Will a Global Subsidy of Artemisinin-Based Combination Treatment (ACT) for Malaria Delay the Emergence of Resistance and Save Lives?

Ramanan Laxminarayan1
Mead Over2
David L. Smith3


The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the view of the World Bank, its Executive Directors, or the countries they represent. Policy Research Working Papers are available online at http://econ.worldbank.org.

This study was initiated by the Finance and Resources Working Group of the Roll Back Malaria (RBM) Partnership, following its meeting on September 8-9, 2004 in Washington, DC. It was financed by the RBM Partnership Secretariat and managed by the Health, Nutrition and Population (HNP) Department of the World Bank. The study benefited greatly from an advisory group convened by Olusoji Adeyi, Chair of the Finance and Resources Working Group of RBM. Phil Musgrove provided extremely useful comments. The authors also acknowledge support from Suprotik Basu and Samantha Naidoo. Yvette Atkins provided administrative assistance.

1 Corresponding Author: Resources for the Future, 1616 P St, NW, Washington DC 20036, Ramanan@rff.org
2 The World Bank, 1818 H St, NW, Washington DC, meadover@worldbank.org
3 Fogarty International Center, NIH, smitdave@helix.nih.gov
Summary

Artemisinin-based combination treatments (ACTs) are seen as an important tool in the global effort to roll back malaria. With rapidly increasing parasite resistance to chloroquine in many parts of the world, there is greater international recognition of the need for both a different antimalarial and a coordinated malaria treatment strategy to ensure that resistance does not needlessly cut short the useful therapeutic life of any successor drug to chloroquine. The effectiveness of antimalarial drugs is a global public good, of particular value in malarious regions that also are among the most economically impoverished parts of the world. Inappropriate drug use in neighboring countries reduces the incentive of any given country to deploy drug regimens that may be rapidly undermined by resistance originating outside their borders. Therefore, a case can be made for globally coordinated action to protect the effectiveness of these valuable drugs. Translating this case to one for a global subsidy is not straightforward. On the one hand, in the absence of such a subsidy to ensure that ACTs are comparably priced to monotherapies, increasing monotherapy of artemisinin and other antimalarials that would be used along with artemisinin in ACT will hasten the demise of this drug. On the other hand, a global subsidy would greatly increase the use and potential misuse of ACTs and could result in resistance emerging at a more rapid rate.

This study finds that a subsidy to ACTs is likely to slow the rate of emergence of resistance to artemisinin and partner drugs, even if such a subsidy were to increase the use of ACTs significantly. This conclusion is robust to alternative assumptions regarding the responsiveness of demand to the lower price for ACTs and a wide range of epidemiological and economic parameters. However, the simulation results show that a subsidy for two or more ACT combinations is likely to be much more cost-effective than a subsidy to a single ACT. The only consideration is that the drugs used as partners to artemisinin be unrelated to each other and to artemisinin in mechanism of action and in genetic bases of resistance, so that a single mutation cannot encode resistance to both components. Such a subsidy program for ACTs, administered globally, that reduces reliance on any single combination, and discourages monotherapy, not only of artemisinin but of any effective antimalarial that could potentially be used as partner drug with artemisinin, is likely to be effective (and cost-effective) both in buying time for ACTs and in saving lives.
1. Introduction and Rationale for Study

In 2002, the Institute of Medicine’s (IOM) Board on Global Health convened a committee to examine the economics of alternative strategies to treat malaria given the reality that chloroquine and its successors were no longer effective in most malarious regions of the world. The IOM Committee was faced with the task of recommending a strategy that would enable the widespread use of effective drugs to fight the disease, while at the same time delaying the emergence of resistance to artemisinin compounds for as long as possible. In its final report, titled Saving Lives, Buying Time, the Committee noted that there were important reasons for the immediate introduction of artemisinin-based combination treatments (ACTs); notably the need to minimize the use of artemisinin-monotherapy (AMT) which could potentially engender rapid resistance to this valuable class of drugs, and the need to switch to a highly effective drug that could save lives that were being lost each year due to ineffective antimalarial treatments (Arrow, Panosian et al. 2004). Furthermore, there was a need for an internationally coordinated policy since the evolution of resistance to artemisinin-derivatives due to poor treatment choices made in any single country had the potential to compromise the effectiveness of ACTs everywhere else in the world. With these objectives in mind, the Committee recommended the institution of an international fund that would buy ACT from producers at $1 a dose and re-sell to anyone at one-tenth of that price, and would ensure a stable demand for ACTs and scale economies that could ultimately bring down the price of these drugs.

During subsequent deliberations at the World Bank in September 2004, it was pointed out that subsidies for any drug (including ACTs) create substantial benefits for the individual patients who take the drug, but typically entail negative externalities in the form of resistant pathogens: an international public “bad”. As the IOM committee itself had noted (on page 84), “if ACTs become as inexpensive to consumers as chloroquine is now, chances are they will be used as frequently as chloroquine, including over-use for febrile illnesses due to causes other than malaria. All overuse increases the probability that artemisinin-resistant parasites will arise and spread.”
Therefore, the IOM committee’s recommendation to encourage (through a subsidy), the greater use of ACTs would delay resistance only if the subsidized ACT would crowd out the practice of artemisinin monotherapy (AMT) or monotherapy of a partner drug (PMT), which could lead to more rapid development of drug resistance. More specifically, if the benefits of crowding out monotherapies outweighed the resistance related costs of greater use of ACTs, a global subsidy of ACTs would contribute to the expansion of a global public good, i.e., antimalarial efficacy.

There are a number of other reasons for subsidizing ACTs, as pointed out by the IOM Committee. An upfront subsidy would ensure a stable demand for ACTs, encourage suppliers to invest in scaled up production of raw materials and lower the price of ACTs. The centrally administered subsidy would lower the costs of standardization, procurement and distribution of the drugs. Moreover, a subsidy would achieve other objectives including discouraging counterfeit antimalarial drugs and allowing low-income countries to choose ACTs as their first line treatment without needing the promise of renewable funding from donors. None of these arguments is specific to ACTs or even malaria, however. They could be made just as easily for a global subsidy for sulfadoxine-pyrimethamine to treat malaria, for ciprofloxacin to treat diarrheal disease, or for anti-epileptic drugs to treat epilepsy. The global public goods based subsidy argument for ACTs is tied to the expectation that a low-cost provision of ACT would discourage AMT and PMT and would therefore delay the emergence of resistance to ACTs.

---

4 Support for the public goods argument could be bolstered if effective treatment with ACT combinations were to result in reductions in malaria transmission. Evidence of a transmission reduction effect has been found in south-east Asia (Nosten, Vugt et al. 2000) but has not been confirmed in Sub-Saharan Africa.  
5 The effectiveness of antimalarials (and other drugs to treat infectious diseases) is not a pure public good, which is defined as being non-excludable and non-rival in consumption. Rather, it is non-excludable, though rival in consumption (i.e., one individual’s use of antimalarial effectiveness diminishes its effectiveness for others) and can be classified as a global, common property resource in the same way as fisheries or climate stability. The extent to which antimalarial effectiveness is a public good also depends on the extent to which resistance is transmitted across borders. Since resistance to artemisinin compounds has not been detected, it is difficult to predict whether resistance will arise independently in a number of locations or if it will occur in a few select locations before spreading around the globe, as was the case with commonly encountered mutations that conferred resistance to chloroquine (pfcrt) and sulfadoxine-pyrimethamine (dhfr and dhps). If the latter were the case, then the resistance that could arise in a single country or region could jeopardize the value of artemisinin drugs in all countries where they will be used.
The basic question motivating the current study is whether the benefits of a global subsidy of ACTs through discouraging the use of monotherapy, both of artemisinin and of partner drugs to artemisinin, outweigh the drawbacks of potentially faster development of resistance to ACTs because of their widespread use.

2. Study Objectives

The primary objective of this report is to explore if and under what conditions on ACT use and on the counterfactual of AMT and PMT use, a global subsidy for ACTs would prolong the useful therapeutic life of artemisinin as an antimalarial. A broader question is whether a global subsidy for ACTs would save lives and avert morbidity (compared to doing nothing), and if so, at what cost? Underlying these objectives is the basic question of how the speed with which an artemisinin-resistant strain of *P. falciparum* develops and spreads depend upon the relative prices, substitution elasticities in consumption between ACT, AMT and partner drug monotherapies, and availabilities of these different drugs.

3. Methods

The question of the optimal introduction of ACTs, and the extent to which the effectiveness of ACTs is a global public good is addressed through a mathematical model of malaria transmission, resistance, immunity and economics. The analysis models the sensitivity of the optimal introduction of ACT to a counterfactual where AMT and PMT are used in a small proportion of patients, in the absence of ACT. Critical parameters here are the elasticities which characterize the response of consumer demand to the lower price of ACT, both in increasing their demand for ACT and in substituting ACT for AMT and PMT. Reasonable ranges of cross-elasticities of demand between ACT and AMT, as well as with other antimalarials are explored (see Appendix B).

A societal perspective is adopted for the purpose of determining optimal policy. Input parameters for cost of illness and treatment are drawn from the existing literature. For the purpose of this study, treatment costs were restricted to drug costs alone. Non-drug treatment costs were ignored, but including these would only dilute the importance of drug costs and should not change any of the results qualitatively. Cost of treatment
coverage is assumed to be linear in treatment costs and this may have some bearing on our qualitative results. One alternative is to use a non-linear function with decreasing marginal costs of coverage initially when coverage is small, and increasing marginal costs of coverage when all the easily accessibly patients have been reached and additional patients are more expensive to treat. Simulations are run for a ten-year planning horizon. For cost-effectiveness calculations, the streams of costs and health benefits are discounted at three percent.

Finally, the quality of ACT treatment, including the sensitivity and specificity of diagnosis, and treatment compliance is an important consideration but one that is outside the purview of this report. Compliance with different drug regimens may vary due to difference in the patient (i.e., education, household income, etc.), in the quality of service delivery because of differences in duration of treatment, side-effects, rapidity with which a cure is experienced and other factors; however, there is little quantitative evidence on how compliance is influenced; in general, the role of treatment compliance in the evolution of resistance remains poorly understood.

3.1 Malaria Model

An entomological-epidemiological model was developed to model transmission of malaria, and emergence and spread of single and multi-drug resistance to antimalarials. Details of the model are in Appendix A.

3.2 Treatment Demand

The demand for ACTs for malaria treatment depends on the efficacy of ACTs and its cost to patients. Other factors that may influence the decision to seek treatment include age, sex, geographical location, degree of malaria endemicity, proximity to a trained health care provider, and other socio-cultural factors and are assumed to be the same for

---

While there is a general notion that improved compliance lowers the likelihood of emergence of resistance, it is possible that this is true only for de novo resistance that emerges in infected individuals receiving antimalarials, and not for epidemic resistance, which is transmitted between individuals.
all treatments. The interaction of economic (behavioral) and epidemiological variables is self-evident. If ACTs are too expensive, they are less likely to be used widely and will have no impact on malaria deaths or evolution of resistance. Conversely, the greater the degree of malaria transmission, the faster and more lasting the immunity that is acquired to the disease, and for those individuals there will be a lower likelihood of symptomatic disease and less need to seek drug treatment (Hastings 1997).

Typically, only a fraction of patients infected with malaria parasites are symptomatic and are likely to be in the market for treatment. We assumed this fraction to be 20 percent in our model. The demand for a specific antimalarial treatment is a function of the quality-adjusted price of other antimalarials that are available. We can write the fraction of symptomatic malaria patients treated with antimalarial X (AMT) in the most general terms as a function of the price of AMT, $p_X$, price of PMT, $p_Y$, and the price of the combination ($p_{XY}$).

$$f_s = f(p_X, p_Y, p_{XY})$$

A subsidy that will decrease $p_{XY}$ has two effects. First, it elicits increased use of the antimalarial $XY$ (i.e., ACT) according to the assumed demand schedule. Figure 1 displays the assumed demand schedule relating $f_{xy}$ to the price of combination therapy (the curve labeled Elast = 3) and two alternatives (Elast = 4 and Elast = 2). Second, it encourages substitution away from monotherapies to the ACT combination $XY$. The net result of these two effects will typically be to increase overall use of antimalarials. Although there are no data to support the magnitude of these shifts in antimalarial use patterns in response to a subsidy to ACT, some plausible own and cross price elasticities are used (see Appendix B). In general, useful operational research to support the implementation of a global ACT subsidy may be to estimate the demand for antimalarial therapy, as a function of the perceived efficacy of therapy, price and other factors.$^7$

---

$^7$ Among the other relevant factors is the income level of the household. A benefit of the proposed subsidy is to increase access to antimalarials by the poorest. Yet these may be the same patients who are most likely
3.3 Parameter Values and Scenarios Considered

All entomological and epidemiological parameter values considered are in Table 1 and are justified elsewhere (Laxminarayan 2004). Model sensitivity was tested over a wide range of economic, epidemiological and entomological parameters. Price of AMT was assumed to be $1.00 per treatment course (3 days) and price of PMT was assumed to be $0.30 per treatment course. Price of ACT was $1.30 in the baseline (no-subsidy) case.

Table 2 describes the assumed levels of treatment coverage for different levels of subsidy (also see Appendix B). Under scenario A with no subsidy, 4% of all symptomatic cases of malaria are treated with AMT (drug X) and 16% are treated with the partner drug Y as monotherapy and 1% are treated with ACTs.

In scenario B, a partial subsidy (assumed to be $0.80 per dose) is provided to ACTs, which drops the cost to $0.50 per dose. As a consequence, under the baseline assumption on demand elasticity the use of AMT and PMT both drop by almost a third (to 3.4 % and 13.7 % respectively), while 15.1% of symptomatic infections are treated with ACTs. Overall coverage with antimalarials increases from 21% of symptomatic infections in the no-subsidy case to 32% under the partial subsidy. These changes in coverage levels correspond to an own price elasticity of ACT of -1.97 and a cross-price elasticity with respect to AMT and PMT of 0.17.

With a broader subsidy program (Scenario C) under which each dose of ACT is subsidized by $1.00 each (for a final price of $0.30 per treatment course), the use of ACTs increases to 45 % while the use of compromising monotherapy drops to 2.2 % and 8.9 % for AMT and PMT respectively. These coverage levels correspond to an own price elasticity of -1.53 for ACTs and a cross-price-elasticity of 0.46.

Scenario D is one under which there is a two-year delay in implementing a full subsidy program, during which time monotherapy with an artemisinin-derivative and the

---

Filmer 2002).
partner drug Y continues at levels seen in the no-subsidy scenario. We model this scenario using the same overall coverage response assumptions as in Scenario A for the first two years and Scenario C thereafter.

Two additional scenarios (E and F) were modeled. In E, a partial subsidy program is introduced for two types of ACTs (for instance, artesunate+amodiaquine and artesunate+Sulfadoxine-pyrimethamine). Overall coverage of ACT and monotherapies is identical to that in scenario B. In F, a full subsidy program is implemented for these two types of ACTs, and use of monotherapies is effectively discouraged as in scenario C.

4. Results

Figure 2 shows treatment failure rates (proportion of treated patients who fail to receive treatment) for a sample population of 1 million. Under the no-subsidy scenario (A), in orange, treatment failure rates rises rapidly but levels off after a short period of time. The lower treatment failure rate at equilibrium is because of the relatively fewer people who receive treatment without a subsidy program. The brown and red lines show projected treatment failure rates for a partial and full subsidy program respectively. Treatment failures rise more quickly under the partial subsidy program because of the use of compromising monotherapies alongside ACTs. With the full subsidy program, the rise in treatment failure rate is delayed because of lower compromising monotherapy but equilibrates at a relatively higher rate because of the greater proportion of patients receiving treatment. Failure rates under scenario D (delayed subsidy of ACTs) show that this is likely to be the worst strategy. Failure rates arise early, following the same path as the no subsidy case. The treatment failure rate drops at year 2 following the introduction of a full subsidy to ACT, but climbs back quickly because of the previous use of the artesunate and partner drug as monotherapy.

Treatment failure rates for the two other scenarios involving heterogeneous ACT use are also shown in purple and cyan. A partial subsidy program that increases use of two ACTs in which the partner drugs are unrelated results in a slower increase in treatment failure rates compared to scenarios A through D. A full subsidy program that covers two
ACTs with unrelated partner drugs does even better both in delaying the emergence of resistance and in the equilibrium level of resistance.

Table 3 shows malaria related deaths, deaths averted, discounted treatment and infection costs and discounted costs for averted mortality and morbidity for each of the six scenarios for baseline parameter values. Scenario A, with no subsidy, resulted in the most number of deaths over the 10-year horizon for the population of 1 million. A partial subsidy to ACTs of 80 cents per treatment course in scenario B was effective in lowering the number of deaths by roughly 3% at a cost of roughly $12,500 per death averted and lowering disease burden at an incremental cost-effectiveness ratio of $360 per DALY (disability-adjusted life year) averted. The predicted cost of averting a DALY in terms of the subsidy applied to increase ACT use was $271 at a rate of $0.80 per treatment.

A full subsidy of $1.00 per treatment course of ACT in scenario C could avert roughly 6,000 deaths over the 10-year horizon at a treatment cost of roughly $43,000 per death averted. The marginal cost-effectiveness ratio of this intervention was less favorable than for a partial subsidy, however, at $1,254 per DALY averted. The subsidy cost per DALY averted was $846\(^8\).

Scenario D, a delayed subsidy after some period of AMT and partner drug monotherapy resulted in only slightly fewer deaths compared to A, but at a greater cost (relative to a subsidy policy implemented today) of $53,494 per deaths averted.

Scenarios E and F were by far the most effective at lowering deaths and burden at a low cost. A partial subsidy to two ACT combinations lowered deaths by roughly 13% compared to the no-subsidy scenario at a cost of only $2,593 per death averted. The cost-effectiveness ratio of averted malaria burden was also low at $75 per DALY averted. When a full subsidy was applied to two ACT combinations in which the drugs used as

\(^8\) Note that in scenarios C and F the incremental subsidy costs per DALY averted exceeds the incremental treatment costs per DALY averted. This is possible because subsidy costs (unlike treatment costs) are zero for the baseline scenario A.
partners to artemisinin were unrelated to each other, the expected treatment cost of an averted death was $8,955 and the cost of averting a DALY was $259.

Next we tested the sensitivity of these results to the demand elasticity parameter (see table 4). If demand was relatively unresponsive to price, deaths averted was relatively lower with a partial subsidy because of the failure to drive out AMT and PMT. The treatment cost for averted morbidity and mortality was considerably low with the lower demand elasticity of –2 compared to elasticities of –3 and –4.

A full subsidy to ACTs that resulted in only a small increase in overall treatment coverage to 15% was more effective than a partial subsidy in saving lives but less cost-effective. If the same subsidy were able to increase coverage of ACT to 78% (which exceed the Abuja targets) under the demand elasticity assumption of –4, while completely discouraging any use of compromising monotherapy, treatment cost/death averted would increase to $81,963 and treatment cost/DALY averted would increase to $2,371, because of the greater likelihood of emergence of resistance to the ACT combination. Furthermore, unlike with a partial subsidy, more lives were saved with a full subsidy when the demand for treatment was less sensitive to price.

Two results remained qualitatively unchanged regardless of the responsiveness of demand to price. First, a delay in instituting a subsidy program was always inferior both in saving lives and in cost-effectiveness compared to all other scenarios. Second, the use of two combinations averted many more deaths and was more cost-effective than a use of a single combination.

5. **Implications for Policy**

The results indicate that a subsidy of ACTs is likely to be useful in extending the therapeutic life of artemisinin and partner drugs even if the subsidy results in expanding use of ACTs for reasonable ranges of own and cross-price elasticity of treatment demand. While the emergence of resistance is delayed by the use of combinations, it is highly sensitive to the choice of partner drug used with artemisinin in the combination. Using a
partner antimalarial to which resistance has already evolved or is likely to evolve very quickly is likely to lead to more rapid emergence of resistance to the combination. This finding is consistent with those from other recent studies (Watkins, Sibley et al. 2005).

Cotemporaneous use of a compromising monotherapy (either AMT or PMT) was not as detrimental to the emergence of resistance as one might have anticipated. Although resistance emerged faster with cotemporaneous compromising monotherapy (scenario B), the rate of growth of resistance was limited by the deployment of ACTs. This aspect of the model’s results should be explored further in future research.

The model results indicate that when treatment coverage is relatively more sensitive to drug price, moving from the partial to the full subsidy results in fewer averted deaths and diminished cost-effectiveness. This is a consequence of our motivating argument that greater use of ACTs could expedite the emergence of resistance. Depending on the level of demand elasticity (which would have to be empirically determined), it is possible to identify a level of subsidy at which the greatest number of deaths averted. It is important to recognize that a larger subsidy is not necessarily a good thing if it excessively encourages the use and misuse of ACT.

We found that a delayed introduction of a full subsidy program was possibly the worst scenario both from the perspective of averted deaths and cost-effectiveness. The delay would permit continued use of AMT and PMT and emergence of low-level resistance, which would then be magnified through intense selection pressure with the introduction of a full subsidy program.

An important implication of the results is that using a single combination in all regions places greater selection pressure for parasites to become resistant to that combination. Use of different combinations relieves the selection pressure for resistance to evolve to any single combination. In general, the idea of using the same ACT combination worldwide deserves serious reconsideration. Also, different ACT combinations may, if priced lower than monotherapies, be effective in driving out monotherapies by offering
consumers a choice of different antimalarials with different dosing schedules and other attributes.

A possible secondary benefit of increasing ACT use is the potentially reduced selection pressure for chloroquine resistance, but this not explicitly modeled in this report. ACT would clear CQ-resistant strains, and reduce the frequency of CQ resistance. After some time, CQ could become useful again. However, since CQ resistance could be re-introduced from another country, a switch to a CQ only strategy would be undesirable because resistance would probably return to its former frequency. A better strategy would be to continue to use ACTs and CQ in roughly equal amounts so that ACTs reduce the selection pressure on CQ, and CQ reduces the selection pressure on ACTs. Simultaneous use of two agents reduces the selective advantage of resistance to either agent and benefits both. These benefits arise because of the interactions with a third player—drug sensitive malaria. Neither CQ resistant nor ACT resistant malaria has an advantage over drug-sensitive malaria when no drug is used.

The same principles apply to the simultaneous use of two ACT agents, but in that case, the benefits of decreased selection pressure can additionally delay the time to the appearance of resistance to each ACT combination. In principle, contemporaneous use of three or more ACT agents would be even better than two because resistance to each antimalarial would be reduced by the use of every other antimalarial in competition relative to drug-sensitive malaria. We did not evaluate whether the use of an ACT triple (artemisinin combined with two partner drugs) would be more effective than the use of two ACT doubles. The relevant tradeoffs are the longer delay to the appearance of malaria that is resistant to all drugs in the combination weighed against the expense of subsidizing each additional agent in the combination and the added selection pressure from increasing the use of each agent in the combination.

A number of caveats apply to these results. First, they are only relevant from a qualitative standpoint and the quantitative aspects of deaths averted, time to emergence of resistance and other such outcomes are highly sensitive to model parameters. Also, further, more complex models are needed to explore specific phenomena, such as that of
ACTs on transmission potential. For instance, increased coverage using ACT could result in significantly reduced transmission both from a reduction in the total number of circulating parasites (included in this model), and also from lower gametocytes in circulation because of the effect of the artemunate in ACT (not considered here). Also, within the context of a particular structure of consumer demand, we find that the cost-effectiveness results are sensitive to the responsiveness of demand to price. If demand is highly responsive, additional lives are saved at the cost of an earlier spread of artemisinin-resistant strains of plasmodium. Within the assumptions used here, a global subsidy would be cost-effective across a range of plausible demand elasticities. However, the attractiveness of a global subsidy would be somewhat reduced when demand is highly responsive. The increased use of ACTs resulting in response to the subsidy would speed the development of resistance so much that more lives would be lost from ineffective antimalarials than would be gained from the increased utilization.

Second, optimal policy is likely to be affected by the potential availability of cheaper, synthetic, ACT combinations in the near future. Less expensive ACT combinations that may be available a few years from now may lower the value of introducing expensive versions of the combination drugs today if there are other drugs that we can rely on a few years from now. However, artemisinin monotherapy has the potential to introduce cross-resistance that may reduce the efficacy of these new, synthetic artemisinin-based combination treatments. In order to quantitatively consider the impact of the availability of synthetic artemisinin drugs on optimal policy today, one would need to know (with some degree of approximation) the expected date of arrival of new synthetic antimalarials and their likely efficacy.

Another unknown is the potential availability of a malaria vaccine. The impact of vaccine availability on optimal treatment policy today is similarly ambiguous. On the one hand, a vaccine lowers the benefits of an effective drug in the future since the caseload will be reduced. However, it is possible that an effective vaccine could in combination with effective treatment result in local eradication of the disease in some areas. Thus

---

9 Reduced transmission with effective treatment has been shown to be the case with the treatment of
there may be interesting non-linearities in the benefit function that could alter the costs of resistance and influence current use policies.

6. Conclusions

The basic objectives of the study were to examine whether a subsidy for ACTs would save lives and reduce malaria (compared to a counterfactual in which AMT and PMT would be used), and if so at what cost. Our summary conclusions are as follows.

1. We find that, regardless of the degree of responsiveness of antimalarial consumption to price, a subsidy to ACT would save lives even considering that a subsidy could hasten the arrival of parasite resistance to artemisinin-based drugs. Over a ten year planning horizon, even a partial subsidy that allows for some use of compromising monotherapy (AMT and PMT) saves lives and leads to a disease burden compared to a no-subsidy scenario. In fact, we find that for moderate to high elasticities of demand a larger subsidy is not necessarily more cost-effective than a smaller one and there exist optimal subsidy levels at which the marginal cost-effectiveness of treatment and the number of deaths averted are greatest.

There are two implications of this finding for moving ahead on ACT subsidies. First, achieving a high level of subsidy for ACTs should not stand in the way of instituting a more modest subsidy immediately. Second, the introduction of price subsidies should be accompanied by studies of demand response and of the quality of malaria service delivery, which would enable a more precise determination of the optimal subsidy level.

2. A delay (even by two years) in instituting a subsidy for ACTs would exacerbate resistance engendered by use of AMT and PMT prior to the introduction of a subsidy, and would lead to faster resistance to ACTs. Therefore, a subsidy to ACTs should be implemented sooner rather than later.

tuberculosis (Blower, Koelle et al. 1999).
Postponing action is likely to result in diminished cost-effectiveness of future treatment using ACTs.

3. A global subsidy for two or more ACTs is likely to be far more effective in delaying the onset of resistance and saving lives than reliance on a single or even a limited number of combinations. Such an intervention is highly cost effective at averting deaths and disease burden compared to a policy of doing nothing and is always better than subsidizing a single drug combination. The global subsidy program should support locally appropriate combination and treatment strategies, while keeping in mind the importance of transboundary spillovers of actions/decisions undertaken by any single country.

The rationale for considering ACT effectiveness to be a global public good is that resistance may arise in any single country and that use of AMT or even of ACT may serve to speed up the rate at which the resistant strain proliferates worldwide. A worst-case scenario would be one in which poor malaria policymaking results in the emergence of artemisinin resistance in a single country or region. This would compromise treatment in all countries where ACTs are introduced. Therefore, there is need for investment in surveillance so that the global subsidy program can be nimble in detecting emergence of resistance and changing drug combinations.

References


Appendix A: Mathematical Model of Malaria Transmission and Resistance – A Brief Description

We extend a simple compartmental model of malaria transmission (Koella 1991) to include the evolution of resistance to an antimalarial encoded at a single locus (Bonhoeffer, Lipsitch et al. 1997; Koella and Antia 2003; Laxminarayan 2004). \( X, Y, Z, XY, XZ \) and \( XYZ \) denote proportions of the population infected with a strain that is resistant to drugs \( X, Y \) and \( Z \), and drug combinations \( XY, XZ, YZ \) and \( XYZ \) respectively. In our model, artemisinin is represented by drug \( X \) and partner drugs are denoted by \( Y \) and \( Z \) respectively.

\( U \) denotes the uninfected population, \( W \) denotes the population infected with a wild-type or susceptible strain, and \( R \) denotes the immune population. Immunity would be lost at a constant per-capita rate, but it is boosted by re-exposure to malaria. In this model, immune individuals may or may not be infected with malaria, but they do not transmit it.

\[
\frac{\delta X}{\delta t} = h_X U + s_X f_X \rho W - \rho (f_Y + f_Z + f_{XY} + f_{XZ}) X - q_X X - \delta X
\]

\[
\frac{\delta Y}{\delta t} = h_Y U + s_Y f_Y \rho W - \rho (f_X + f_Z + f_{XY} + f_{XZ}) Y - q_Y Y - \delta Y
\]

\[
\frac{\delta Z}{\delta t} = h_Z U + s_Z f_Z \rho W - \rho (f_X + f_Y + f_{XY} + f_{XZ}) Z - q_Z Z - \delta Z
\]

\[
\frac{\delta XY}{\delta t} = h_{XY} U + \rho \left( s_X (f_Y + f_{XY}) + s_Y (f_X + f_{XY}) + f_{XY}s_{XY} W \right) - \rho (f_Z + f_{XZ}) XY - (q_{XY} + \delta) XY
\]

\[
\frac{\delta XZ}{\delta t} = h_{XZ} U + \rho \left( s_Z (f_X + f_{XZ}) + s_X (f_X + f_{XZ}) + f_{XZ}s_{XZ} W \right) - \rho (f_Y + f_{XY}) XZ - (q_{XZ} + \delta) XZ
\]
\begin{equation}
\frac{\delta YZ}{\delta t} = h_{yz}U + \rho(s_zyf_z + s_zyf_y) - \rho(f_x + f_{xy})YZ - (q_{yz} + \delta)YZ
\end{equation}

\begin{equation}
\frac{\delta XYZ}{\delta t} = h_{xyz}U + \rho(f_ys_yxZ + f_xsf_xY) + \rho(f_{xyz}s_{xyz}Y + f_{xy}s_{xy}Z) - (q_{xyz} + \delta)XYZ
\end{equation}

\begin{align*}
\frac{\delta R}{\delta t} &= q_wW + q_xX + q_yY + q_zZ + q_{xy}XY + q_{xz}XZ + q_{xyz}XYZ \\
&\quad - \gamma(h_r(1-U-R)) - \delta R
\end{align*}

\begin{equation}
\frac{\delta W}{\delta t} = h_wU - (f_x + f_y + f_{xy} + f_{xz})W - q_wW - \delta W
\end{equation}

\begin{equation}
I = W + X + Y + Z + XY + XZ + YZ + XYZ
\end{equation}

$f_x, f_y, f_{xy}, f_{xz}$ refer to fractions of symptomatic patients who are treated with drugs $X$ and $Y$, and combinations $XY$ and $XZ$ respectively. $\rho$ is the rate at which treated individuals return to the susceptible state. $q_x, q_y, q_{xy}, q_{xz}$ are rates at which patients recover spontaneously and acquire immunity. Resistance arises \textit{de novo} when an infected patient is treated with a drug or drug combination. In the model described, the rate of acquisition of primary resistance is denoted by $s_i$ where $i$ denotes the drug or drug combination. Resistance spreads when mosquitoes transmit a resistant parasite to the susceptible population. The net rate of transmission increases as the frequency of resistance increases and resistant parasites are selected in preference to wild-type parasites in the presence of antimalarials. Patients die at a rate $\delta$ and deaths are balanced by births into the uninfected population. Population is normalized to 1. Resistance imposes a fitness cost on the parasite, and is represented in the model by the differential rates of spontaneous resolution of infection in the absence of an antimalarial.

As in previous literature in this area, we assume that the mosquito dynamics operate on a much faster time-scale than the human dynamics, so that the mosquito population
can be considered to be at equilibrium with respect to changes in the human population, and its dynamics can be collapsed into the inoculation rate (Koella 1991). The derivation of the inoculation rate is as follows and is provided in Koella (1991). The inoculation rate is given by:

\[ h = ma_b w \]

where \( m, a, \) and \( b \) are as defined earlier and \( w \) is the proportion of mosquitoes with sporozoites in their salivary glands. The population of mosquitoes comprises \( w, \) a proportion, \( v, \) of mosquitoes that are infected but have not yet developed sporozoites, and uninfected proportion given by \( 1 - v - w. \) The equations of motion for \( v \) and \( w \) are provided in Koella (1991). If the mosquito population is considered to be at equilibrium, then we can solve for the steady state values of \( v \) and \( w \) as

\[ v = \frac{ab_2 e^{-\mu t} (e^{-\mu t} - 1)}{\mu + a l} I \]
\[ w = \frac{ab_2 e^{-\mu t} I}{\mu + a l} \]

Substituting back into expression (A.1) and indexing by each drug or drug combination, we get

\[ h_i = ma^2 b b_2 e^{-\mu t} \frac{I}{\mu + a l} \]

where \( i = W, X, Y, Z, XY, XZ, XYZ, \) \( a \) is the biting rate (number of bites per female mosquito per night), \( b_i \) is the infectiousness of humans to mosquitoes, \( b_2 \) is the susceptibility of mosquitoes to humans, \( m \) is the mosquito density (number of mosquitoes
per human), $\tau$ is the incubation period of parasites in the mosquito, and $\mu$ is the mortality of mosquitoes. $I$ denotes the proportion of infected individuals in the human population (as defined in equation 1.10), $y_{w}$ is the fraction of infected patients with a susceptible strain, and $y_{r}$ carry a resistant strain ($y_{w} + y_{r} = y$). Rate of loss of immunity, $\gamma (h)$ is a function of the transmission coefficient and can be represented by

$$\gamma (h) = \frac{he^{-h\theta}}{1-e^{-h\theta}}$$

where $\theta$ is the average number of years without exposure after which immunity is lost and the person returns to the susceptible state (Aron 1988). Infected individuals who are treated successfully (because they carry a sensitive parasite) return to the susceptible state. There is some evidence that the benefit of effective treatment is accompanied by a loss of immunity (Pringle and Avery-Jones 1966; Cornille-Broger, Mathews et al. 1978). Treatment, therefore, retards expansion of the immune class. The use of ACTs or some similar effective treatment strategy does not alter transmission intensity in this model, but reduces the number of parasites in circulation by reducing the duration of illness.

Modeling the evolution of resistance to malaria is complicated by a number of factors including the fact that there are a number of different strains of the malarial parasite, which could slow down the evolution of drug resistance, superinfection of hosts with multiple parasite strains, and poor understanding of immunity to malaria. These represent some important limitations of the current model. The qualitative implications of these limitations are discussed elsewhere (Koella and Antia 2003).
Appendix B: A Model of Demand for Antimalarial Drugs

To our knowledge no empirical studies of the responsiveness of the demand for antimalarial medications to their prices are available. One study used a willingness-to-pay approach to estimate the elasticity of the demand for a vaccine to prevent malaria in the range -.08 to -.64. (Whittington, Cropper et al. 2000). Intuition suggests that the elasticity of demand for treatment should be smaller than that for prevention, because treatment relieves acute symptoms, while prevention avoids uncertain future illness. However, the safest approach will be to study the sensitivity of simulation results to a range of elasticity assumptions.

In order to simplify our model of the demand for antimalarial medication, we assume that those with symptoms must choose among four available treatment options: artemisinin monotherapy (AMT), partner drug monotherapy (PMT), combination therapy (ACT) and self-treatment. In order to structure and systematize the sensitivity analysis, we adopt a simple, often-used structural form for the demand for alternative medications, which was originally developed by McFadden to model consumer choice among transportation modes (McFadden 1974). The relationship between the proportion of those with symptoms who choose treatment option $i$ and the prices of that and the other alternative options is given by the equation:

\[
 f_i = \frac{e^{(\alpha_X + \beta \log(P_i))}}{\sum_j e^{(\alpha_j + \beta \log(P_j))}} \quad j \in \{X, Y, XY, Self\}
\]

where $f_i$ is the proportion (of those with symptoms) demanding treatment $i$, $P_i$ is the price of option $i$, the four parameters $\alpha_X$, $\alpha_Y$, $\alpha_{XY}$ and $\alpha_{Self}$ capture idiosyncratic features of the four options that affect demand, such as their perceived quality or distance from the consumer, and the parameter $\beta$ captures the responsiveness of demand to price. For the present situation, where little is known about the nature of demand for these drugs, it is useful that responsiveness is captured by only a single parameter.
Since the results of the modeling effort are likely to vary by the degree of demand responsiveness, we select three values of $\beta$ for which to run the models. The central assumption is that $\beta = -3$. As shown in Table 2, this value of $\beta$ corresponds to an arc elasticity of demand between -1.5 (Scenario B) and -2 (Scenario C). While these elasticities seem high, recall that some of the increased demand for ACT in response to a decrease in its price comes from a decrease in the use of the monotherapies. The arc cross price elasticities in this model range between 0.2 (Scenario B) and 0.5 (Scenario C). When taking these substitution possibilities into account, the net increase in antimalarial use in response to a ten percent decrease in the price of ACT will be between two and five percent.

Figure 1 presents the demand curves that are consistent with the hypothetical structure of demand for the three posited values of $\beta$. The horizontal axis is a logarithmic scaling of the share of those with symptomatic malarial (or febrile illness) who select ACT rather than another antimalarial or self-treatment. These curves display the downward slope typical of all demand curves, as price decreases cause an increase in demand. The magnitude of that increase differs across the three demand curves, with the flattest one, which corresponds to $\beta = -4$, being the most elastic or responsive. Two convenient attributes of this demand structure are that the proportion of people choosing each of the four alternatives must always add to one and no individual proportion can equal or exceed one. The latter feature is illustrated in the figure by the way the vertical right-hand axis bounds all the demand curves even as the price of ACT approaches zero.
Figure 1. Demand curves for ACT
Figure 2: Time paths of treatment failure for different scenarios over ten years
Table 1: Parameter values for simulations

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>Biting rate (number of bites per female mosquito)</td>
<td>100 per year</td>
</tr>
<tr>
<td>$b_1$</td>
<td>Proportion of infectious bites on humans that produce an infection</td>
<td>0.8</td>
</tr>
<tr>
<td>$b_2$</td>
<td>Proportion of bites by susceptible mosquitoes on infected humans that produce an infection</td>
<td>0.8</td>
</tr>
<tr>
<td>$m$</td>
<td>Number of female mosquitoes per human host</td>
<td>0.1</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Mosquito mortality rate</td>
<td>36 per year</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Incubation period of parasites in mosquito</td>
<td>10 days</td>
</tr>
<tr>
<td>$q_w$</td>
<td>Spontaneous rate of recovery from sensitive strain</td>
<td>0.7/year</td>
</tr>
<tr>
<td>$q_X, q_Y$</td>
<td>Spontaneous rate of recovery from single drug-resistant strain</td>
<td>0.7/year</td>
</tr>
<tr>
<td>$q_{XY}, q_{XZ}$</td>
<td>Spontaneous rate of recovery from strain resistant to two drug</td>
<td>0.84/year</td>
</tr>
<tr>
<td>$q_{XZ}$</td>
<td>Spontaneous rate of recovery from multi-drug resistant strain</td>
<td>1/year</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Malaria attributable mortality</td>
<td>0.2/year</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Excess rate of recovery of treated individuals</td>
<td>12/year</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Rate of loss of immunity</td>
<td>2/year</td>
</tr>
<tr>
<td>$s_X$</td>
<td>Rate of emergence of <em>de novo</em> resistance to artemisinin monotherapy</td>
<td>$10^{-12}$</td>
</tr>
<tr>
<td>$s_Y, s_Z$</td>
<td>Rate of emergence of de novo resistance to partner drug monotherapies</td>
<td>$10^{-3}$</td>
</tr>
<tr>
<td>$s_{XY}, s_{XZ}$</td>
<td>Rate of emergence of de novo resistance to combinations</td>
<td>$10^{-15}$</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Underlying population death rate</td>
<td>$3.3 \times 10^{-5}$</td>
</tr>
</tbody>
</table>

### Table 2: Treatment coverage for different scenarios

<table>
<thead>
<tr>
<th></th>
<th>Scenario A: No Subsidy</th>
<th>Scenario B: Partial Subsidy (ACT with compromising monotherapy)</th>
<th>Scenario C: Full Subsidy (ACT Now)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Price of ACT, Pxy</strong></td>
<td>1.3</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td><strong>Elasticity parameter</strong></td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
</tr>
<tr>
<td><strong>Coverage shares:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACT</td>
<td>0.01</td>
<td>0.316</td>
<td>0.781</td>
</tr>
<tr>
<td>Artemisinin monotherapy</td>
<td>0.04</td>
<td>0.028</td>
<td>0.009</td>
</tr>
<tr>
<td>Partner monotherapy</td>
<td>0.16</td>
<td>0.111</td>
<td>0.035</td>
</tr>
<tr>
<td>Self treatment</td>
<td>0.79</td>
<td>0.546</td>
<td>0.175</td>
</tr>
<tr>
<td><strong>Arc elasticities of demand</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Own price elasticity of ACT</td>
<td>NA</td>
<td>-2.11</td>
<td>-1.56</td>
</tr>
<tr>
<td>Cross price elasticities</td>
<td>NA</td>
<td>0.41</td>
<td>1.02</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Scenario D: Full Subsidy (ACT Later)</th>
<th>Scenario E: Partial Subsidy for two ACTs</th>
<th>Scenario F: Full Subsidy for two ACTs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Price of ACT, Pxy</strong></td>
<td>0.3</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td><strong>Price of ACT, Pxz</strong></td>
<td>NA</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td><strong>Coverage shares:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACT (XY)</td>
<td>0.390</td>
<td>0.158</td>
<td>0.390</td>
</tr>
<tr>
<td>ACT (XZ)</td>
<td>0.390</td>
<td>0.158</td>
<td>0.390</td>
</tr>
<tr>
<td>Artemisinin monotherapy</td>
<td>0.009</td>
<td>0.028</td>
<td>0.009</td>
</tr>
<tr>
<td>Partner x monotherapy</td>
<td>0.018</td>
<td>0.055</td>
<td>0.018</td>
</tr>
<tr>
<td>Partner y monotherapy</td>
<td>0.018</td>
<td>0.055</td>
<td>0.018</td>
</tr>
<tr>
<td>Self treatment</td>
<td>0.175</td>
<td>0.546</td>
<td>0.175</td>
</tr>
</tbody>
</table>

**Note:** Where the coverage share, $f_i$, denotes the proportion of the infected population that is treated with drug $i$, $X$ is an artemisinin derivative, $Y$ and $Z$ are partner drugs that can also be used as monotherapy. See Appendix B for an explanation of the elasticity parameter. NA: Not applicable.
Table 3: Summary results for the six scenarios for 10 year planning horizon and 1 million population for baseline elasticity (-3)

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Deaths Averted (Compared to Scenario A)</th>
<th>Discounted Treatment Costs ($)</th>
<th>Discounted Cost of Infection ($)</th>
<th>Treatment $/Death Averted</th>
<th>Treatment $/DALY averted</th>
<th>Subsidy $/DALY averted</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>233,584</td>
<td>55,857,080</td>
<td>66,364,850</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>227,114</td>
<td>140,980,400</td>
<td>64,082,000</td>
<td>12,429</td>
<td>360</td>
<td>271</td>
</tr>
<tr>
<td>C</td>
<td>227,500</td>
<td>338,090,700</td>
<td>64,194,440</td>
<td>43,346</td>
<td>1,254</td>
<td>856</td>
</tr>
<tr>
<td>D</td>
<td>228,155</td>
<td>340,190,300</td>
<td>64,593,100</td>
<td>53,494</td>
<td>1,548</td>
<td>1,055</td>
</tr>
<tr>
<td>E</td>
<td>203,902</td>
<td>126,130,300</td>
<td>57,331,940</td>
<td>2,593</td>
<td>75</td>
<td>61</td>
</tr>
<tr>
<td>F</td>
<td>203,562</td>
<td>301,387,400</td>
<td>57,225,460</td>
<td>8,955</td>
<td>259</td>
<td>181</td>
</tr>
</tbody>
</table>

Table 4: Sensitivity analysis with respect to demand elasticity for the six scenarios for 10 year planning horizon and 1 million population

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Deaths Averted (Compared to A)</th>
<th>Treatment $/Death Averted</th>
<th>Treatment $/DALY averted</th>
<th>Subsidy $/DALY averted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elasticity</td>
<td>-2</td>
<td>-4</td>
<td>-2</td>
<td>-4</td>
</tr>
<tr>
<td>B</td>
<td>4,784</td>
<td>6,338</td>
<td>6,392</td>
<td>29,633</td>
</tr>
<tr>
<td>C</td>
<td>6,507</td>
<td>5,644</td>
<td>13,585</td>
<td>81,963</td>
</tr>
<tr>
<td>D</td>
<td>4,728</td>
<td>5,459</td>
<td>20,717</td>
<td>93,777</td>
</tr>
<tr>
<td>E</td>
<td>27,525</td>
<td>29,792</td>
<td>926</td>
<td>6,354</td>
</tr>
<tr>
<td>F</td>
<td>29,638</td>
<td>30,628</td>
<td>2,881</td>
<td>15,602</td>
</tr>
</tbody>
</table>