section, I develop a simple social learning model of adoption behavior in which individuals learn about the effectiveness of new malarial treatment from their own health outcomes and those of their neighbors who have adopted in the past. In each period, acutely ill individuals make adoption choices based on the common prior on the new treatment's effectiveness and the costs and (potentially heterogeneous) returns to adoption. Part of this return depends on the misdiagnosis of malaria, which occurs with known probability. I show that in this context, misdiagnosis negatively affects the rate of adoption and the speed of learning.

2.1 Setup

The model is in discrete time, and time periods are indexed by t . Consider a village composed of a set N of individuals, indexed by i , with $|N| = n$. In each period, a randomly chosen subset $N_t \subseteq N$ of individuals ($|N_t| = n_t \leq n$) falls acutely ill. Each acutely ill individual draws at random a malarial status, which is unobserved to the individual himself, as well as to the other villagers. Each individual in N_t then makes an adoption decision, and realizes a health outcome. The adoption choice and the ensuing health outcome are observed by all villagers.³ Information contained in the choices and outcomes of acutely ill individuals in period t is then used to form a posterior belief on the effectiveness of the new therapy; this posterior is then carried over into period $t + 1$. Note that all probabilities below (m, p, p) are conditional on being acutely ill. I omit this conditionality for brevity's sake.

2.2 Definitions

I begin by defining the following terms:

- Let $M \in \{0, 1\}$ be a random variable determining malarial status. $M = 1$ indicates the presence of malaria, and $M = 0$ indicates no malaria. Let $m = \Pr(M = 1)$. M_{it} is the realization of M for individual $i \in N_t$.
- $h_{it} \in \{0, 1\}$ denotes the adoption choice for $i \in N_t$, where $h_{it} = 1$ denotes adoption, and $h_{it} = 0$ denotes non-adoption.
- Let $D \in \{D^b, D^g\}$ be a random variable determining the length of illness, where $D^g < D^b$, i.e., the good health outcome is a speedier recovery from illness. D_{it} is the realization of D for $i \in N_t$.
- $\theta \in \{0, 1\}$ is the true effectiveness of the therapy, where $\theta = 1$ denotes effective and $\theta = 0$ denotes ineffective.
- The common period- t belief about effectiveness is: $q_t = \Pr(\theta = 1 | q_{t-1}, \{h_{it-1}, D_{it-1} | i \in N_{t-1}\})$.⁴

The probability of good and bad health outcomes being realized depends on the effectiveness of treatment, adoption choice, and malarial status. If the individual adopts the new therapy ($h_{it} = 1$) and has malaria, he will recover quickly if $\theta = 1$ (i.e., if the new therapy is effective). If it is ineffective, he will recover quickly with probability p , capturing the possibility that even ineffective therapy works some of the time. If the individual does not adopt ($h_{it} = 0$), he recovers quickly with probability p , reflecting the fact that alternative antimalarial treatments are relatively ineffective.⁵

³In addition, I assume that individual-specific characteristics, which play a role in the utility maximization problem described below, are common knowledge. Note that because the prior is common and choices and outcomes are public information, there is no private value to experimentation, as often seen, for example, in bandit models of learning (see, e.g., Jones and Gittins (1972) and Banks and Sundaram (1994)).

⁴I make the simplifying assumption of a common prior q_0 coming into the first period. Of course, it is likely that there will be a distribution of priors, some "optimistic" and some "pessimistic," when a new product is introduced (see, e.g., Kondor and Ujhelyi (2005)). I suppress this feature to focus on heterogeneity in realized outcomes arising from unobserved malarial status. If priors q_i were drawn from some distribution, and the draws were known for all i , then the model would essentially be equivalent to the simplified model I present, since observed characteristics are accounted for when extracting information via social learning. If prior beliefs are heterogeneous and private, however, this equivalence would no longer necessarily hold.

⁵Note that for simplicity, I equate the effectiveness of non-adoption conditional on malaria with the effectiveness of adoption when $\theta = 0$ conditional on malaria.

If the individual does not have malaria, regardless of adoption, he will recover quickly with probability p (despite having adopted the wrong therapy). This parameter captures the fact that some acutely ill individuals who do not have malaria may recover regardless of intervention—for example, those who caught a common cold—while some may need specific treatment for the underlying causes of their fevers, e.g. in the case of pneumonia.

These conditional probabilities are summarized in the equation below:

$$\Pr(D_{it}=D^g|\theta, h_{it}, M_{it})=h_{it}(M_{it}(\theta+(1-\theta)p)+(1-M_{it})\tilde{p})+(1-h_{it})(M_{it}p+(1-M_{it})\tilde{p}). \quad (1)$$

The impact of misdiagnosis on the learning process, as it turns out, depends on the relative magnitudes of p and $\tilde{p}$. Note from above that the lower is the effectiveness of the outside option (i.e., of existing therapy), the smaller p will be. The magnitude of p depends on the most prevalent causes of non-malarial fevers. This may differ significantly depending on geography, climate, demographic characteristics and baseline health of the population in question.

Evaluated over the distribution of M , the probabilities of receiving good or bad outcomes when the therapy is of high or low efficacy are the following:

$$\Pr(D_{it}=D^g|\theta=1; h_{it}=1)=m+(1-m)\tilde{p} \quad (2)$$

$$\Pr(D_{it}=D^g|\theta=0; h_{it}=1)=mp+(1-m)\tilde{p} \quad (3)$$

$$\Pr(D_{it}=D^b|\theta=1; h_{it}=1)=(1-m)(1-\tilde{p}) \quad (4)$$

$$\Pr(D_{it}=D^b|\theta=0; h_{it}=1)=m(1-p)+(1-m)(1-\tilde{p}). \quad (5)$$

2.3 Timing

The model begins in period 0, at which time a new therapy of unknown quality $\theta \in \{0, 1\}$ is introduced. Individuals learn about this quality (or alternatively, effectiveness) parameter over time by observing the history of adoption choices and realized health outcomes. I assume that in $t = 0$, all individuals begin with a common initial belief, $q_0 = \Pr(\theta = 1)$. For every period $t > 0$, the timing of the model is as follows:

1. All individuals enter period t with a common belief distribution, summarized by q_t , over quality.
2. A subset N_t of villagers fall acutely ill, and each draws a malarial status $M_{it} \in \{0, 1\}$, which is unobserved to the ill individual himself, as well as to the other villagers.
3. Each acutely ill individual makes an adoption choice $h_{it} \in \{0, 1\}$.

4. The resulting outcomes and adoption choices $\{D_{it}, h_{it} | i \in N_t\}$ are observed by all individuals.
5. The common belief distribution is updated, and a posterior belief q_{t+1} on the probability of effectiveness is formed.
6. Period $t + 1$ begins, and the process repeats.

2.4 Expected utility maximization

Utility is given as $u_i(C) - P(h)$, where the function u_i is increasing in consumption, C , and varies across individuals i . $P(h)$ is the price of health care at option h , and is measured in utils.⁶ The budget constraint is $C = w_i(\Omega_i - D)$, where w_i is the individual's wage rate and Ω_i is the amount of time he would work if fully healthy. This individual-level heterogeneity is perfectly observed by all individuals. The individual's expected utility maximization problem is thus $\max_{h \in \{0,1\}} \mathbb{E}_t(u_i(C) - P(h))$ subject to $C = w_i(\Omega_i - D)$, where $\mathbb{E}_t$ is the expectation taken using all known information up to and including period t .

Define $\bar{u}_i = u_i(w_i(\Omega_i - D^g))$ as utility under the good health outcome, and $\underline{u}_i = u_i(w_i(\Omega_i - D^b))$ as utility under the bad outcome. Expanding the expected value above using the definition of D from equation 1 and collecting terms, the maximization problem can be expressed as the following: individual i adopts in period t if and only if

$$q_t m(1 - p)(\bar{u}_i - \underline{u}_i) > P(1) - P(0). \quad (6)$$

Define $\kappa_i = \bar{u}_i - \underline{u}_i$ and $\Delta P = P(1) - P(0)$. The utility maximization problem can then be expressed as a simple cutoff rule: the acutely ill individual adopts if and only if the current-period prior on effectiveness exceeds a person-specific cutoff value:

$$h_{it} = 1 \left(q_t > \frac{\Delta P}{m(1 - p)\Delta u_i} \right). \quad (7)$$

I denote $\kappa_i = \frac{\Delta P}{m(1 - p)\Delta u_i}$. This cutoff responds in intuitive ways to changes in the model's parameters. An increase in the relative cost of adoption (ΔP) increases κ_i (i.e. makes adoption less likely). An increase in the rate of misdiagnosis ($1 - m$) increases κ_i . An increase in the effectiveness of the outside option p also increases the cutoff. Finally, an increase in the utility difference between quick and slow recovery from illness decreases κ_i .

2.5 Misdiagnosis and the adoption rate

In each period, let us denote the number of individuals who adopt as $n_{1t} = \sum_{i \in N_t} \mathbf{1}(q_t > \kappa_i)$

and those who do not as $n_{0t} = n_t - n_{1t}$. Define the period- t rate of adoption as $r_t = \frac{n_{1t}}{n_t}$, that is,

⁶I include the price of care directly in the utility function to simplify algebra: this assumption makes calculating a cutoff value for adoption more straightforward. If the price of health care were factored into the budget constraint (as we normally characterize the problem), then I would need to make an assumption on the functional form of the utility function to be able to show the existence of only one cutoff value and characterize that value.

the fraction of sick individuals who adopt in a given period. Proposition 1 below states that as the rate of misdiagnosis ($1 - m$) increases, the rate of adoption decreases. (See the appendix for proofs of propositions.)

Proposition 1 r_t is weakly decreasing in $(1 - m)$.

2.6 Misdiagnosis and the rate of learning

Next, I study how beliefs evolve over time through learning, and how misdiagnosis changes the learning process. Define the log-likelihood ratio of q_t as

$$\lambda_t = \log \left(\frac{q_t}{1 - q_t} \right). \quad (8)$$

From period to period, the log-likelihood ratio (equivalently, the belief q_t) evolves as individuals update the prior by incorporating new information contained in $\{h_{it-1}, D_{it-1} | i \in N_{t-1}\}$. Applying Bayes' rule, the updating equation can be expressed as:⁷

$$\lambda_{t+1} = \lambda_t + \sum_{k \in \{g,b\}} \sum_{i \in N_t} h_{it} \mathbf{1}(D_{it} = D^k) \log \left(\frac{\Pr(D_{it} = D^k | \theta = 1)}{\Pr(D_{it} = D^k | \theta = 0)} \right). \quad (9)$$

Bayes' rule aggregates information gained from the outcomes of period t adopters and converts it into (log) probability points. h_{it} scales the function of outcomes in equation 9 above because only adopters' outcomes are informative. The inner summation is over all sick individuals, while the outer is over good and bad outcomes. For each good outcome, the belief that the therapy's effectiveness is high ($\theta = 1$) should become stronger; for each bad outcome, that belief should become weaker.

Using the expressions for the probabilities above from equations 2 through 5, I rewrite the above equation as:

$$\lambda_{t+1} - \lambda_t = n_{1t}^g \log \left(\frac{m + (1 - m)\tilde{p}}{mp + (1 - m)\tilde{p}} \right) + n_{1t}^b \log \left(\frac{(1 - m)(1 - \tilde{p})}{m(1 - p) + (1 - m)(1 - \tilde{p})} \right), \quad (10)$$

where $n_{1t}^g = \sum_{i \in N_t} h_{it} \mathbf{1}(D_{it} = D^g)$ and $n_{1t}^b = \sum_{i \in N_t} h_{it} \mathbf{1}(D_{it} = D^b)$, so that $n_{1t}^g + n_{1t}^b = n_{1t}$.

Intuitively, if an individual adopts and realizes the good health outcome, the common prior on effectiveness should be revised upwards; if the adopter realizes the bad outcome, the opposite should happen. Finally, if no individual adopts, then no new information about effectiveness is revealed, and thus beliefs should not change. Whether beliefs about effectiveness go up or down from t to $t + 1$ depends on the proportion of adopters experiencing good and bad outcomes, and the magnitudes of the terms in logs, which reflect *how much* the belief should be scaled up or down for each individual adopter who experiences, respectively, a good or bad outcome.

⁷See Chamley (2004) for more details on the application of Bayes' rule in the canonical social learning model.

To determine how misdiagnosis changes the rate of learning, I examine the expected drift in the log-likelihood ratio conditional on $\theta = 1$, denoted as $\mathbb{E}\lambda_{t+1} - \lambda_t | \theta = 1$ (Chamley, 2004).⁸

Let $x_g := \log \left(\frac{m+(1-m)\tilde{p}}{mp+(1-m)\tilde{p}} \right) > 0$ and $x_b := \log \left(\frac{(1-m)(1-\tilde{p})}{m(1-p)+(1-m)(1-\tilde{p})} \right) < 0$. From equation 10 the expected drift can be expressed as:⁹

$$\mathbb{E}(\lambda_{t+1} - \lambda_t | \theta = 1) = x_g \mathbb{E}(n_1^g | \theta = 1) + x_b \mathbb{E}(n_1^b | \theta = 1). \quad (11)$$

Now, I study how this expected drift varies with the rate of misdiagnosis, $1 - m$; this exercise enables us to understand how the rate of learning changes when misdiagnosis increases. The following proposition states that the way in which the expected drift varies with the misdiagnosis rate depends on the magnitudes of p and $\tilde{p}$, i.e., the extent to which the existing malarial treatment (the outside option) is effective, compared to the rate at which non-malarial fevers resolve without intervention. Intuitively, the proposition states that if the existing treatment is sufficiently ineffective, higher misdiagnosis will generate slower learning.

Proposition 2 When $p > \tilde{p}$, $\mathbb{E}\lambda_{t+1} - \lambda_t | \theta = 1$ is weakly decreasing in $1 - m$.

Table 1 provides some empirical justification for the assumption on the ordering $p < \tilde{p}$ mentioned above. I define the empirical analog for p in two ways. In the first instance, p is defined as the proportion of individuals reporting acute illness in the 14 days before being surveyed who had malaria, who did not visit the health facility, and who were well by the time of survey. In the second, the restriction on health facility usage in the definition of p is loosened for all individuals living outside the treatment district and all individuals pre-intervention. Under both definitions, p is compared to $\tilde{p}$, defined as the proportion of individuals reporting acute illness in the 14 days before survey who did not have malaria and who were well by the time of survey. I compare p and $\tilde{p}$ for the whole sample, and only in the comparison districts.

Table 1 shows that under both definitions, and under both sample restrictions, $p < \tilde{p}$. The difference $p - \tilde{p}$ is significantly less than 0 in all but one of the four resulting cases. This evidence lends support to the assumption that, at least in this particular setting, the ordering $p < \tilde{p}$ holds true.

In summary, this model makes two key predictions. First, greater misdiagnosis discourages adoption via the lower expected benefits of adoption. Second, greater misdiagnosis decreases the rate of learning by introducing excess noise in the learning process.¹⁰ In the following sections, I test these predictions in the context of the introduction of ACT in Tanzania.

⁸Conditioning on $\theta = 1$ reflects the fact that the drift should be calculated for the true state, which in the case of effective therapy is $\theta = 1$.

⁹For brevity, I suppress the t subscript on n_{1t}^g and n_{1t}^b from now on.

¹⁰Note that this is, by proposition 2, only true in the case when $p < \tilde{p}$. As shown in Table 1, there is some support for this case in my empirical context.

3 Intervention, data, and definitions

3.1 ACT intervention

3.1.1 Policy context—Health care markets in Tanzania, as in many low-income countries, are highly informal. Formal health facilities (hospitals, regional health centers, and drug dispensaries) provide care for less than half of all illness episodes; this fraction is particularly low in rural areas. Most patients self-medicate, seeking care and purchasing medication from informal pharmacies, drug shops, kiosks, traditional healers, as well as family and friends. The vast majority of interaction with the health system is for curative as opposed to preventative care; among those seeking curative care, the most common symptom is fever (Adhvaryu and Nyshadham, 2012).

As most travel is done by foot, travel cost (utility cost and/or opportunity cost of time) is the most significant determinant of cost of care. This effectively makes formal health facilities much more expensive than self-treatment (Gertler et al., 1987). Nearly all pecuniary health care costs are out-of-pocket, as no formal insurance markets exist. Infants and young children, pregnant women, the elderly, and severely ill patients tend to select into formal care.¹¹ Care in the formal sector is delivered by health workers with varied training and experience. Fully trained doctors are few and far between, particularly in rural areas. Diagnosis and treatment are often low quality, both due to the paucity of training of workers and the frequent lack of adequate diagnostic tools and medicine stocks (Klemick et al., 2009).

The majority of malaria diagnosis in low-income countries is done on a presumptive basis, diagnosing malaria from clinical symptoms (primarily fever) alone (Amexo et al., 2004). Given that only a fraction of fevers are malarial, even in endemic malarial regions, the rate of non-malarial fevers treated as malaria is high (Reyburn et al., 2004).

At the time of the ACT intervention, the front line malaria treatment in Tanzania (and much of sub-Saharan Africa) was sulphadoxine pyrimethamine (SP). When it was introduced throughout the African continent in 2000, SP was extremely effective, as demonstrated by randomized controlled trials in malaria-endemic settings, in the treatment of *falciparum* malaria, the most prevalent type of malaria parasite in Africa. By 2003, parasitic resistance had reduced the effectiveness of SP by half in most of the places it was once effective (Malisa et al., 2011).

3.1.2 Summary of IMPACT-TZ intervention—The Interdisciplinary Monitoring Project for Antimalarial Combination Therapy in Tanzania (IMPACT-TZ) was implemented by The Centers for Disease Control and Prevention (CDC) and the Ifakara Health Institute (IHI) in Rufiji, a rural district in southeast Tanzania, from February 2003 to the end of 2006.¹² Under the auspices of the program, artesunate plus sulphadoxine pyrimethamine (the particular combination therapy used in this implementation trial), was prescribed to all

¹¹See, e.g., Cohen et al. (2012).

¹²For more detail on the implementation trial and household surveys, please refer to Kachur et al. (2001), Kachur et al. (2004), and Njau et al. (2008).

individuals seeking care at government- or NGO-operated health facilities with fever or a recent history of fever.

3.1.3 Rufiji district—Rufiji district is located in the Coast region, just under 200 km from Dar es Salaam, the commercial capital of Tanzania. It is divided into 6 divisions, which are subdivided into 19 wards containing 94 registered villages. The district had a population of about 212,000 at the time of ACT implementation, spread over about 13,000 square kilometers. The primary economic activity is subsistence agriculture. Villages are generally organized into clusters of family homes. Household farms are usually located some distance from these central clusters. Major staples farmed are cassava, maize, rice, and millet; the main cash crops are cashew nuts and a variety of fruits (mangoes, oranges, coconuts, etc.). Malaria transmission occurs throughout the year; incidence and prevalence of malaria are highest during the two rainy seasons. At the time of ACT implementation, Rufiji district had 59 health facilities (including hospitals, health centers, and dispensaries); of these, 46 were public, 10 were non-profits run by missions, and 3 were for-profit facilities (Njau et al., 2008).

3.1.4 Training, sensitization, and campaign launch—The 56 public and private non-profit health facilities in Rufiji and their surrounding catchment areas underwent preparation prior to ACT implementation. New drug stock log books were distributed. Drawing from five focus groups, a local artist created posters, wall charts, and other information, education, and communication materials for publicity and community sensitization to ACT. Dosing envelopes with age-specific dosing instructions were developed. Health worker training began in January 2003. Two staff from each health facility underwent an initial centralized training covering prescribing protocols, and the importance of emphasizing adherence to the ACT regimen. Trained staff members then conducted similar training sessions at their own facilities. A refresher course was done in mid 2004, 18 months after the start of the program. A public launch for ACT introduction was held in August 2003 in Ikwiriri, the largest market in Rufiji. Following this, sensitization meetings were held by the research team in smaller market areas.

3.1.5 Drug supplies—Artesunate and SP were not co-formulated (i.e., combined into a single pill, as most ACTs are now). The SP supply chain, through the Government of Tanzania's Medicine Stores Department, remained in place through the ACT intervention. Artesunate (Arsumax, 50 mg tablets) was sourced from Sanofi-Aventis, a large French drug manufacturer, via a separate supply chain (Njau et al., 2008). Artesunate was shipped to Kibiti Health Centre, a central health facility in Rufiji where the Medicine Stores Department made initial deliveries of SP into Rufiji. From here, it was shipped to the distal facilities via carriers hired by the research team. ACT was only available through the 56 public and mission health facilities in Rufiji district. Due to the CDC and IHI's continuous monitoring efforts, leakage of ACTs into the private sector was virtually nonexistent (Njau et al., 2008).

3.1.6 Prescription—The program's protocols specified that ACT was to be prescribed for every non-pregnant patient above the age of 2 months who presented at a public health

facility with a fever or recent history of fever (Kachur et al., 2004). In the event that artesunate was not available, health workers were advised to prescribe SP monotherapy instead. Dosing was age-specific; four age groups were used: children under 1 year, children between ages 1 and 5, school-aged children (under 15), and adults. The regimen entailed taking artesunate and SP on the first day of treatment (which was done at the health facility in the presence of the clinician immediately following diagnosis), and artesunate only on the second and third days (Njau et al., 2008).

Thwing et al. (2011) document ACT dispensing practices among health centers and dispensaries in Rufiji before and during the intervention, from 2002 to 2005. The authors show that once the intervention began, 75 to 80 percent (depending on the quarter of survey) of patients with fever or recent history of fever were prescribed ACT, and of those prescribed, virtually all patients received the correct age-specific dosage (Thwing et al., 2011). Only 10% of febrile patients received SP monotherapy during the intervention, and most of this gap occurred during the second quarter of 2003, during which there was a brief artesunate stock-out.

3.2 Data

3.2.1 Overview of survey data—This study uses data from household surveys conducted before and after the introduction of ACT, in Rufiji, the treatment district, and in two comparison districts, Kilombero and Ulanga. Households in villages which were part of the Demographic Surveillance System (DSS) in the treatment and comparison districts were sampled randomly to be surveyed.

The Rufiji and Kilombero/Ulanga DSSs are continuous monitoring infrastructures that track births, deaths, demographic and socioeconomic characteristics for all households in specific, contiguous geographic areas of these districts. General surveillance happens in rounds, such that each household in the DSS areas is visited approximately every 4 months. DSS areas within the three districts were chosen non-randomly. In Rufiji, the DSS covers 31 villages, a population of about 85,000; the DSS is serviced by 18 public and NGO health facilities. The DSS site in Kilombero/Ulanga (known as the Ifakara DSS) covers 25 villages, 13 in Kilombero and 12 in Ulanga, also with a population of about 85,000; the site is serviced by 15 public and NGO health facilities.¹³

The IMPACT-TZ household surveys used the DSS census to randomly sample households from Rufiji, Kilombero, and Ulanga for more in-depth study prior to and after the introduction of ACT in Rufiji's health facilities. A new random sample (sampled at the household level) was done for each survey round. I use rounds in which blood slide data were collected, namely, the 2001 pre-intervention round and the 2004 post-intervention round. There was no survey round in 2003, the year that the new therapy was introduced. The treatment and comparison districts are geographically contiguous but separated by a large game reserve (the Selous Reserve).

¹³For more detail on the Rufiji and Ifakara DSSs, see Mwageni et al. (2002) and Armstrong Schellenberg et al. (2001), respectively.

I focus my analysis on the sample of individuals who reported being acutely ill with fever in the past two weeks. The individual module of the household survey asks questions about treatment-seeking following an episode of fever; I construct the dependent and independent variables of interest based on the answers of individuals who were acutely ill in the recent past to questions about their health care choices and health outcomes.

Table 2 reports the means and standard deviations of variables used in analysis for the sample of individuals who reported being acutely ill with fever starting in the two weeks prior to survey.¹⁴ I present summary statistics for the whole sample, and split by treatment versus comparison districts, and finally results of tests for statistical differences in means across the two groups.

The total number of surveyed individuals in the selected rounds is nearly 17000, divided roughly evenly across the treatment and comparison districts. Of these individuals, a little over 1700 reported being ill with fever that began in the two weeks preceding survey, and also have non-missing values for education, age, and reference group misdiagnosis level (defined in section 5). This is the main sample used in analysis.

The unit for the age variable is years. It is constructed by subtracting the survey date from the date of birth of the individual, and dividing by 365.25. In all regression analyses, I include linear and quadratic terms in age as controls. The average age in the sample is about 24.5.

The average years of completed education of household heads is approximately 4.25. The household survey only asked for the educational attainment of the household head. This attainment is reported in years, but for all analyses, I divide education into four categories (no education, less than primary, completed primary only, and more than primary), and include dummy variables for each category in the regressions as controls (the omitted category is always no education). Note that primary completion is equivalent to 7 years of schooling. In this sample, the majority of household heads had not completed primary school. About 12 percent of the sample self-reported fever in the 2 weeks preceding survey.

Blood was drawn for all individuals consenting to a malaria test at the time of survey. Blood slides were sent to a lab to confirm presence or absence of *P. falciparum*, the most prevalent form of the malaria parasite in this region of Tanzania. Approximately 22 percent of surveyed individuals tested positive for malaria at the time of survey. See the appendix (section A) for a more detailed description of collection and analysis of the blood slide data.

Table 2 also reports pre-intervention means for key health care choice and health outcome variables. Pre-intervention, about 22 percent sought care at a formal-sector health center, hospital or dispensary (public and NGO). In the treatment district, those who sought treatment at a health facility after acute illness received ACT. I thus use formal-sector care usage as a proxy for ACT adoption in the treatment district.¹⁵ The rest of the individuals, those who did not visit the aforementioned options but who reported being acutely ill with

¹⁴The means are further restricted to the sample for whom education, age, and lagged average misdiagnosis level are non-missing, as this is the main sample on which analysis is run.

fever, went to a medicine shop, street doctor, general store, kiosk, traditional healer, private laboratory, or could have used modern or traditional medicines from home or from a neighbor, or could have sought no treatment at all.

Approximately 80 percent of individuals reporting recent fever reported that the fever had resolved by the time of survey. Among those individuals whose fevers had resolved, the length of illness is about 3.2 days. Among individuals who sought health facility care, the average duration of illness was about 4.2 days. Among those who did not choose health facility care, the length of illness is shorter—just under 3 days. This difference likely reflects the role of selection into health facility usage by individuals with more severe illnesses; even if the quality of care at health facilities is higher, realized outcomes across for health facility goers will be worse if this selection based on severity of illness is strong enough.

Columns 2 and 3 of Table 2 report summary statistics separately for the treatment and comparison districts, and Column 4 reports the differences in means across these two groups. On demographic characteristics, the treatment and comparison groups look very different. The average age among acutely ill individuals is higher in the treatment district compared to the comparison, and the average educational attainment of household heads is lower. On other dimensions, however, the two groups look quite similar. Malaria positivity is comparable across treatment and comparison, as is the proportion reporting recent fever. Pre-intervention means in outcomes are also not statistically different across the two groups.

Overall, then, the treatment and comparison districts appear different on demographic characteristics, but fairly similar with regard to epidemiological conditions, and, prior to ACT intervention, on health care choices and health outcomes as well. The empirical strategy laid out in section 5 attempts to control for the differences that do exist by controlling for survey respondents' demographics, as well as differences across space and time in unobserved factors that may be correlated with these characteristics.

3.3 Definitions of key empirical variables

3.3.1 ACT Adoption—I define ACT adoption as the choice of health facility care among individuals reporting recent fever in the treatment district (Rufiji) post-intervention. This is, perhaps, the most natural definition of adoption in the empirical context, for the following reasons. First, in the treatment district, ACT was only available through public health facilities.

Second, this definition is appropriate because the key choice made by each acutely ill individual is whether and where to seek treatment—at a health facility vis-a-vis in the informal sector. While the patient may be able to exercise additional agency once at the health facility, the treatment she receives is, at that point, in large part decided by the health worker who administers her care. Thus, at the individual level, the primary decision of seeking out ACT as treatment or not—i.e., the adoption decision—is made through the choice of care.

¹⁵Njau et al. (2008) confirm, via surprise visits to health facilities and drug stores in the treatment district, that 1) there was no leakage of ACT into the informal sector, and 2) that all individuals presenting with fevers at government and NGO health facilities were prescribed ACT.

Third, as per the CDC/IHI program protocols, surveyed individuals in the treatment district who self-reported fever and visited a health facility should have received ACT. It is, however, possible that the policy may not have been enforced in all cases.

Two main possibilities are of concern. First, artemisinin stock-outs may have prevented some febrile patients from accessing ACT. Second, in certain cases clinical judgement may have overruled the prescription policy. For example, if an individual visits the clinic with cough and shortness of breath along with fever, the health worker may conclude that pneumonia is more likely than malaria to be the underlying cause of the symptoms, and may therefore prescribe antibiotics in lieu of ACT. As summarized in section 3.1.6, the results from Thwing et al. (2011) suggest that both of these issues do not appear to be concerns in this trial.¹⁶

In sum, defining adoption via health care choice is appropriate in this context because the choice of care is the patient's primary decision, and for individuals with fever in the treatment district, choosing care at a public or mission health facility resulted in access to ACT in the vast majority of cases.

3.3.2 Lagged adoption and outcomes—I begin by introducing some notation.

Suppose individual i in village j falls sick with acute illness on date t . He makes an adoption choice, $h_{ijt} \in \{0, 1\}$ after falling ill, and his eventual health outcome, D_{ijt} , is measured as the length of illness in days. Note that if the individual is still ill when surveyed, the length of his illness will not be recorded (i.e., it will be coded as missing). I discuss the ramifications of this right-censoring of the distribution of the length of illness at the end of this section, and present evidence that it does not bias the estimate of the learning effect.

In line with the theory, the empirical model should reflect the intertemporal nature of the learning process: sick individuals should use the past health outcomes of adopters in their learning reference groups to update their priors on the quality of the new therapy. To construct the empirical analog, I must first define reference groups for learning.

My definition exploits the plausibly exogenous timing and location of survey enumeration to construct groups based on geographic and temporal proximity to the sick individual i . For geographic proximity I use the individual's village j , under the assumption that when individuals make health care choices, they learn from their fellow villagers who made similar choices in the past. Note that the survey instrument did not include data on social networks, which would have better reflected individuals' reference groups.¹⁷ For temporal proximity, I use information on the date the individual's acute illness began (t , as defined above). In particular, I assume that individuals falling sick on date t look back in time up to m days at the outcomes of adopters *in their village* who fell sick and made health care

¹⁶Even the brief stockout in 2003, in the same quarter that trial began, was well before surveying in the treatment and comparison districts began, in May 2004.

¹⁷I test the robustness of the main adoption and learning results to narrowing the definition of the reference group to: 1) the individual's household, and 2) groups defined by village x religious affiliation (Muslim or Christian). The results are reported in appendix Table 9. I report adoption and learning results by misdiagnosis level. Differences in adoption and learning across reference group misdiagnosis levels are not statistically different from the baseline estimates, though differences are slightly smaller than baseline estimates (contrary to expectations).

choices from date $t-m$ to date $t-1$. At the end of this section, I present evidence that the order of survey enumeration was plausibly random.

Let N_{jt} be the set of individuals who fell ill in village j on date t . Let $N_{jt}^1 \subseteq N_{jt}$ denote the set of all individuals with fever in village j on date t who adopted the new therapy, and $N_{jt}^0 \subseteq N_{jt}$ denote those who did not adopt, such that $N_{jt}^1 \cup N_{jt}^0 = N_{jt}$. Then $\bar{D}_{j,(t-m,t)}^1$ and $\bar{D}_{j,(t-m,t)}^0$, the average length of illness for adopters and non-adopters, respectively, from dates $t-m$ to t , are defined as follows:

$$\bar{D}_{j,(t-m,t)}^1 = \frac{\sum_{a=1}^m \sum_{i \in N_{j,t-a}^1} D_{ij,t-a}}{\sum_{b=1}^m |N_{j,t-b}^1|} \quad (12)$$

$$\bar{D}_{j,(t-m,t)}^0 = \frac{\sum_{a=1}^m \sum_{i \in N_{j,t-a}^0} D_{ij,t-a}}{\sum_{b=1}^m |N_{j,t-b}^0|}. \quad (13)$$

Equivalently, lagged average adoption is defined as the following:

$$\bar{h}_{j,(t-m,t)} = \frac{\sum_{a=1}^m \sum_{i \in N_{j,t-a}} h_{ij,t-a}}{\sum_{b=1}^m |N_{j,t-b}|}. \quad (14)$$

I make four notes regarding these definitions. First, defining the reference group in this way implies that the group's choices and outcomes vary in general at the level of village $\times$ day on which illness began. Thus it is possible that two individuals surveyed on the same day in the same village may have different lagged outcome and lagged adoption realizations if their illnesses began on different dates.

Second, I only average over observations for whom the length of illness is recorded, i.e., for those whose acute illnesses are complete by the date of survey. At the end of this section, I check that right-censoring of the length of illness is non-differential across treatment and control districts and misdiagnosis categories.

Third, if, for some village j , choice $k \in \{0, 1\}$, and time span $(t-m, t)$, $\sum_{b=1}^m |N_{j,t-b}^k| = 0$, I replace the missing value of $\bar{D}_{j,(t-m,t)}^k$ with the average length of illness for all individuals who chose k between $t-m$ and t as calculated across the health facility catchment area (a group of spatially proximate villages). If this value is missing as well, then I replace it with the same average across the entire district (further broadening the definition of the individual's reference group for these observations). This process is necessary for less than 10 percent of acutely sick individuals.

Fourth, in the estimates presented in section 6, I use a lag of 6 weeks ($m = 42$ days). As a check, I rerun the main analyses using 4, 5, 7 and 8 week lag lengths, and the results are qualitatively similar. These results are reported in the appendix (appendix Tables 5 and 6).

3.3.3 Misdiagnosis of malaria—I construct a measure of the extent of malaria misdiagnosis in each acutely ill individual's reference group, by exploiting individual-level blood slide data on malaria positivity collected at the time of survey enumeration. Define $p_{ijt} \in \{0, 1\}$ as the result of the blood slide test for individual i in village j surveyed on date t , where $p_{ijt} = 1$ denotes that the individual was found positive for malaria, and $p_{ijt} = 0$ denotes negative. The average level of reference group misdiagnosis is then calculated as the following:

$$\bar{M}_{j,(t-m,t)} = \frac{\sum_{a=1}^m \sum_{i \in N_{j,t-a}} (1 - p_{ij,t-a})}{\sum_{b=1}^m |N_{j,t-b}|}. \quad (15)$$

Intuitively, misdiagnosis is defined as the proportion of acutely ill individuals in the reference group who were confirmed as negative for malaria via blood slide test. I use this definition in the majority of the analysis presented in the next section. In the appendix (last two Columns of appendix Table 4), I re-do the main analyses using the misdiagnosis level for health facility goers only, and find similar results.¹⁸ This finding is consistent with a check I ran for differences in malaria positivity across health facility goers and those who did not choose health facility care in the pre-intervention round of the survey. The prevalence among health facility goers is 0.211 (SD = 0.427), while the prevalence among those who did not choose health facility care is 0.239 (SD = 0.409). The difference in means is .028 (SE = .0361), which is statistically indistinguishable from 0.

Figures 1 and 2 depict the variation in malaria misdiagnosis in two ways. First, to get a sense for seasonal fluctuations in misdiagnosis, in figure 1 I plot reference group misdiagnosis, as defined above, with the date of survey. The x-axis in this figure is the week of the year in which survey enumeration occurred, and the y-axis is the proportion of fevers who visited the health facility who were misdiagnosed (i.e., were non-malarial). I use the pooled data from both rounds of the survey (2001 and 2004). As figure 1 clearly shows, there is strong seasonal fluctuation in the frequency of malaria misdiagnosis. Average misdiagnosis varies from about 0.65 to above 0.85 over the course of the survey period (approximately May–September). These fluctuations emphasize the need to flexibly account for seasonal variation in order to properly isolate a learning effect.¹⁹

On the other hand, the possibility exists that flexible spatial and seasonal effects would absorb too much of the variation in misdiagnosis, not leaving enough to estimate the

¹⁸The ideal measure of misdiagnosis would be the proportion of non-malarial fevers actually treated with ACT. Since this is unavailable, I use a proxy measure—the extent of misdiagnosis in one's reference group—and check that the results are robust to restricting this measure to health facility goers in the reference group, as shown in appendix Table 4.

¹⁹Relatedly, trends in misdiagnosis over time are significantly different across treatment and comparison districts. The point estimate on the difference in trends (treatment x post-intervention) is 0.175, indicating a 22 percent relative change above the mean misdiagnosis level. This rather large change underscores the necessity of controlling for reference group misdiagnosis trends, which I do for all regression analyses.

learning effect with precision. Figure 2 explores this issue by plotting the histogram of residual variation in malaria misdiagnosis once village $\times$ week fixed effects have been absorbed.²⁰ Figure 2 shows that there is substantial variation in misdiagnosis even within narrowly defined village by time effects. This figure plots the histogram of residuals from a regression of misdiagnosis on village $\times$ week fixed effects; the residuals are thus by construction bounded between -1 and 1 . The residual variation comprises about 27 percent of the overall variation in misdiagnosis.

The main difficulty with defining misdiagnosis in this way is that the blood slide data were collected at the time of survey enumeration, not when the individual's illness actually began. The time lag between the start of illness and the time of survey could generate discrepancies in *prior* and *measured* malarial status. Specifically, two cases are problematic. First, some individuals who had malaria at the start of their illness may have received effective enough treatment that their observed malarial status (at the time of survey) is negative. Second, some individuals who did not have malaria at the start of their illness may have acquired it sometime between the start of illness and date of survey.

Let us address these two cases in turn. First, I argue that the proportion of acutely ill individuals whose prior malarial status was positive and measured status was negative is not large. Of 1891 acutely ill individuals with non-missing blood slide data, 1429 (about 75 percent) were measured as negative for malaria. Of these individuals, 1075 (about 57 percent of the total) self-medicated or chose to visit a health facility, where they could have obtained an antimalarial. The rest either sought/received no treatment or received traditional (non-antimalarial) treatment. Of these, some had malaria at the start of their illnesses and some did not. Supposing that this proportion is equal to the baseline rate of malarial prevalence among acutely ill individuals in the sample (about 22 percent), we are left with approximately 237 individuals (about 13 percent of the total 1891).

For these individuals, the only way to transition from positive to negative malarial status in the time between the start of sickness and time of survey is by taking effective treatment. The proportion of individuals receiving effective treatment is a function of the (local) resistance to antimalarial therapies other than ACT. For example, in many areas of Tanzania, more than half of *P. falciparum* samples had acquired resistance to sulphadoxine pyrimethamine (SP) by 2004 (Malisa et al., 2011). In the extreme case, one might argue that only ACT was maximally effective, meaning that only those individuals in the treatment district post-intervention would plausibly change from positive to negative malarial status in the above defined time span. Taking 50 percent as a more conservative estimate, we are left with about 119 individuals, which is approximately 6 percent of the total 1891.

The second scenario, in which individuals acquire malaria sometime between the start of illness and the time of survey, is likely very small, given that the average time elapsed (i.e. the average number of days after the start of illness individuals were surveyed) is just over 6 days. A very small proportion of individuals are likely to have acquired malaria during this short a time span on average.

²⁰I project misdiagnosis on the full set of village $\times$ week dummies and take the residual, which is then plotted in the figure.

Given these two facts, it is likely that for the great majority of acutely ill individuals (who comprise the reference groups for learning I construct here), measured malaria status (at the time of survey) equals prior malaria status (at the start of illness).

In a set of robustness checks, covered in detail in the appendix, I define two alternative measures of misdiagnosis that address the potential issues still outstanding from the discussion above. In particular, I first restrict the calculation of the lagged average misdiagnosis level to only those individuals in the reference group whose illnesses started within one week of survey. This restriction further shrinks the length of time between the prior malarial status and measured status that is the source of error. Second, for each individual in the sample, I predict malaria positivity using a nonparametric function of age, and use only the predicted component of malaria status to calculate the misdiagnosis rate in the reference group. As shown by the results described in the appendix, changing the definition of misdiagnosis in these ways does not substantially alter the main results.

4 Descriptive Evidence

I present a variety of summary statistics across treatment and comparison districts, pre- and post-intervention, and across individuals whose reference groups experiences high and low misdiagnosis. These statistics are meant to give the reader a better sense of some basic patterns in the raw data. Overall, the conditional means reported in the following Tables 1) bolster the assumption that health care behaviors across treatment and comparison districts were similar prior to ACT intervention, and 2) support the model's predictions regarding changes in health care choices and learning following ACT intervention.

4.1 Pre-intervention health facility usage

Table 3 breaks down pre-intervention rates of health facility usage among individuals reporting recent fever episodes. The first and second rows report pre-intervention means in treatment and comparison districts, respectively. The Columns further break down these means by reference group misdiagnosis status, split at the median misdiagnosis level. The third row and third Column report differences across treatment v. comparison districts, and high v. low misdiagnosis reference groups, respectively.

Across all four groups, the rate health facility usage is around 21–24 percent. Both single and double differences are small: the largest difference is about 3 percent. *t*-tests confirm that single and double differences were not significantly different from 0. Overall, these means bolster the assumption that pre-intervention health care behaviors were similar, not only across treatment and comparison groups, but also across individuals whose reference groups experienced different levels of misdiagnosis.

4.2 Pre-intervention trends

I then look at data from the two pre-intervention rounds, 2001 and 2002, using simple difference-in-differences regressions to study whether key variables of interest were trending differentially across treatment and comparison districts before the introduction of ACT. I regress each variable on a treatment $\times$ post-intervention interaction term and the

main effects of treatment and post-intervention, plus a constant. Standard errors are clustered at the village level.

Results are reported in Table 4. Columns 1 and 2 use age and years of schooling as the dependent variable, respectively, and make use of the whole pre-intervention sample. The variables in Columns 3–5 – health facility usage, recovery from illness by date of survey, and the length of illness in days – are only defined for individuals who reported recent illness (beginning in the two weeks preceding survey). The coefficients on the treatment $\times$ post-intervention dummy in each of the 5 regressions is not significantly different from 0, indicating that trends in these variables were indeed similar across treatment and comparison districts before the introduction of ACT. Note that several of the variables had a common trend in the region, but of course these trends are absorbed in the year dummies.

4.3 Did ACT work? Evidence from changes in length of illness

The first panel of Table 5 reports means in the length of illness (in days) pre- and post-intervention for individuals with and without malaria, in the sample of treatment district health facility goers. These statistics are meant to gauge the extent to which the introduction of ACT at treatment district health facilities changed illness lengths. I include differences across individuals with and without malaria because ACT introduction should only affect recovery times for malarial patients.²¹

Column 1 in this first panel shows pre- v. post-intervention means for length of illness for treatment district health facility goers with malaria. There is a substantial decrease in days of sickness over time, of about 1.65 days, for this group. Column 2 shows means for treatment district health facility goers without malaria. Here, very little decline is seen in the length of illness (the difference over time is about 0.53 days). The double difference (-1.12) echoes this pattern, demonstrating that the majority of the decline over time in the lengths of illness happened in malarial patients. These patterns, though not precisely estimated, are consistent with the idea that ACT should cause declines in the average length of illness for those with malaria (for whom ACT is very effective as treatment), but should not affect health outcomes on average for non-malarial patients.

The second panel of Table 5 shows a similar comparison of means of illness length in days, now for health facility goers with malaria across treatment and comparison districts before and after ACT. The first row of this panel (treatment district means pre- and post-intervention) is identical to the first row of the previous panel. The second row shows the same set of means for the comparison districts. Here we find that there was a modest decrease in illness length over time in the comparison districts as well – .86 days on average, or about half the size of the decline in the treatment district. It bears mention that there was, as I document, also a massive compositional shift in health facility usage in the treatment district relative to the comparison, and selection into health facility usage is likely linked with severity of illness, so we should be careful to interpret these results merely as suggestive evidence of ACT's impact.

²¹Of course, as discussed previously, malarial status is measured at the time of survey, so the same problem of incorrect categorization applies here.

4.4 Trends in health facility usage

Table 6 shows trends in health facility usage across treatment and comparison districts over time. This Table is meant to quantify the effect of ACT introduction on health facility usage rates in the raw data.

Pre-intervention health facility usage rates among individuals with fever were similar across treatment and comparison districts (Column 1). The difference in the third row of Column 1 confirms that the two rates were not significantly different. Post-intervention, health facility usage increased in both groups, but increased by substantially more in the treatment district (Column 2). Again, the difference in the third row confirms that post-intervention facility usage rates were significantly higher in the treatment district compared to control.

The differences in Column 3 show that while the comparison district health facility usage rate grew slightly over time, the growth in the treatment district was much larger (health facility usage nearly doubled in the treatment district). The double difference affirms that the change over time in health facility usage was much greater in the treatment district.

To get a sense for this change over time, I plot the single difference coefficients (the difference in health facility across treatment and comparison districts in a particular year) over time in figure 3. I include all five rounds of the survey, rather than just the two I use in the main results, because data on health facility usage was present throughout the survey, whereas malaria parasitemia was only collected in 2001 and 2004.

The trends in figure 3 are quite telling. First, we see that health facility usage is fairly stable in 2001 and 2002, before ACT introduction. After ACT (denoted by the red line in 2003), health facility usage spikes dramatically in the treatment district relative to the comparison. But after this initial-year spike, facility usage declines sharply over time, such that by 2006, the difference in facility usage across treatment and comparison districts is no different than it was at baseline. This spike followed by a lack of sustained adoption is consistent with the story that learning is impeded by misdiagnosis, and that longer-term adoption rates suffer as a result.²²

4.5 Trends in association between health facility usage and lagged illness length

Table 7 reports trends over time in the association between health facility usage and the average length of illness in individuals' reference groups.²³ This Table is meant to illustrate the effect of ACT introduction on the extent to which reference group illness length affects health care decisions in the subsequent period. Each cell contains a coefficient from a regression of health facility usage on the lagged difference in days sick across health facility goers and patients seeking self-treatment; village $\times$ week fixed effects; education; and a quadratic in age. The coefficient on lagged difference in illness length is reported for each sub-sample – treatment v. comparison districts, pre- and post-intervention.

²²This decline over time may represent a variety of factors, some related to behavior (e.g., optimistic priors) and some to implementation (e.g., there was a brief supply shock in 2006, which could have eroded public confidence in ACT availability at health facilities).

²³The latter variable (D) is the difference between the average length of illness for health facility goers and patients who self-treated or received no treatment at all.

Pre-intervention (Column 1), there is a weak positive association overall between health facility usage and reference group illness length (most prominently displayed in the treatment district). This relationship may be due, for example, to common preferences for health or common time- and location-specific health shocks, affecting both the index individual and his reference group. Post-intervention (Column), this positive relationship arises strongly in the comparison districts. In the treatment district, however, the association switches sign, becoming negative, a change that is consistent with learning behavior. The difference across time in the treatment district (row 1, Column 3) is strongly significant and negative, and in the comparison districts (row 2, Column 3) is weakly positive. Likewise, the post-intervention difference (row 3, Column 2) and the double difference (row 3, Column 3) are negative and significantly different from 0.

The first difference estimates of the “learning effect” range from about 0.10 to 0.13 – that is, a one day decrease in the average length of sickness for facility goes in a patient’s reference group increases the probability of adoption by about 10 percentage points. The double difference estimate is substantially larger, at nearly 0.18, due to the weakly positive trend in the association observed in the comparison districts over time (row 2).

Overall, Table 7 provides evidence that individuals in the treatment district began learning from their reference groups, i.e., their health care choices became appropriately responsive to changes in the health outcomes of their reference groups, after the new therapy became available. In the comparison districts, the trend went (weakly) in the reverse direction. In the following sections, I estimate this learning effect in a regression framework, and study the interaction of learning and the rate of misdiagnosis in the adoption of ACT.

5 Empirical strategy

The goal of this section is to develop an empirical strategy to test the implications of the learning model using data from household surveys before and after the introduction of the ACT pilot program. The main implications of the model are that when misdiagnosis of malaria is more prevalent, the adoption rate should be lower and learning—that is, the relationship between current-period adoption and previous-period health outcomes of adopters—should take place more slowly.

5.1 Misdiagnosis and the adoption rate

In this section, I develop an empirical test of the prediction that individuals whose reference groups faced higher misdiagnosis rates should be less likely to adopt. I use a difference in differences approach to measure the adoption rate, comparing health facility usage in the treatment and comparison groups before and after the intervention. I attribute the differential change in health facility usage over time in the treatment vis-a-vis the comparison group to the introduction of ACT. This differential change in the sample of acutely ill individuals thus defines the adoption rate. The checks presented at the end of this section ensure that the differential trends across groups are not due to differential selection into acute illness.

I first estimate the following difference in differences specification to estimate the average adoption rate in the pooled sample (γ):

$$h_{ijr} = \gamma T_j P_r + \alpha_r + \beta_j + \zeta \bar{M}_{j,(t-m,t)} + \mathbf{X}'_{ijr} \delta + \varepsilon_{ijr}. \quad (16)$$

Here i denotes individual, j denotes village (and β_j are village fixed effects), r denotes round of survey (and α_r are round fixed effects), and $\mathbf{X}_{ijr}$ is a vector of individual- and village-level controls, including the following variables: week of survey dummies to capture week-by-week seasonal variation; dummies for categories of educational attainment of the household head; and a quadratic term in age. The misdiagnosis rate faced by the individual's reference group (denoted $\bar{M}_{j,(t-m,t)}$) is also included in the specification.²⁴

I denote the dummy variable for health facility usage as h_{ijr} , which equals 1 if the individual sought care at a health facility, and 0 if the individual sought care at an informal care option or did not seek care at all. T_j is a treatment district dummy and P_r is a post-intervention dummy, which equals 1 in post-intervention rounds. The coefficient of interest is γ , the difference in differences estimate of the impact of ACT introduction on health facility usage.

In all regression analyses, I use a block bootstrap with 500 repetitions, clustered at the village level, to correct standard errors for the use of constructed regressors.²⁵ In a robustness check, I re-do the main analyses allowing for arbitrary correlation in the error term within health facility catchment areas, which are slightly larger than villages. The results are reported in appendix Table 13; these results show that estimates are precise even when allowing for this broader arbitrary correlation. Please see section N of the appendix for more details on how catchment areas are defined.

I then interact the treatment $\times$ post-intervention term with $\bar{M}_{j,(t-m,t)}$, include the second-order interactions, and estimate a triple interaction specification in the pooled sample, measuring the differential adoption rate across reference-group misdiagnosis rates:

$$h_{ijtr} = \gamma_1 \bar{M}_{j,(t-m,t)} T_j P_r + \gamma_2 \bar{M}_{j,(t-m,t)} T_j + \gamma_3 \bar{M}_{j,(t-m,t)} P_r + \gamma_4 T_j P_r + \gamma_5 \bar{M}_{j,(t-m,t)} + \alpha_r + \beta_j + \mathbf{X}'_{ijr} \delta + \varepsilon_{ijr}. \quad (17)$$

The coefficient of interest is γ_1 , measuring differential adoption across reference groups with differing levels of misdiagnosis.

5.2 Learning effect estimation

5.2.1 Estimating equation—I begin with equation 7 from the theoretical model, describing the cutoff rule governing the period- t adoption decision: $h_{it} = \mathbf{1}(q_t > \kappa_i)$. The decision rule specifies that individual i adopts if and only if q_t is greater than an individual- and illness episode-specific cutoff κ_i .

²⁴ t denotes the date of survey; t subscript suppressed in the remainder of the specification.

²⁵Note that I should be taking into account arbitrary correlation in the error term at the district level, as the intervention was rolled out across the entire treatment district (Rufiji). However, given that there is only one treatment district and two control districts, allowing for this is not feasible, and thus I cluster standard errors at the village level instead.

From equation 10, we know that log-likelihood of the current-period belief

$(\lambda(q_t) = \log\left(\frac{q_t}{1 - q_t}\right))$ is an additive function of the log likelihood of the previous period's belief $(\lambda(q_{t-1}))$, and a weighted average of the outcomes of previous adopters $(x_g n_1^g + x_b n_1^b)$, where the weights x_g and x_b both depend on the model's parameters.

Since λ is an increasing function of q , it has an inverse (λ^{-1}) , which is also increasing in q .

Equation 7 can thus be expressed as $h_{it} = 1(\kappa_i < \lambda^{-1}(\lambda(q_{t-1}) + x_g n_1^g + x_b n_1^b))$, or equivalently, applying the inverse again, as:

$$h_{it} = 1\left(\lambda(\kappa_i) < \lambda(q_{t-1}) + x_g n_1^g + x_b n_1^b\right). \quad (18)$$

I use the above equation as a starting point for estimating a learning effect.²⁶ The two main issues I encounter are dealing with the belief q_{t-1} , and creating an empirical analog for the averaged outcomes of last period's adopters $(x_g n_1^g + x_b n_1^b)$.

q_{t-1} is unobserved. If left in the error term, it will generate bias in the learning effect estimate, since it in part determines outcomes in $t - 1$ via selection into adoption in that period. I proxy for $\lambda(q_{t-1})$ with the adoption rate in $t - 1$, h_{t-1} .²⁷ I use a simple average of outcomes (days of sickness) of last period's adopters $(\bar{D}_{t-1}^1)$ instead of the weighted average the model suggests, since the empirical analogs to "good" and "bad" outcomes are not well defined. Finally, I proxy for $\lambda(\kappa_i)$ with individual characteristics $\mathbf{X}_i$, village $\times$ time fixed effects η_{jw} (where j denotes village and w denotes week of survey), plus an error term ε .²⁸ Substituting the empirical versions of these proxies into equation 18, as defined in section 3.3.2, I obtain:

$$h_{ijt} = 1\left(\varepsilon_{ijt} < \gamma_1 \bar{D}_{j,(t-m,t)}^1 + \gamma_2 \bar{h}_{j,(t-m,t)} - \eta_{jw} - \mathbf{X}_i' \beta\right). \quad (19)$$

The coefficient of interest is γ_1 , which reflects the size of the learning effect – the extent to which information gleaned from the previous-period outcomes of adopters influences the current-period probability of adoption.

5.2.2 Contagion-based autocorrelation and selection into adoption—Of course, it is unlikely in reality that the only way in which past adopters' outcomes and current adoption probabilities are linked is through the learning process. We might expect that the learning effect estimate, γ_1 above, would be biased due to what Manski (1993) terms

²⁶The general definition of a learning effect here is the extent to which the probability of adoption changes for a given change in the outcomes of previous adopters. The specific definition depends on the outcome measure used; for example, if k adopters in $t - 1$

switch from good to bad outcomes (i.e., n_1^g decreases by k and n_1^b increases by k), the probability of adoption should decrease as a function of $k(x_g - x_b)$ in t .

²⁷ $h_{t-1} = \frac{1}{c} \sum_{i \in N_{t-1}} 1(q_{t-1} > \kappa_i)$ is, admittedly, an imperfect proxy for $\lambda(q_{t-1})$, but it will certainly be correlated, allowing us to absorb some of the variation in q_{t-1} .

²⁸Recall that κ_i is a function of the relative price of adoption; the rate of malarial prevalence conditional on fever; the probability of quick recovery from malaria after treatment with a non-ACT; and the utility differential between the good and bad outcome.

correlated effects. Individuals likely share common preferences for health with their reference group; have similar stocks of health as well as options for health care; and are exposed to the same local disease environment. Moreover, since this disease environment is often highly seasonal, outcomes and health care choices could be locally autocorrelated due to, for example, persistently heavy rainfall or high temperatures. Finally, it is possible that individuals are constantly engaged in learning about (spatially or temporally) local malarial prevalence; if this is the case, reference group outcomes might contain signals of local prevalence rates, and thus the relationship between previous outcomes and current-period adoption probabilities would not be due to learning about ACT's effectiveness, but rather about prevalence.

My empirical strategy aims to disentangle learning from these correlated effects. I take several steps to address bias arising from the fact that sick individuals and their reference groups have similar characteristics (health stocks, common shocks, preferences, availability of health care options, etc.).

First, as shown in equation 19, I flexibly control for local, time-varying unobserved factors by introducing the full set of village $\times$ week-of-survey fixed effects, denoted η_{jw} . These dummies stringently account for fluctuations in the local environment that simultaneously drive current-period adoption and past-period health outcomes. I am thus restricting attention to variation in lagged outcomes (as defined above) within village $\times$ week cells.²⁹

But even within fixed effect cells, geographically and temporally local shocks—epidemics, weather fluctuations, drug stock-outs and the like—could potentially bias learning effect estimates. To deal with this possibility, the second aspect of my empirical strategy is to difference the past outcomes of adopters and non-adopters:

$\Delta \bar{D}_{j,(t-m,t)} := (\bar{D}_{j,(t-m,t)}^1 - \bar{D}_{j,(t-m,t)}^0)$. To the extent that sick adopters and non-adopters in the same reference group are affected equally by these common shocks, differencing their outcomes will remove the effect of the common shock from the health outcome measure.³⁰

Finally, the possibility that the health outcomes of adopters and non-adopters may indeed *not* react in the same way to shocks must be reckoned with. Adoption is inherently driven by choice, and the unobserved characteristics of sick individuals—for example, the severity of their illness or their preferences for health—likely drive their adoption choices, and will also be correlated with the way in which they react to a common shock. As a result, shocks which affect current-period choices may be correlated with $\bar{D}_{j,(t-m,t)}$ as well.

To account for this possibility, I exploit data on the choices and outcomes of 1) individuals in the comparison districts, and 2) individuals before the introduction of the new therapy. The intuition behind my strategy is that, for these individuals, the correlation between

²⁹Since the lagged adoption and outcomes variables are defined at the level of village $\times$ day the illness began, there exists substantial variation in these variables across *days and individuals* within fixed effect cells, as shown in figure 2 in section 3.

³⁰Thus I posit that individuals learn from the *differential* outcomes of adopters as compared to non-adopters in their reference groups. Note that although this distinction cannot be made in the theoretical model, as for simplicity I assumed that the only intertemporal link between outcomes and adoption is through learning, it is nevertheless crucial to make in the empirical setting, in which common autocorrelated shocks may bias estimates of the learning effect.

current-period health care choices (h_{ijt}) and previous-period (differential) health outcomes ($\bar{D}_{j,(t-m,t)}$) should represent only the spurious effects induced by common shocks, since the therapy was not introduced to these individuals, so no learning effect should be present for these groups. To purge the coefficient on $\bar{D}_{j,(t-m,t)}$ of these spurious effects, I interact the variable with a treatment $\times$ post-intervention dummy (denoted $T_j P_r$), as well as the main effects—treatment (T_j) and post-intervention (P_r), where r denotes round of survey. Representing equation 19 as a linear probability model, the resulting triple difference specification is:

$$h_{ijtrw} = (\gamma_1 T_j P_r + \gamma_2 T_j + \gamma_3 P_r + \gamma_4) \Delta \bar{D}_{j,(t-m,t)} + \eta_{jw} + \gamma_5 \bar{h}_{j,(t-m,t)} + \mathbf{X}'_{ijt} \beta + \varepsilon_{ijtrw}. \quad (20)$$

The learning effect is captured by the coefficient γ_1 . $\mathbf{X}$ includes the following variables: dummies for categories of educational attainment of the household head; a quadratic term in age; and the date on which the individual fell acutely ill.

5.3 Misdiagnosis and the size of the learning effect

I first estimate a learning effect via equation 20 in the pooled sample. Then, to test the prediction that misdiagnosis decreases the rate of learning, I estimate the difference in the learning effect across differing reference group misdiagnosis levels. The coefficient of interest (γ_1) is on the interaction of the learning effect estimate and the reference group misdiagnosis rate. The interaction specification estimated in the pooled sample is:

$$\begin{aligned} h_{ijtrw} = & (\gamma_1 T_j P_r \\ & + \gamma_2 T_j + \gamma_3 P_r + \gamma_4) \Delta \bar{D}_{j,(t-m,t)} \bar{M}_{j,(t-m,t)} \\ & + (\gamma_5 T_j P_r + \gamma_6 T_j + \gamma_7 P_r + \gamma_8) \Delta \bar{D}_{j,(t-m,t)} \\ & + (\gamma_9 T_j P_r \\ & + \gamma_{10} T_j + \gamma_{11} P_r + \gamma_{12}) \bar{M}_{j,(t-m,t)} \\ & + \gamma_{13} \Delta \bar{D}_{j,(t-m,t)} \bar{M}_{j,(t-m,t)} \\ & + \gamma_{14} \bar{M}_{j,(t-m,t)} \\ & + \gamma_{15} \Delta \bar{D}_{j,(t-m,t)} \\ & + \eta_{jw} + \gamma_{16} \bar{h}_{j,(t-m,t)} \\ & + \mathbf{X}'_{ijt} \beta + \varepsilon_{ittrw}. \end{aligned} \quad (21)$$

6 Results

6.1 Misdiagnosis and ACT adoption

Table 8 reports regression results of ACT adoption in the pooled sample, as well as triple difference estimates of differences in adoption across reference group misdiagnosis levels. Column 1 reports the ACT adoption rate, i.e., the double difference in health facility usage over time and across treatment and comparison districts. The coefficient estimate is about 0.18 and is significantly different from 0; this magnitude is roughly in line with the double difference estimate of unadjusted means from Table 6, which was approximately 0.16. The size of the effect is large, about 75% of treatment group mean pre-intervention. In other

words, health facility usage was quite elastic to the introduction of ACT in the treatment district.

Columns 2–5 in Table 8 report results of triple difference specifications in which reference group misdiagnosis level, defined in four ways, is interacted with the treatment $\times$ post-intervention dummy (as well as lower-level interactions). In Column 2, the continuous misdiagnosis level interaction is reported. The negative coefficient on the triple interaction term indicates that adoption was smaller for individuals whose reference groups faced more misdiagnosis. To get a sense of the magnitude of this effect, increasing misdiagnosis by 10 percentage points decreases adoption by about 6.3 points.

Columns 3–5 use different definitions of the reference group misdiagnosis variable to generate heterogeneous effects across different cutoffs, from very low misdiagnosis (< 25 th percentile) to relatively high (< 75 th percentile). These cutoffs help us to understand where the heterogeneity in adoption across reference group misdiagnosis level is most pronounced. The triple difference coefficients in Columns 3–5 suggest that the biggest differential in adoption comes when placing the cut at the 25th percentile (Column 3) – that is, comparing very low misdiagnosis reference groups to the rest of the sample. Note also that there is a clear downward trend in the triple difference coefficients: as the cutoff level goes from 25th to 75th percentile, the adoption differential across low and high misdiagnosis groups shrinks. I make these remarks only suggestively, because precision on estimates in Columns 3–5 is lacking.

6.2 Estimates of the learning effect

In Table 9, I report estimates of the learning effect. Column 1 reports estimates of the baseline specification, equation 20. The learning effect estimate is the coefficient on the interaction of the differential illness length across adopters and non-adopters $\times$ treatment district $\times$ post-intervention. The learning effect estimate is large relative to the mean health facility usage rate and is precisely estimated. The interpretation of this estimate is that narrowing the difference in the length of illness across past adopters and non-adopters by 1 day *increases* the future adoption probability by just over 20 points. If we take the rough double difference estimate of the “impact” of ACT introduction on the length of illness from Table 5, for 0.64 day reduction in the length of illness, future adoption increases by about 14 percentage points.

Columns 2 and 3 report the results of robustness checks, in which lagged differences in demographic characteristics and symptoms across adopters and non-adopters, constructed in the same way as the lagged differences in the length of illness, are interacted with treatment and post-intervention dummies. These Columns report estimations of the augmented learning specification in equation 32, described in more detail in appendix section M.

In Column 2, I estimate the above specification using age, education of the household head and a wealth index (generated via principal components analysis) as x variables. In Column 3, I include demographic characteristics as well as the number of additional self-reported symptoms, a measure of the severity of illness, averaged over reference group adopters and non-adopters. The estimates in Columns 2 and 3 suggest that the addition of these lagged

differences and their interactions does not affect the magnitude or significance of the learning effect (the Column 2 estimate is -0.25 , and the Column 3 estimate is -0.22).

6.3 Misdiagnosis and the learning effect

Table 10 reports estimates of the learning effect by reference group misdiagnosis level (estimates of equation 21 in the pooled sample). I use the same four functional representations of misdiagnosis level as shown in Table 8. Column 1 shows the continuous misdiagnosis level interaction. The interaction coefficient is positive, large, and precisely estimated. To give a sense of the magnitude of the difference of the learning effect across misdiagnosis levels, a 10 percentage point increase in reference group misdiagnosis level diminishes the learning effect (i.e., makes the sum of the main effect and interaction coefficients less negative) by about 6 percentage points.

Columns 2–4 show specifications with binary definitions of reference group misdiagnosis level, with cuts at the 25th, 50th, and 75th percentile, respectively. In each Column, the difference in learning effect coefficients across “high” v. “low” misdiagnosis categories is always significantly different from 0. The magnitudes fluctuate slightly (ranging from about -0.20 to -0.24) but are not statistically different from each other. The interpretation of these coefficients is that future adoption is about 20 percentage points more sensitive to past outcomes of adopters when the individual’s reference group happened to experience a “low” misdiagnosis level.

6.4 Checks

I relegate an in-depth discussion of robustness checks, selection tests, and falsification exercises to the appendix. Below, I provide a list and brief summaries of each check.

Timing of survey enumeration—I test whether key model variables significantly affect the timing of survey enumeration within village $\times$ time fixed effect cells. Results are presented in appendix Table 1. In general, none of the observables is significantly associated with timing of enumeration.

Truncation of length of acute illness—I examine the determinants of truncation of reported illness length (due to the illness not having ended by the time of survey). Results are presented in appendix Table 2. Reference group misdiagnosis level (main effect) is significantly associated with truncation, but this precision declines when village $\times$ time fixed effects are added.

Sample selection checks—I examine the determinants of self-reporting of acute illness and blood slide testing for malaria. Results are presented in appendix Table 3. None of the important variables (treatment and post-intervention dummies, and their interactions with reference group misdiagnosis level) are significantly associated with self-reporting of acute illness. The (treatment $\times$ post $\times$ misdiagnosis level) triple interaction term and the (treatment $\times$ post) dummy are weakly associated (at 10% significance) with being tested for malaria.

Alternative definitions of reference group misdiagnosis level—I test the robustness of the main results to changing the definition of reference group misdiagnosis levels. I use two alternative definitions: the first calculates average misdiagnosis using only the malarial statuses of recent illnesses (starting 1 week or less before the date of survey), and the second uses predicted presence of malaria based on a non-parametric function of age. Results are presented in appendix Table 4. The main results on adoption and learning remain significant under the first definition, but lose precision under the second definition, though the magnitudes of the coefficients are unchanged.

Alternative lag lengths—I test the robustness of the main adoption and learning results to changing to length of the lag used to define the reference group. I vary the lag length from 4 to 8 weeks (the main specifications used a 6 week lag). Results are presented in appendix Tables 5 and 6. The adoption result remains significant for 5, 7, and 8 week lags, and all coefficients are fairly similar to the main results. The learning result remains significant for all four alternative lag lengths.

Alternative reference group windows—I test the robustness of the main results to changes in definition of the reference group's window. I first restrict analysis to individuals with reference groups with a full 6 weeks of data (i.e., all individuals surveyed 6 weeks or more past the first date of survey in their village). I then separately estimate adoption and learning effects in early and late windows of survey enumeration, to get a sense of the evolution of these effects over time. Results are presented in appendix Table 7. When the same window length is used for all individuals, the main results on adoption and learning lose significance, though the coefficients remain of similar magnitude. When two separate windows are used to estimate evolution of effects over time, I don't find much substantial difference across early and late periods in the extent to which misdiagnosis changes adoption and learning. For both of these exercises, one should exercise caution in interpreting the results, as only 40% of the sample can be used for the former (same window), and 50% for the latter (the sample is split into two here).

Incremental addition of controls—I test whether the main results are sensitive to controls used in the specification. The results are reported in appendix Table 8. The main results on adoption and learning do not survive if village and survey round (for adoption) and village $\times$ time (for learning) effects are removed.

Alternative reference group definitions—I test the robustness of the main results to two alternative definitions of the reference group, and two alternative definitions of illness: 1) the individual's household, and 2) groups defined by village $\times$ religious affiliation (Muslim or Christian). The results are reported in appendix Table 9. The adoption result remains significant when using the religion-based definition, but loses significance for the household reference group definition, and the magnitude cuts by about half, though it is not significantly different from the baseline (village-defined) result. The significance and magnitude for the learning coefficient survives for both definitions.

Alternative illness episode definitions—I test the robustness of the main results to two alternative definitions of the average illness experiences of the reference group: 1) with

lower-bound values for censored illness episodes included, and 2) using a binary for “well at the time of survey” (after reporting recent fever) instead of number of days sick as the main outcome. The results are reported in appendix Table 10. The main results survive for both definitions.

Robustness to additional controls—I test the robustness of the main results to the addition of two important potential confounders: 1) the time elapsed since reference group has sought care (measured as the average distance in days between the index individual and her reference group); and 2) congestion (measured as the number of health facility users in the individual’s reference group). Results are reported in appendix Table 11. The significance and magnitude of the main results survive the addition of both controls.

Placebo check using future reference group—Finally, I perform a placebo test in which I create reference groups based on future behavior and outcomes, rather than past. I then construct the same variables related to average adoption, health outcomes, and misdiagnosis levels in the individual’s reference group, as before, using the new, future-defined reference group, and regress these variables in the same main specifications to test for adoption and learning effects. Results are presented in appendix Table 12. The main results on adoption and learning are not significant when the future-defined reference group is used.

7 Conclusion

In this study, I demonstrate how the acceptance and adoption of effective technologies can hinge on the way in which they are allocated. In the case of malaria therapy, I show that the misdiagnosis of malaria affects individuals’ beliefs and subsequent adoption patterns through learning. When individuals are uncertain about the effectiveness of new therapy, misdiagnosis of malaria makes it more difficult to extract a signal about quality from the health outcomes of adopters. It also scales down the expected benefit of the therapy, since individuals who are unsure that they have malaria know that even if the therapy is effective, they will only realize its benefits if they actually have the disease. In both these ways, poor diagnostic policy can discourage the adoption of new malaria therapy even if the therapy is clinically effective.

I develop a strategy to test these hypotheses empirically, using household survey data from a pilot program through which ACT was prescribed at health facilities in Tanzania. I find evidence that 1) individuals whose reference groups experienced idiosyncratically more frequent misdiagnoses have lower adoption rates, and 2) individuals learn from the health outcomes of past adopters, but misdiagnosis decreases the extent of this learning.

This study underscores the importance of targeting therapeutic innovations through diagnosis. Introducing appropriate diagnostic technology can speed up learning and generate higher average returns for those who adopt, since a higher fraction of these individuals will be positive for the disease that the therapy is designed to treat. Given the decreased efficiency of information transfer that is induced by misdiagnosis, there is public value to diagnostic technology, and thus the optimal subsidy level should be positive. Future research

should be undertaken to ascertain what exactly the magnitude of the public subsidy should be.

Recent studies on the effectiveness of rapid diagnostic tests (RDTs) for malaria in Tanzania show that the introduction of RDTs at public clinics (Reyburn et al., 2007) and drug shops (Cohen et al., 2012) reduces the rate of misdiagnosis by about half, but that a substantial fraction—about one-half—of individuals who are diagnosed negative for malaria by RDT still buy or receive ACT.³¹ My results are complementary to this latter study: I leverage non-experimental variation and detailed data on a large sample to document the potential importance of inappropriately targeting ACTs, and Cohen et al. (2012) experimentally evaluates methods for improving the targeting of ACTs.

The next frontiers in this field are 1) understanding why some individuals' behaviors are not influenced by the test results, and 2) structuring the roll-out of RDTs in the public and private sectors such that patients' and providers' incentives for ACT adoption after diagnosis are taken into account. For example, drug shop owners selling both RDTs and ACTs may have a financial incentive to push the therapy to customers despite a negative diagnostic result. This incentive could be balanced by concurrently rolling out a treatment package for non-malarial illnesses, or by offering incentives for referrals to a nearby formal sector health facility. Structuring this roll-out correctly on a wide scale should lead not only to better outcomes for both malarial and non-malarial patients, but also, in the longer term, to less rapid build-up of parasitic resistance and higher acceptance and more appropriate usage of ACTs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

I gratefully acknowledge funding from the NIH/NICHD (5K01HD071949) and the Yale MacMillan Center Directors Award. I am deeply indebted to S. Patrick Kachur for his incredible hospitality at the CDC Mission in Dar es Salaam, Tanzania, and to the CDC and the Ifakara Health Institute for granting access to their household survey data. I thank Prashant Bharadwaj, Michael Boozer, Jessica Cohen, Rahul Deb, Pascaline Dupas, James Fenske, Adrian de la Garza, A. Mushfiq Mobarak, Anant Nyshadham, Mark Rosenzweig, Maher Said, T. Paul Schultz, and Christopher Udry for their advice and guidance. I thank participants in seminars at the World Bank, Johns Hopkins, Cornell, Columbia, Chicago, Emory, Yale, Stanford, RAND, Oxford, PACDEV, NEUDC, the NBER Health Care Summer Institute, NOVAFRICA, Berkeley, Penn, the African Studies Association Meetings, Pittsburgh/CMU, Michigan, and Maryland for helpful comments. Thanks also to Imran Rasul and four anonymous referees for greatly improving the paper through the review process. I am grateful to Rashid Khatib and Mujobu Moyo for their assistance in gathering data. All errors are my own.

References

- Adhvaryu A, Nyshadham A. Returns to treatment in the formal health care sector: Evidence from tanzania. Technical report, working paper. 2012
- Amexo M, Tollhurst R, Barnish G, Bates I. Malaria misdiagnosis: effects on the poor and vulnerable. *The Lancet*. 2004; 364(9448):1896–1898.

³¹In the latter study, Cohen et al. (2012) experimentally vary access to RDTs with the goal to better target ACTs to patients who need them most. Targeting does indeed improve significantly, though the revelation of disease information does not change all individuals' choices, suggesting market frictions, informational barriers, or other obstacles to adoption and acceptance of RDTs are at play.

- Armstrong Schellenberg J, Mukasa O, Abdulla S, Marchant T, Lengeler C, Kikumbih N, Mshinda H, Nathan R. The ifakara demographic surveillance system. INDEPTH Monograph Series: Demographic surveillance systems for assessing populations and their health in developing countries. 2001; 1:159–164.
- Arrow, KJ.; Panosian, C.; Gelband, H., et al. Saving lives, buying time: economics of malaria drugs in an age of resistance. National Academies Press; 2004.
- Ashraf N, Berry J, Shapiro JM. Can higher prices stimulate product use? evidence from a field experiment in zambia. *American Economic Review*. 2010; 100:2383–2413.
- Baird JK. Effectiveness of antimalarial drugs. *New England Journal of Medicine*. 2005; 352(15):1565–1577. [PubMed: 15829537]
- Bandiera O, Rasul I. Social networks and technology adoption in northern mozambique. *The Economic Journal*. 2006; 116(514):869–902.
- Banks JS, Sundaram RK. Switching costs and the gittins index. *Econometrica: Journal of the Econometric Society*. 1994:687–694.
- Björkman-Nyqvist M, Svensson J, Yanagizawa-Drott D. The market for (fake) antimalarial medicine: Evidence from uganda. Technical report, Working paper. 2013
- Bleakley H. Malaria eradication in the americas: A retrospective analysis of childhood exposure. *American Economic Journal: Applied Economics*. 2010:1–45.
- Chamley, C. Rational herds: Economic models of social learning. Cambridge University Press; 2004.
- Chanda P, Hamainza B, Moonga HB, Chalwe V, Banda P, Pagnoni F. Relative costs and effectiveness of treating uncomplicated malaria in two rural districts in zambia: implications for nationwide scale-up of home-based management. *Malar J*. 2011; 10(159):151–152. [PubMed: 21645385]
- Cohen J, Dupas P. Free distribution or cost-sharing? evidence from a randomized malaria prevention experiment. *Quarterly Journal of Economics*. 2010; 125(1):1–45.
- Cohen J, Dupas P, Schaner SG. Price subsidies, diagnostic tests, and targeting of malaria treatment: evidence from a randomized controlled trial. Technical report, National Bureau of Economic Research. 2012
- Conley TG, Udry CR. Learning about a new technology: Pineapple in ghana. *The American Economic Review*. 2010:35–69.
- Cutler D, Fung W, Kremer M, Singhal M, Vogl T. Early-life malaria exposure and adult outcomes: Evidence from malaria eradication in india. *American Economic Journal: Applied Economics*. 2010:72–94.
- Dupas P. Short-run subsidies and long-run adoption of new health products: evidence from a field experiment. Technical report, National Bureau of Economic Research. 2013
- Foster AD, Rosenzweig MR. Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of political Economy*. 1995:1176–1209.
- Gertler P, Locay L, Sanderson W. Are user fees regressive?: The welfare implications of health care financing proposals in peru. *Journal of Econometrics*. 1987; 36(1):67–88.
- Hong SC. The burden of early exposure to malaria in the united states, 1850–1860: malnutrition and immune disorders. *The journal of economic history*. 2007; 67(04):1001–1035. [PubMed: 19081796]
- Jones, DM.; Gittins, JC. A dynamic allocation index for the sequential design of experiments. University of Cambridge: Department of Engineering; 1972.
- Kachur SP, Abdulla S, Barnes K, Mshinda H, Durrheim D, Kitua A, Bloland P. Letters to the editors. *Tropical Medicine & International Health*. 2001; 6(4):324–325. [PubMed: 11348524]
- Kachur SP, Khatib RA, Kaizer E, Fox SS, Abdulla SM, Bloland PB. Adherence to antimalarial combination therapy with sulfadoxine-pyrimethamine and artesunate in rural tanzania. *The American journal of tropical medicine and hygiene*. 2004; 71(6):715–722. [PubMed: 15642960]
- Klemick H, Leonard KL, Masatu MC. Defining access to health care: evidence on the importance of quality and distance in rural tanzania. *American Journal of Agricultural Economics*. 2009; 91(2): 347–358.
- Kondor P, Ujhelyi G. Social learning with subjective communication and self-selection. Manuscript, Univ. Houston. 2005

- Kremer M, Miguel E. The illusion of sustainability. *The Quarterly Journal of Economics*. 2007; 122(3):1007–1065.
- Lindblade KA, Steinhardt L, Samuels A, Kachur SP, Slutsker L. The silent threat: asymptomatic parasitemia and malaria transmission. *Expert review of anti-infective therapy*. 2013; 11(6):623–639. [PubMed: 23750733]
- Lucas AM. Malaria eradication and educational attainment: evidence from paraguay and sri lanka. *American Economic Journal: Applied Economics*. 2010; 2(2):46–71. [PubMed: 23946866]
- Malisa AL, Pearce RJ, Mutayoba BM, Abdullah S, Mshinda H, Kachur PS, Bloland P, Roper C. The evolution of pyrimethamine resistant dhfr in plasmodium falciparum of south-eastern tanzania: comparing selection under sp alone vs sp+ artesunate combination. *Malaria journal*. 2011; 10(1): 317. [PubMed: 22029848]
- Manski CF. Identification of endogenous social effects: The reaction problem. *The review of economic studies*. 1993; 60(3):531–542.
- Munshi K. Social learning in a heterogeneous population: technology diffusion in the indian green revolution. *Journal of Development Economics*. 2004; 73(1):185–213.
- Mwagani E, Momburi D, Juma Z, Irema M, Masanja H. Rufiji dss, tanzania. *Population, Health, and Survival at INDEPTH Sites*. 2002
- Njau JD, Goodman CA, Kachur SP, Mulligan J, Munkondya JS, Mchomvu N, Abdulla S, Bloland P, Mills A. The costs of introducing artemisinin-based combination therapy: evidence from district-wide implementation in rural tanzania. *Malaria journal*. 2008; 7(1):4. [PubMed: 18179716]
- Reyburn H, Mbakilwa H, Mwangi R, Mwerinde O, Olomi R, Drakeley C, Whitty CJ. Rapid diagnostic tests compared with malaria microscopy for guiding outpatient treatment of febrile illness in tanzania: randomised trial. *BMJ: British Medical Journal*. 2007; 334(7590):403.
- Reyburn H, Mbatia R, Drakeley C, Carneiro I, Mwakasungula E, Mwerinde O, Saganda K, Shao J, Kitua A, Olomi R, et al. Overdiagnosis of malaria in patients with severe febrile illness in tanzania: a prospective study. *BMJ: British Medical Journal*. 2004; 329(7476):1212.
- Sachs J, Malaney P. The economic and social burden of malaria. *Nature*. 2002; 415(6872):680–685. [PubMed: 11832956]
- Somi MF, Butler JR, Vahid F, Njau J, Kachur SP, Abdulla S. Is there evidence for dual causation between malaria and socioeconomic status? findings from rural tanzania. *American Journal of Tropical Medicine and Hygiene*. 2007; 77(6):1020–1027. [PubMed: 18165515]
- Tarozzi A, Mahajan A, Blackburn B, Kopf D, Krishnan L, Yoong J. Micro-loans, insecticide-treated bednets and malaria: Evidence from a randomized controlled trial in orissa (india). *Economic Research Initiatives at Duke (ERID) Working Paper*. 2011; (104)
- Thwing J, Njau J, Goodman C, Munkondya J, Kahigwa E, Bloland P, Mkikima S, Mills A, Abdulla S, Kachur S. Drug dispensing practices during implementation of artemisinin-based combination therapy at health facilities in rural tanzania, 2002–2005. *Tropical Medicine & International Health*. 2011; 16(3):272–279. [PubMed: 21226795]
- Tougher S, Ye Y, Amuasi JH, Kourgueni IA, Thomson R, Goodman C, Mann AG, Ren R, Willey BA, Adegoke CA, et al. Effect of the affordable medicines facility? malaria (amfm) on the availability, price, and market share of quality-assured artemisinin-based combination therapies in seven countries: a before-and-after analysis of outlet survey data. *The Lancet*. 2012

A Proofs of propositions

Proposition 1 states that r_t is weakly decreasing in $(1 - m)$.

Proof. Consider two levels of misdiagnosis, $1 - m' > 1 - m''$. From above, we know that κ_i is increasing in $1 - m$; thus $\kappa_i|_{1 - m'} > \kappa_i|_{1 - m''}$. But this implies that for a given q_t , $i \in N_t$ $\mathbf{1}(q_t > \kappa_i|_{1 - m'}) < \mathbf{1}(q_t > \kappa_i|_{1 - m''})$, or $n_{1t}|_{1 - m'} < n_{1t}|_{1 - m''}$. Dividing by n_t on both sides, we obtain the desired result: $r_t|_{1 - m'} > r_t|_{1 - m''}$.

Proposition 2 states the following: When $p > p^*$, $\mathbb{E}(\lambda_{t+1} - \lambda_t | \theta = 1)$ is weakly decreasing in $1 - m$.

Proof. Using equations 2 and 4, $\mathbb{E}(n_1^g | \theta = 1)$ and $\mathbb{E}(n_1^b | \theta = 1)$ can be expressed as:

$$\mathbb{E}(n_1^g | \theta = 1) = (m + (1 - m)\tilde{p})n_{1t} \quad (22)$$

$$\mathbb{E}(n_1^b | \theta = 1) = (1 - m)(1 - \tilde{p})n_{1t}. \quad (23)$$

Substituting the above expected value expressions into equation 11, we obtain

$$\mathbb{E}(\lambda_{t+1} - \lambda_t | \theta = 1) = (x_g(m + (1 - m)\tilde{p}) + x_b(1 - m)(1 - \tilde{p}))n_{1t}. \quad (24)$$

Consider first the derivative of $\Gamma := x_g(m + (1 - m)\tilde{p}) + x_b(1 - m)(1 - \tilde{p})$ with respect to m in equation 24 above. This derivative can be expressed as

$$\frac{\partial \Gamma}{\partial m} = \frac{\partial x_g}{\partial m}(m + (1 - m)\tilde{p}) + \frac{\partial x_b}{\partial m}(1 - m)(1 - \tilde{p}) + (x_g - x_b)(1 - \tilde{p}). \quad (25)$$

Evaluating $\frac{\partial x_g}{\partial m}$ and $\frac{\partial x_b}{\partial m}$ and plugging the expressions into expression 25 above, we obtain

$$\frac{\partial \Gamma}{\partial m} = (\tilde{p} - p)(e^{x_g} - e^{x_b}) + (x_g - x_b)(1 - \tilde{p}). \quad (26)$$

Thus when $p > p^*$, $\frac{\partial \Gamma}{\partial m} > 0$, since $x_g > x_b$.

Now consider equation 24. Take two levels of misdiagnosis, $1 - m'' < 1 - m'$. The difference in $\mathbb{E}(\lambda_{t+1} - \lambda_t | \theta = 1)$ evaluated from $1 - m''$ (initial) to $1 - m'$ (final) is

$$\Delta \mathbb{E}(\lambda_{t+1} - \lambda_t | \theta = 1) = \Delta \Gamma n_{1t}|_{1-m''} + \Delta n_{1t} \Gamma|_{1-m''}. \quad (27)$$

We know $\Gamma < 0$, since $\frac{\partial \Gamma}{\partial m} > 0$ for $p > p^*$. Proposition 1 demonstrates that $n_{1t} > 0$. Finally, $n_{1t}|_{1-m''} > 0$ and $\Gamma|_{1-m''} > 0$. Thus, $\mathbb{E}(\lambda_{t+1} - \lambda_t | \theta = 1) > 0$, as we set out to show.

Figure 1.
Variation in proportion misdiagnosed across week of survey

Figure 2.
Histogram of residual variation in malaria misdiagnosis after removing village x time fixed effects

Figure 3.
Trends in health facility usage in treatment v. comparison districts over time

Table 1
Validating Model Assumption on Ordering of p and p -tilde

	p	p -tilde	Difference ($p - p$ -tilde)	Ha: diff < 0 Pr(T > t)
	(1)	(2)	(3)	(4)
<i>Definition 1</i>				
All	0.739 (0.440)	0.777 (0.417)	-0.037 (0.027)	0.084
Comparison districts only	0.712 (0.454)	0.773 (0.419)	-0.061 (0.040)	0.063
<i>Definition 2</i>				
All	0.749 (0.434)	0.777 (0.417)	-0.028 (0.025)	0.134
Comparison districts only	0.722 (0.449)	0.773 (0.419)	-0.052 (0.037)	0.082

Notes:

Definition 1. p is the proportion of individuals reporting acute illness in the 14 days before survey who had malaria, who did not visit the health facility, and who were well by the time of survey; p -tilde is the proportion of individuals reporting acute illness in the 14 days before survey who did not have malaria and who were well by the time of survey.

Definition 2. p is the proportion of individuals reporting acute illness in the 14 days before survey who had malaria, who did not visit the health facility if in the treatment district post-intervention, and who were well by the time of survey; p -tilde is the proportion of individuals reporting acute illness in the 14 days before survey who did not have malaria and who were well by the time of survey.

Table 2

Summary Statistics

	Whole sample		Treatment district (Rufiji)		Comparison districts (Kilombero & Ulanga)		Difference
	(1)	(2)	(3)	(4)			
Number of individuals (1)	16887	8146	8741				
Number of individuals reporting recent fever (2)	1738	894	844				
	Mean	SD	Mean	SD	Mean	SD	SE
Age (years)	24.343	21.364	25.186	22.888	23.558	19.808	-1.628***
Educational attainment of household head (years)	4.240	3.354	3.134	3.441	5.270	2.915	2.136***
Malaria positivity	0.222	0.416	0.217	0.412	0.226	0.418	0.009
Proportion reporting fever in 2 weeks preceding survey	0.123	0.328	0.126	0.332	0.120	0.325	-0.006
Outcomes Pre-intervention (2001) (among individuals reporting recent fever)							
Proportion who sought care at health facility (h=1)	0.219	0.414	0.232	0.423	0.204	0.404	-0.028
Proportion well at the time of survey	0.793	0.405	0.792	0.406	0.793	0.405	0.001
Length of illness (days)	3.217	1.925	3.233	2.066	3.199	1.757	-0.033
Length of illness (days) if h=1	4.215	2.069	4.290	2.334	4.114	1.664	-0.176
Length of illness (days) if h=0	2.922	1.778	2.888	1.848	2.959	1.703	0.071

Notes:

p<0.01.

Sample: (1) all surveyed individuals; (2) all individuals self-reporting acute fever (beginning less than 14 days before survey) for whom education, age, and lagged average misdiagnosis level are non-missing.

Table 3
Pre-intervention Health Facility Usage Across Treatment/Comparison Districts for
Individuals with High/Low Misdiagnosis Reference Groups

h =1 if individual sought care at a government or NGO health facility, 0 otherwise

	Ref. group misdiagnosis > median	Ref. group misdiagnosis < median	Difference (High - Low)
Treatment	0.217 (0.413)	0.240 (0.428)	-0.024 (0.041)
Comparison	0.203 (0.403)	0.208 (0.408)	-0.005 (0.041)
Difference (Treatment - Comparison)	-0.014 (0.039)	-0.032 (0.048)	0.019 (0.062)

Notes: ACT was not available at health facilities in either the treatment or comparison districts pre-intervention. Sample: all individuals reporting acute illness (illness beginning less than 14 days before survey) for whom education, age, and lagged average misdiagnosis level are non-missing. Standard deviations are reported in parentheses below means within the four sub-groups; standard errors are reported in parentheses below estimates of differences.

Table 4

Pre-intervention Trends

<i>Dependent variable:</i>	Age	Years of schooling	Sought care at a government or NGO health facility	Sick in last 14 days but well by date of survey	Length of illness (days)
	(1)	(2)	(3)	(2)	(3)
Treatment × post	0.935 (0.606)	-0.0241 (0.313)	-0.0169 (0.0643)	0.0442 (0.0400)	-0.313 (0.290)
Treatment	1.194** (0.512)	-1.732*** (0.288)	0.0482 (0.0441)	0.00120 (0.0299)	0.141 (0.205)
Post	-1.030** (0.432)	0.106 (0.134)	0.0978*** (0.0330)	0.00440 (0.0277)	0.0306 (0.235)
Number of observations	15,713	14,500	1,828	1,826	1,456

Notes:

- *** p<0.01,
- ** p<0.05,
- * p<0.1.

All specifications are estimated using OLS. Standard errors allow for correlation in the error within villages. Only pre-intervention survey rounds (2001 and 2002) are used here. The first two columns are on the whole pre-intervention sample, while columns (3)–(5) are on the sample of individuals reporting illnesses that began in the 2 weeks before survey. Column (5) is on the sample of such individuals who had completed their illnesses by the date of survey and thus could report the completed length of illness.

Table 5
Length of Illness (Days) among Health Facility Goers

	Pre-intervention (2001)	Post-intervention (2004)	Difference (Post - Pre)
<i>In the treatment district:</i>			
Malaria	4.963 (4.071)	3.313 (1.401)	-1.650 (1.057)
No malaria	4.976 (2.546)	4.442 (2.418)	-0.534 (0.371)
Difference (Malaria - No malaria)	-0.013 (0.660)	-1.130* (0.623)	-1.116 (0.931)
<i>For malaria-positive individuals:</i>			
Treatment district	4.963 (4.071)	3.313 (1.401)	-1.650 (1.057)
Comparison districts	3.923 (1.847)	3.067 (1.163)	-0.856 (0.575)
Difference (Treatment - Comparison)	1.040 (1.190)	0.246 (0.464)	-0.794 (1.374)

Notes:

*
p<0.1.

ACT was available at health facilities in the treatment district post-intervention. Sample: all individuals reporting acute illness (illness beginning less than 14 days before survey) and visiting government or NGO health facility in treatment district (Rufiji) for whom education, age, and lagged average misdiagnosis level are non-missing.

Table 6
Health Facility Usage Pre-/Post-intervention in Treatment/Comparison Districts

h =1 if individual sought care at a government or NGO health facility, 0 otherwise

	Pre-intervention (2001)	Post-intervention (2004)	Difference (Post - Pre)
Treatment	0.232 (0.423)	0.435 (0.496)	0.203*** (0.031)
Comparison	0.204 (0.404)	0.251 (0.434)	0.047 (0.029)
Difference (Treatment - Comparison)	0.028 (0.028)	0.184*** (0.032)	0.156*** (0.042)

Notes:

p<0.01.

ACT was available at health facilities in the treatment district post-intervention. Sample: all individuals reporting acute illness (illness beginning less than 14 days before survey) for whom education, age, and lagged average misdiagnosis level are non-missing.

Table 7
Association of Health Facility Usage (h) and Lagged Days Sick (D)

h =1 if individual sought care at a government or NGO health facility, 0 otherwise D = Lagged difference in days sick across h=1 and h=0

	Pre-intervention (2001)	Post-intervention (2004)	Difference (Post - Pre)
Treatment	0.0420* (0.0246)	-0.0539* (0.0284)	-0.0974*** (0.0372)
Comparison	-0.00549 (0.0329)	0.0738*** (0.0272)	0.0771* (0.0426)
Difference (Treatment - Comparison)	0.0499 (0.0406)	-0.126*** (0.0397)	-0.176*** (0.0568)

Notes:

p<0.01.

ACT was available at health facilities in the treatment district post-intervention. Each cell represents regressions of health facility usage on the lagged difference in days sick across h=1 and h=0, village × week FE, education, and a quadratic in age. Sample: all individuals reporting acute illness (illness beginning less than 14 days before survey) for whom education, age, and lagged average misdiagnosis level are non-missing.

Table 8

ACT Adoption by Reference Group's Level of Misdiagnosis

Dependent var: $h=1$ if individual sought care at a government or NGO health facility, $h=0$ otherwise

	Adoption by misdiagnosis level				
	Adoption				
		Continuous	Misdiagnosis < 25th percentile	Misdiagnosis < 50th percentile	Misdiagnosis < 75th percentile
	(1)	(2)	(3)	(4)	(5)
<i>Ref. group misdiagnosis level $\times$ Treatment $\times$ post</i>					
Treatment		-0.623* (0.357)	0.236 (0.154)	0.0812 (0.136)	-0.0458 (0.135)
Post		0.437 (0.297)	-0.162 (0.132)	-0.0432 (0.0984)	0.0218 (0.0993)
Treatment $\times$ post		0.529** (0.268)	-0.301** (0.127)	-0.0680 (0.101)	-0.00545 (0.0895)
Ref. group misdiagnosis level	0.181** (0.0736)	0.683** (0.313)	0.149** (0.0735)	0.131 (0.0814)	0.210* (0.113)
Fixed effects	-0.0386 (0.0797)	-0.416 (0.257)	0.216* (0.116)	0.0142 (0.0796)	-0.00199 (0.0660)
Number of observations	1,738	1,738	1,738	1,738	1,738

Village & survey round

Notes:

*** $p < 0.01$,

** $p < 0.05$,

* $p < 0.1$.

All specifications are estimated using OLS. Reference group for learning is all sick individuals in village within 6 weeks of date illness began. Standard errors are bootstrapped (500 repetitions) and allow for correlation in the error within villages. Specifications control for week-of-survey dummies to capture seasonal variation, dummies for categories of educational attainment of the household head, and a quadratic term in age. Column (2) uses a continuous reference group misdiagnosis variable. Columns (3), (4), and (5) use dummy variables for reference group misdiagnosis levels less than 25th, 50th, and 75th percentile, respectively.

Table 9
Learning Effect Estimates

Dependent var: $h=1$ if individual sought care at a government or NGO health facility, $h=0$ otherwise

	Baseline	Baseline + lagged differences in demographics	Baseline + lagged differences in demographics and symptoms
	(1)	(2)	(3)
$D \times$			
Treatment $\times$ post	-0.217*** (0.0686)	-0.250*** (0.0874)	-0.215** (0.0966)
Treatment	0.0945** (0.0455)	0.132** (0.0622)	0.105 (0.0690)
Post	0.0870* (0.0527)	0.121* (0.0723)	0.0919 (0.0840)
D	-0.0248 (0.0370)	-0.0512 (0.0539)	-0.0253 (0.0623)
Ref. group misdiagnosis level	0.0355 (0.112)	0.00869 (0.129)	0.0588 (0.146)
Fixed effects	<i>Village $\times$ year $\times$ week</i>		
Number of observations	1,680	1,586	1,554

Notes:

p<0.01,

**
p<0.05,

*
p<0.1.

D = Lagged difference in days sick across $h=1$ and $h=0$. All specifications are estimated using OLS. Reference group for learning is all sick individuals in village within 6 weeks of date illness began. Proportion of acutely ill individuals in reference group visiting health facility is included as a control. Standard errors are bootstrapped (500 repetitions) and allow for correlation in the error within villages. Additional controls are dummies for categories of educational attainment of the household head and a quadratic term in age. Column (1) reports results from the baseline learning effect specification. Columns (2) and (3) test robustness to addition of lagged differences (across facility users and non-users) in demographics (age, education and asset index) and number of symptoms, respectively.

Table 10
Learning Effect Estimates by Reference Group's Level of Misdiagnosis

Dependent var: $h=1$ if individual sought care at a government or NGO health facility, $h=0$ otherwise

	Continuous	Misdiagnosis < 25th percentile	Misdiagnosis < 50th percentile	Misdiagnosis < 75th percentile
	(1)	(2)	(3)	(4)
($D \times \text{treatment} \times \text{post}$) $\times$ Ref. group misdiagnosis level	0.609 *** (0.220)	-0.225 ** (0.100)	-0.194 *** (0.0715)	-0.240 ** (0.109)
$D \times \text{treatment} \times \text{post}$	-0.702 *** (0.218)	-0.147 ** (0.0653)	-0.119 * (0.0664)	-0.0860 (0.0777)
Ref. group misdiagnosis level	0.379 (0.499)	0.164 (0.275)	0.00617 (0.111)	-0.0353 (0.104)
Fixed effects	<i>Village $\times$ year $\times$ week</i>			
Number of observations	1,680	1,680	1,680	1,680

Notes:

 $p < 0.01$,

**
 $p < 0.05$,

*
 $p < 0.1$.

D = Lagged difference in days sick across $h=1$ and $h=0$. $D \times \text{treatment} \times \text{post}$ = Lagged difference in days sick interacted with treatment and post-intervention dummies. All specifications are estimated using OLS. All pairwise combinations of interactions and main effects not displayed above are included in specifications. Proportion visiting health facility in reference group is also included in all specifications. Column (1) uses a continuous reference group misdiagnosis variable. Columns (2), (3), and (4) use dummy variables for reference group misdiagnosis levels less than 25th, 50th, and 75th percentile, respectively. Standard errors are bootstrapped (500 repetitions) and allow for correlation in the error within villages. Additional controls are dummies for categories of educational attainment of the household head and a quadratic term in age.