High-throughput analysis of antimalarial susceptibility data by the WorldWide Antimalarial Resistance Network (WWARN) in vitro analysis and reporting tool.


To cite this version:

HAL Id: pasteur-00861904
https://hal-riip.archives-ouvertes.fr/pasteur-00861904
Submitted on 13 Sep 2013
High-Throughput Analysis of Antimalarial Susceptibility Data by the WorldWide Antimalarial Resistance Network (WWARN) In Vitro Analysis and Reporting Tool


Published Ahead of Print 22 April 2013.

Updated information and services can be found at: http://aac.asm.org/content/57/7/3121

These include:

SUPPLEMENTAL MATERIAL
Supplemental material

REFERENCES
This article cites 51 articles, 31 of which can be accessed free at: http://aac.asm.org/content/57/7/3121#ref-list-1

CONTENT ALERTS
Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), more»

Information about commercial reprint orders: http://journals.asm.org/site/misc/reprints.xhtml
To subscribe to to another ASM Journal go to: http://journals.asm.org/site/subscriptions/
High-Throughput Analysis of Antimalarial Susceptibility Data by the WorldWide Antimalarial Resistance Network (WWARN) In Vitro Analysis and Reporting Tool

Charles J. Woodrow,a,b Sabina Dahlström,c Richard Cooksey,a Jennifer A. Flegg,a Hervé Le Nagard,9 France Mentré,d Claribel Murillo,a Didier Ménard,f François Nosten,g Kanlaya Sriprawat,g Lise Musset,h Neils B. Quashie,i,j Rick M. Fairhurst,k Sam L. Nsobya,m Veronique Sinou,n Harald Noedl,o Bruno Pradines,p,q,r Jacob D. Johnson,s Philippe J. Guerin,a Carol H. Sibley,a,t Jacques Le Bras,c,u,v

WorldWide Antimalarial Resistance Network (WWARN), Centre for Tropical Medicine, Nuffield Department of Clinical Medicine, University of Oxford, Churchill Hospital, Oxford, United Kingdomq; Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailandp; WorldWide Antimalarial Resistance Network (WWARN) In Vitro Module, Hôpital Bichat-Claude Bernard, Paris, Francev; INSERM, Université Paris Diderot, UMR-S 738, Paris, Francev; International Center for Medical Research and Training (CIDEIM), Cali, Colombiaa; Malaria Molecular Epidemiology Unit, Institut Pasteur in Cambodia, Phnom Penh, Cambodiaf; Shoklo Malaria Research Unit, Mae Sod, Tak, Thailandf; Laboratoire de Parasitologie, Centre National de Référence de la Chimiorésistance du Paludisme pour la Région Antilles-Guyane, Institut Pasteur de la Guyane, Cayenne, French Guiana; Centre for Tropical Clinical Pharmacology and Therapeutics, University of Ghana Medical School; and Epidemiology Department, Noguchi Memorial Institute for Medical Research, University of Ghana, Accra, Ghana; Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, USAa; National Center for Parasitology, Entomology and Malaria Control, Phnom Penh, Cambodia; Department of Medicine, Makerere University, Kampala, Uganda; aix-Marseille Université, UMR-MD3, Marseille, France; Institut of Specific Prophylaxis and Tropical Medicine, Medical University of Vienna, Vienna, Austria; Unité de Parasitologie, Institut de Recherche Biomédicale des Armées, Marseille, France; Aix Marseille Université, URMI TE, UM63, CNRS 7278, IRD 198, INSERM 1095, Marseille, Franced; Centre National de Référence du Paludisme, Marseille, France; U.S. Army Medical Research Unit—Kenya, Walter Reed Project, Kismu, Kenya; Department of Genome Sciences, University of Washington, Seattle, Washington, USA; Université Paris Descartes (Paris 5), Paris, France; and Institut de Recherche pour le Développement (IRD), UMR 216, Paris, France.

Assessment of in vitro susceptibility is a fundamental component of antimalarial surveillance studies, but wide variations in the measurement of parasite growth and the calculation of inhibitory constants make comparisons of data from different laboratories difficult. Here we describe a Web-based, high-throughput in vitro analysis and reporting tool (IVART) generating inhibitory constants for large data sets. Fourteen primary data sets examining laboratory-determined susceptibility to artemisinin derivatives and artemisinin combination therapy partner drugs were collated from 11 laboratories. Drug concentrations associated with half-maximal inhibition of growth (IC50s) were determined by a modified sigmoid Emax model-fitting algorithm, allowing standardized analysis of 7,350 concentration-inhibition assays involving 1,592 isolates. Examination of concentration-inhibition data revealed evidence of apparent paradoxical growth at high concentrations of nonartemisinin drugs, supporting amendment of the method for calculating the maximal drug effect in each assay. Criteria for defining more-reliable IC50s based on estimated confidence intervals and growth ratios improved correlation coefficients for the drug pairs mefloquine-quinine and chloroquine-desethylamodiaquine in 9 of 11 and 8 of 8 data sets, respectively. Further analysis showed that maximal drug inhibition was higher for artemisinins than for other drugs, particularly in ELISA (enzyme-linked immunosorbent assay)-based assays, a finding consistent with the earlier onset of action of these drugs in the parasite life cycle. This is the first high-throughput analytical approach to apply consistent constraints and reliability criteria to large, diverse antimalarial susceptibility data sets. The data also illustrate the distinct biological properties of artemisinins and underline the need to apply more sensitive approaches to assessing in vitro susceptibility to these drugs.

The mission of the WorldWide Antimalarial Resistance Network (WWARN) is to enhance the quality, quantity, and geographic extent of drug susceptibility data available to the malaria control community via a global data repository. Laboratory-based assessment of parasites in culture (“in vitro”) enables measurement of the intrinsic drug susceptibility of Plasmodium falciparum without the confounding effects of host pharmacokinetics, immunity, and genetics (1). Parasites with reduced antimalarial susceptibilities can be established in continuous culture, allowing the investigation of molecular mechanisms of resistance as well as the assessment of susceptibility to other antimalarial agents (2). In an era when artemisinin combination therapies (ACTs) are recommended treatment for falciparum malaria worldwide, additional considerations apply. While the use of a combination is beneficial in therapeutic terms, resistance to either partner alone can develop without an immediate reduction in clinical treatment efficacy. Assessment of drug susceptibility in parasites isolated directly from patients provides an opportunity to detect resistance to each individual drug at a relatively early stage, potentially allowing appropriate action before clinically relevant drug failure occurs (1).
One challenge facing the in vitro field is that culture-based assessment of parasite susceptibility has undergone a natural evo-

tion since techniques for studying chloroquine (CQ) resistance were established more than 4 decades ago (3). The basic measure-

ment of drug susceptibility is the growth of parasites in the presence of a range of concentrations of a given drug, expressed as the con-
centration of the drug needed to suppress growth to 50% of that observed in the absence of the drug (50% inhibitory concentra-

tion [IC50]). A wide variety of readout methods for assessing parasite growth have been described (4, 5), including microscopic
 assessment (6), incorporation of radiolabeled hypoxanthine (7), production of the highly expressed P. falciparum proteins lactate
 dehydrogenase (LDH) (8, 9) and histidine-rich protein 2 (HRP2) (10), and methods involving DNA detection, such as SYBR green
 fluorescence (11, 12) and flow cytometry (13). This variety of techniques reflects practical and financial considerations that de-
 fines a specific need for different assays in different settings.

All methods for phenotyping parasite responses outside the host are to some degree surrogates for in vivo phenomena, and
 although each new technique has been validated against a standard (generally hypoxanthine incorporation), differences be-
 tween the methods clearly exist. Longitudinal estimates from the same lab measured consistently over time are still informative (2, 4),
 but the comparison of data from different laboratories remains a major challenge. The use of control reference clones holds the
 potential to reduce this problem (2) but has rarely been achieved over a substantive time frame.

Differences in computational methods also compromise the comparison of results from different laboratories. Investigators
calculate inhibitory constants by a variety of means, including algorithms within software packages and freely available tools
based on log probit (14), polynomial (10), and sigmoid inhibition (15) models. In addition, some assays exhibiting poor growth are
misleading and should not be used as a basis for defining drug resistance (4, 15). Standardized methods to address these issues
have been reported on occasion (16), but in general, the classifi-
ation of concentration-inhibition curves remains a time-con-\nsuming and potentially subjective process involving visual inspec-
tion of individual curves. The need to examine parasites isolated directly from patients precludes repeated studies of individual
parasite isolates, further compounding these difficulties.

This work describes the development of an in vitro analysis and reporting tool (IVART) capable of producing inhibitory constants
for large in vitro data sets in a rapid, automated manner via a Web interface. We first sought biological evidence to better define key
 elements of this tool and therefore collated a wide-ranging collection of raw data obtained in a variety of global locations, generat-
ing perhaps the most diverse data set of this type so far assembled. Systematic examination of concentration-inhibition data from
this range of different assay readouts and drugs informed the choice of appropriate constraints for use in curve fitting. Criteria
for defining a core subset of more-reliable assays were tested by examining correlation coefficients for IC50s from pairs of drugs.

The data also yielded biological insights into the distinct properties of artemisinin derivatives.

### MATERIALS AND METHODS

**Data sets used.** Primary data sets describing the growth of *P. falciparum* in culture at varying drug concentrations in individual wells of 96-well plates were collated, allowing the comparison of data obtained using various assay methods (microscopic assessment, radiolabeled hypoxanthine uptake, HRP2 and Plasmodium LDH [pLDH] enzyme-linked im-

munoassorbent assays [ELISA], and SYBR green). Three groups of drugs were studied: (i) artemisinins found in ACTs, consisting of dihydroart-
emin (DHA), artemether (AM), and artesunate (AS); (ii) ACT partner drugs (or active metabolites) desethylamodiaquine (DQ), lumefantrine (LUM), mefloquine (MQ), and piperaquine (PIP); and (iii) chloroquine (CQ) and quinine (QN), drugs that are no longer recommended for first-
line treatment of *P. falciparum* malaria. Data sets describing at least 40 parasite isolates were considered large enough to be included in this anal-

ysis. Fourteen data sets from 11 laboratories fulfilled these criteria (Table 1). The laboratory methodologies for many of these studies have been described previously (17–27). The primary growth outputs from assays

### TABLE 1 Data sets examined

<table>
<thead>
<tr>
<th>Data set</th>
<th>Method</th>
<th>Location</th>
<th>No. of isolates</th>
<th>Yr</th>
<th>Sample-to-culture delay (h)</th>
<th>Artemisinin(s)</th>
<th>Partner(s)</th>
<th>Other drug(s)</th>
<th>No. of drug-free controls per plate</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3H (0)</td>
<td>Madagascar</td>
<td>315</td>
<td>2006–2007</td>
<td>24–48</td>
<td>DHA</td>
<td>MQ, LUM</td>
<td>MQ, DQ</td>
<td>12, 3 or 12</td>
<td>17</td>
</tr>
<tr>
<td>B</td>
<td>3H (0)</td>
<td>Travelers</td>
<td>421</td>
<td>2010</td>
<td>4–48</td>
<td>DHA</td>
<td>MQ, LUM, MQ</td>
<td>MQ, DQ</td>
<td>12, 4–8</td>
<td>18</td>
</tr>
<tr>
<td>C</td>
<td>3H (0)</td>
<td>French Guiana</td>
<td>83</td>
<td>2008</td>
<td>12–48</td>
<td>AM, AS, DHA</td>
<td>MQ, LUM, MQ</td>
<td>MQ, DQ</td>
<td>9, 2–4</td>
<td>19</td>
</tr>
<tr>
<td>D</td>
<td>H (24)</td>
<td>Thailand</td>
<td>42</td>
<td>2007</td>
<td>4–8</td>
<td>AS, DHA</td>
<td>MQ, LUM, MQ</td>
<td>MQ, DQ</td>
<td>9, 4</td>
<td>20</td>
</tr>
<tr>
<td>E</td>
<td>HRP2</td>
<td>Colombia</td>
<td>57</td>
<td>2006–2007</td>
<td>0–12</td>
<td>AS, DHA</td>
<td>MQ, LUM, MQ</td>
<td>MQ, DQ</td>
<td>12, 4</td>
<td>21</td>
</tr>
<tr>
<td>F</td>
<td>HRP2</td>
<td>Bangladesh</td>
<td>89</td>
<td>2008–2009</td>
<td>0–12</td>
<td>AS, DHA</td>
<td>MQ, LUM, MQ</td>
<td>MQ, DQ</td>
<td>12, 4</td>
<td>21</td>
</tr>
<tr>
<td>G</td>
<td>HRP2</td>
<td>Uganda</td>
<td>77</td>
<td>2010</td>
<td>1–6</td>
<td>AS, DHA</td>
<td>MQ, LUM, MQ</td>
<td>MQ, DQ</td>
<td>9, 2</td>
<td>23</td>
</tr>
<tr>
<td>H</td>
<td>HRP2</td>
<td>Vietnam</td>
<td>48</td>
<td>2010–2011</td>
<td>2–48</td>
<td>DHA</td>
<td>MQ, LUM, MQ</td>
<td>MQ, DQ</td>
<td>9 or 12</td>
<td>24</td>
</tr>
<tr>
<td>I</td>
<td>LDH</td>
<td>Senegal</td>
<td>104</td>
<td>2009</td>
<td>0–12</td>
<td>DHA</td>
<td>MQ, LUM, MQ</td>
<td>MQ, DQ</td>
<td>9</td>
<td>25</td>
</tr>
<tr>
<td>J</td>
<td>LDH</td>
<td>Travelers</td>
<td>195</td>
<td>2009</td>
<td>4–48</td>
<td>DHA</td>
<td>MQ, LUM, MQ</td>
<td>MQ, DQ</td>
<td>9 or 4</td>
<td>18</td>
</tr>
<tr>
<td>K</td>
<td>LDH</td>
<td>Thailand</td>
<td>64</td>
<td>2009</td>
<td>4–8</td>
<td>DHA</td>
<td>MQ, LUM, MQ</td>
<td>MQ, DQ</td>
<td>4</td>
<td>21</td>
</tr>
<tr>
<td>L</td>
<td>SYBR</td>
<td>Cambodia</td>
<td>56</td>
<td>2010</td>
<td>18–24</td>
<td>AS, DHA</td>
<td>MQ, LUM, MQ</td>
<td>MQ, DQ</td>
<td>8</td>
<td>27</td>
</tr>
<tr>
<td>M</td>
<td>SMT</td>
<td>Colombia</td>
<td>57</td>
<td>2006–2007</td>
<td>0–12</td>
<td>DHA</td>
<td>MQ, LUM, MQ</td>
<td>MQ, DQ</td>
<td>4</td>
<td>21</td>
</tr>
<tr>
<td>N</td>
<td>SMT</td>
<td>Ghana</td>
<td>94</td>
<td>2010</td>
<td>0–6</td>
<td>AS</td>
<td>MQ</td>
<td>MQ</td>
<td>12</td>
<td>26</td>
</tr>
</tbody>
</table>

*3H,* isotopic hypoxanthine method (with the timing of addition of hypoxanthine [in hours] given in parentheses); SMT, schizont maturation test.

**Method**

- DHA, dihydroartemisinin; AM, artemether; AS, artesunate.
- DQ, desethylamodiaquine; MQ, mefloquine; LUM, lumefantrine; PIP, piperaquine.
- CQ, chloroquine; QN, quinine.

### References

1. Dha, dihydroartemisinin; AM, artemether; AS, artesunate.
2. DQ, desethylamodiaquine; MQ, mefloquine; LUM, lumefantrine; PIP, piperaquine.
3. CQ, chloroquine; QN, quinine.
4. *References are given for the descriptions of the methodology used at each site (not necessarily the specific data assessed).*
5. *Samples that were obtained from returning travelers presenting to French hospitals and examined at the Centre National de Référence du Paludisme, Paris, France.*
were explored using a random-effects model (Stata, version 11.1; StataCorp), with the drug and the method as fixed effects and the site as a random effect (due to the heterogeneity between sites). Since it was suspected from initial observations that this phenomenon was associated with drugs that are relatively inactive against ring-stage parasites, DHA was used as the reference group for the drug; it was also the most commonly assayed antimalarial drug in current use (1,391 assays across 13 of the 14 data sets). Hypoxanthine incorporation was defined as the reference method.

**High-throughput estimation of IC₅₀.** Curve fitting was undertaken using a sigmoid $E_{\text{max}}$ model. In its general form, this model has four parameters: the IC₅₀ (the 50% effective concentration [IC₅₀] for concentration-inhibition data), a measure of the curve’s steepness at the IC₅₀ (the sigmoidicity factor, or gamma), and the levels of uninhibited growth and maximally inhibited growth (see above).

Code from ICEstimator (15), based on the nls algorithm of R, which performs successive fittings of a sigmoid $E_{\text{max}}$ model to concentration-inhibition data, was adapted within a Google Web Toolkit (GWT) Java-based Web application to perform data transformation, standardized analysis, and reporting of IC₅₀ for each data set. The details of ICEstimator have been described elsewhere (15). Briefly, the primary growth data are first converted to a percentage scale, with baseline growth (no drug) representing 100% and minimum growth (maximum drug inhibition) representing 0%. Following this conversion, the model is constrained at its upper and lower ends to 100% and 0%, respectively, and therefore produces only two parameters: the IC₅₀ and the sigmoidicity factor (gamma). Initial values for the IC₅₀ and gamma are determined by the point at which growth first falls below 50% of control growth (15, 18), and iterations are then undertaken until the limit of improvement is reached. In case of nonfitting (because of a weak dose-response relationship or a paucity of intermediate data points between 100% and 0%, as seen with a very steep slope), curve fitting is attempted again with gamma fixed at 10, based on a previous sensitivity analysis showing that gamma values greater than 10 would not significantly alter IC₅₀ in steep curves (15).

The sigmoid $E_{\text{max}}$ model is focused primarily on determining the IC₅₀ and the slope at this IC₅₀, and all other points on the modeled line are entirely determined by the IC₅₀ and gamma. Points toward the ends of the curve, such as the IC₉₀ and IC₉₅, frequently depart to some degree from the data observed, and for this reason, these values are potentially misleading and are not reported by IVART.

**Assessment of criteria for defining a reliable subset of assays.** It is generally recognized that at least 30% of parasite isolates placed in short-term culture exhibit less than optimal growth due to preexposure to drugs or other factors contributing to reduced parasite viability (4). To detect assays that are less reliable due to such factors, IVART was set up to calculate the ratio of the upper and lower 95% confidence intervals of the IC₅₀ estimate, known as the confidence interval ratio (CIR). A threshold CIR of <3 was selected to define core assays of higher reliability for entry into pooled analyses and association studies. The CIR parameter is not useful in a subset of cases where modeling can be achieved only with a fixed gamma of 10, since this becomes a 1-parameter model, with inevitable narrowing of confidence intervals. For this subset of fixed-gamma assays, the growth ratio (uninhibited growth divided by maximally inhibited growth) was used to define core assays of higher reliability in accordance with previous recommendations (4). For each data set, the main subset of assays in which both the IC₅₀ and gamma were successfully obtained was examined, and the proportion of assays with tight confidence intervals (CIR, <3) was determined at four levels of the growth ratio: <2, 2 to 3, 3 to 5, and >5.

The effect of applying reliability criteria was explored by examining intrainolate Pearson correlations of IC₅₀ for drug pairs in the whole data set and repeating this procedure with increasingly strict criteria.

**Relative efficacies of artemisinins and partner drugs.** Within the 12 nonmicroscopic data sets, the growth ratio (uninhibited growth divided by maximally inhibited growth) for each drug was compared to that for
DHA by using the Mann-Whitney test (with Bonferroni’s correction for the number of drugs). The relative proportions of growth inhibition by DHA and MQ for individual parasite isolates, as illustrated in Fig. 1, were assessed by the Wilcoxon signed-rank test.

RESULTS

Constraints for curve fitting. The 14 data sets contained 7,350 individual drug inhibition assays. Analysis of growth inhibition characteristics revealed that 1,334 (18.1% of total) assays showed evidence of paradoxical growth at high concentrations (a rise in apparent growth of >10% over that at intermediate concentrations). Both the nature of the drug being tested and the readout methodology were clearly associated with this behavior. The four ACT partners (DQ, LUM, MQ, and PIP) were associated with a risk 2- to 3-fold greater (P < 0.001 in all cases) than that for the reference drug, DHA, while CQ and QN showed paradoxical effects at an intermediate level (P < 0.001 in both cases) (see Table S2 in the supplemental material). The phenomenon was seen less commonly with AS than with DHA. The assay method was also relevant; by using hypoxanthine incorporation as the reference group, a paradoxical increase in growth was most commonly encountered in assays based on LDH quantification by ELISA (odds ratio [OR], 1.27; P, 0.007). There was also a trend toward less-frequent occurrence in HRP2 and SYBR green assays that did not reach statistical significance. The issue was not encountered at all in assays assessing schizont maturation by microscopy.

These findings indicate that a paradoxical increase in growth measured at high drug concentrations reflects biological properties of antimalarial drugs rather than being simply an experimental artifact due to equipment or human error. For this reason, it was decided to amend the lower constraint of the sigmoid E_max model to G_min in order to provide a more accurate measure of maximum inhibition, avoiding underestimation of overall drug efficacy and spuriously low IC50s.

Reliability criteria. The default criteria for defining a subset of more-reliable assays consisted of one main criterion, an IC50 confidence interval ratio (CIR) of <3. In the subset of curves with a fixed gamma (1,427 assays [19.4% of the total]), a growth ratio indicating a satisfactory signal/background ratio was used to assess reliability. As expected, there was a clear relationship between the growth ratio and the CIR in the 80.6% of assays where gamma was derived by modeling (i.e., 2-parameter models) (see Fig. S1 in the supplemental material). Growth ratios of 3 to 5 and >5 were associated with very high levels of assays with tight confidence interval ratios (across the 12 nonmicroscopic data sets, the median proportions with CIRs of <3 were 91.3% for a growth ratio of 3 to 5 and 95.3% for a growth ratio of >5). In contrast, a growth ratio of <2 was associated with relatively high proportions (median, 33.9%) of assays with uncertain IC50 estimates (CIR, >3). A growth ratio of 2 to 3 was generally associated with high levels of assays with tight confidence intervals (median proportion, 86.8%), but there did appear to be greater potential for less-reliable IC50 estimates to be accepted at this growth ratio level in hypoxanthine incorporation assays (see Fig. S1 in the supplemental material, data sets A and C).

The effect of applying reliability criteria was assessed by calculating the correlation between IC50,50 for drug pairs within each data set, since it was predicted that application of increasingly strict criteria might improve correlation coefficients. MQ-QN and CQ-DQ provided the most powerful test cases given the number of data sets assessing each drug pair (11 and 8, respectively) and the strong, consistently documented associations between IC50,50 for these drug pairs in field studies (20, 28–33). Exclusion of assays with CIRs of >3 and growth ratios of <2 (for assays with a fixed gamma) led to stronger correlations in 9/11 data sets for MQ-QN and in all 8 data sets for CQ-DQ (Fig. 2); see also Table S3 in the supplemental material). Statistical significance was generally maintained or strengthened in the more-reliable subset despite the reduction in sample size.

Increasingly strict classification of the fixed-gamma subset of assays, involving the exclusion of additional assays with growth ratios of <3 or <5 generally led to relatively small and inconsistent improvements in correlation (Fig. 2; see also Table S3 in the supplemental material). For hypoxanthine assays, where there was concern that unreliable assays might be accepted with growth ratios of 2 to 3 (see above), the mean proportion of additional isolates excluded in this range ranged from 1.8 to 5.1% in the four data sets (examining all drug pairs). The effects of various levels of exclusion criteria in hypoxanthine assays (drug pairs MQ-QN and CQ-DQ) are illustrated in Fig. S2 in the supplemental material.

IVART was therefore set to apply a default growth ratio of 2 to classify the subset of curves with a fixed gamma. When this default growth ratio was combined with the main criterion of a CIR of <3, 6,158 of 7,350 curves (83.8%) met the IVART core criteria.

Relative efficacies of artemisinins and partner drugs. Using only data that conformed to the default IVART reliability criteria (see above), it was possible to discern clear patterns in terms of the growth ratio depending on the assay and the drug (Fig. 3). For example, hypoxanthine-based assays showed the highest growth ratios (typically 5- to 15-fold reductions in growth), while protein-based and SYBR green assays had substantially lower growth ratios (i.e., relatively higher background values). Furthermore, nonartemisinin drugs tended to show lower growth ratios than artemisinin derivatives, an effect more marked with ELISA-based assays.

This issue was explored in more detail by examining the proportion of growth that could be inhibited [(G(C0) - G(min)) / G(C0)] (Fig. 1) by DHA compared to MQ for individual parasite isolates for which both assays passed the core criteria. In all 11 data sets with data for both drugs, MQ inhibited growth significantly less than DHA (P, <0.01 by the Wilcoxon signed-rank test), but the extent of this effect ranged from 1.0 to 30.2% (median values) across data sets (Fig. 4). ELISA-based assays showed substantially greater differences than those based on hypoxanthine incorporation or SYBR green fluorescence.

IVART online. The modeling approaches, constraints, and core criteria described above were adopted within an online version of IVART that is now available for external use at http://www.wwarn.org/toolkit/data-management/ivart. After a one-time registration process, users have access to the tool via a personalized interface that provides secure upload and storage of primary data in a 96-well plate format. IVART incorporates a “Plate Assistant” function to verify data layout and drug concentration information and then undertakes a single-pass analysis that produces graphical (Fig. 5) and spreadsheet reports for each individual assay, along with summaries of reliable-assay subsets by drug and year of study.

The range of drug concentrations in tests can also affect the calculated IC50. In some data sets, the lowest concentration of the drug tested demonstrated more than 50% inhibition, suggesting
that the drug concentration range tested was too high. Such assays still provide useful information on drug sensitivity (as long as they pass core criteria) but are marked in IVART with a “Range High” warning, because reassessment of the range of concentrations used in future assays may be indicated. Similarly, in some assays, growth inhibition appears to be incomplete, even at the highest drug concentration used. A “Range Low” warning is therefore displayed when the level of inhibition at the highest concentration of a drug is ≥10% greater than that at the next lowest concentration. Cautious interpretation of such assay data is required, since the effect may be explained by technical factors, such as underdosing of the drug in the wells or hemolysis of red blood cells, but such assays may also hint at emerging drug resistance.

DISCUSSION
A standardized approach to modeling in vitro data. Differences in laboratory and analytical practices complicate the comparison of data from antimalarial susceptibility assays obtained in different laboratories. This problem can potentially be reduced in a number of ways (2), including the use of validated reference clones (25, 34, 35) and quality-controlled drugs (36). IVART was developed to address a third source of variability by defining a single approach to the calculation of IC₅₀ that could be applied to primary data collected using a range of growth readout methods.

Examination of a wide range of data sets from 11 in vitro testing laboratories confirmed wide variations in experimental method and design. The number and identity of drugs being assessed differs across laboratories, presumably influenced by local patterns of clinical drug usage and susceptibility. This, in turn, affects the number and range of drug concentrations assessed, along with the number of no-drug control wells (critical for establishing a baseline for calculating IC₅₀). In addition, some investigators select specific subsets of control wells from certain rows or columns for each drug, and control growth values may also be derived from wells containing low drug concentrations if these produce higher apparent growth than drug-free wells for any reason. Other potential sources of variation include the use of different models for curve fitting and manual removal of individual points considered to be outliers.

We began the process of standardizing the calculation of IC₅₀ by selecting the sigmoid E_max nonlinear regression model for curve fitting, since this does not involve subjective decisions regarding the form of the inhibition curve (a potential requirement if a polynomial curve is used). The upper and lower bounds of the model were constrained; given the range of plate layouts employed by different investigators, the upper constraint was defined as the mean growth in all wells on the plate with no drug present. Although edge and cross-plate changes in growth have been reported (37), the use of this larger number of drug-free control wells provides a statistical advantage in terms of greater numbers.

Several approaches to defining a lower constraint for curve fitting of antimalarial susceptibility data have been described. In hypoxanthine incorporation assays, the signal in uninfected red blood cells can be used, while the background in ELISA-based assays can be obtained by measuring the baseline antigen present at the start of incubation. However, in practice, these parameters are rarely recorded, and growth at the highest drug concentration, G(C_max), is commonly used; for example, G(C_max) was the method of choice in the original description of the LDH ELISA (9). However, in the data sets examined here, the assumption that the highest concentration of a drug defines its greatest level of inhibition proved incorrect, since nearly one-fifth of assays showed a paradoxical increase in apparent growth at very high drug concentrations. This effect did not appear to be due simply to

FIG 2 Effects of differing reliability criteria on interdrug correlation. Each series of interconnected symbols represents the Pearson correlation coefficients for chloroquine-desethylamodiaquine (CQ-DQ) (8 data sets) and mefloquine-quinine (MQ-QN) (11 data sets). Three levels of exclusion criteria were studied: a, default IVART criteria; b, IVART criteria with additional exclusion of fixed-gamma assays with a growth ratio of 2 to 3; c, IVART criteria with additional exclusion of fixed-gamma assays with a growth ratio of 3 to 5. Filled triangles, significant (P < 0.05) correlation coefficients; open triangles, nonsignificant (P > 0.05) correlation coefficients. Individual correlation coefficients, P values, and numbers of samples are shown in Table S3 in the supplemental material.
noise or outlying values, since it was more marked with nonartemisinin drugs and ELISA-based readouts. One possible explanation for such paradoxical apparent growth at high drug concentrations is that drugs lacking a primary effect on the ring-stage parasite at standard pharmacological concentrations may nevertheless affect the ring-stage parasite in other ways, leading to altered transcriptional responses and increased protein production or nucleic acid uptake. Additional explanations include precipitation of the drug from the solution at high concentrations, plate edge effects (37), and mixed-clone infections (38). Whatever the mechanism of this higher apparent growth at high drug concentrations, the definition of maximum growth inhibition clearly needs to account for cases where maximum inhibition occurs at intermediate concentrations of a drug. Ranking of concentrations in terms of the degree of inhibition allowed the selection of a modified measure of maximal inhibition, based on average growth over the two concentrations with the lowest growth.

**Systematic application of reliability criteria.** Since a proportion of parasites adapt poorly to *in vitro* culture and provide misleading signals, investigators usually define a core subset of more-reliable assays for use in association studies and summary outputs.
The method for defining such assays is rarely described in publications, and when conducted at the level of the individual assay, the decision-making process is likely to be time-consuming and potentially subjective. A key aim of IVART was to promote an objective approach to be applied across whole data sets. IVART uses a confidence interval ratio (CIR) of the IC50 estimate as its main method of defining core assays: a CIR of 1/3 is considered to indicate a reliable assay. However, the CIR is not useful in a subset of assays where initial 2-parameter modeling of the concentration-inhibition data fails and a fixed gamma value of 10 is used (around 20% of all assays); in this scenario, other means of defining reliable assays are required. Measures of goodness of fit were not chosen as IVART’s default criteria because of the clear evidence that in a proportion of assays, the biological properties of drug inhibition produce data that naturally deviate from the classical sigmoid concentration-inhibition curve (see above). Such assays may be robust in terms of signal but nevertheless produce poor scores in goodness-of-fit assessments and would tend to be inappropriately rejected.

Historically, the overall level of signal to background (uninhibited to maximally inhibited growth, known as the growth ratio) has been recommended as a means of defining reliable curves (4). Examination of the relationship between the growth ratio and the confidence interval ratio across the data sets indicated that a threshold growth ratio of 2 would lead to acceptance of very few unreliable assays for ELISA- and SYBR green-based assays, and this was adopted within the default criteria of IVART. However, it was noted that there was a greater potential to accept less-reliable data in hypoxanthine-based assays, where the signal-to-background ratio is usually much higher than 2. This is also consistent with previous suggestions for reliability criteria in hypoxanthine-based studies, for which a growth ratio of 5 was proposed (4). In this study, when more-restrictive criteria were applied in hypoxanthine-based data sets, leading to the exclusion of assays with growth ratios of 1/3 or 1/5, relatively few additional isolates were excluded (since such growth ratios are rarely encountered in hypoxanthine-based data sets, and the growth ratio is applied only to the minority fixed-gamma subset). Accordingly, the correlation

<table>
<thead>
<tr>
<th>Drug</th>
<th>IC50</th>
<th>Lower CI</th>
<th>Upper CI</th>
<th>CI Ratio</th>
<th>Gamma</th>
<th>Meets Core Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>CQ</td>
<td>650.97</td>
<td>528.21</td>
<td>773.73</td>
<td>1.46</td>
<td>5.04</td>
<td>yes</td>
</tr>
<tr>
<td>DHA</td>
<td>0.58</td>
<td>0.51</td>
<td>0.64</td>
<td>1.25</td>
<td>2.73</td>
<td>yes</td>
</tr>
<tr>
<td>DQ</td>
<td>90.82</td>
<td>85.41</td>
<td>96.24</td>
<td>1.13</td>
<td>10.00</td>
<td>yes</td>
</tr>
<tr>
<td>LUM</td>
<td>0.42</td>
<td>0.35</td>
<td>0.48</td>
<td>1.40</td>
<td>2.26</td>
<td>yes</td>
</tr>
<tr>
<td>PIP</td>
<td>4.84</td>
<td>4.06</td>
<td>5.63</td>
<td>1.39</td>
<td>5.71</td>
<td>yes</td>
</tr>
<tr>
<td>QN</td>
<td>82.34</td>
<td>68.8</td>
<td>95.87</td>
<td>1.39</td>
<td>2.18</td>
<td>yes</td>
</tr>
</tbody>
</table>

FIG 5 Excerpt from an output PDF file showing results for a Ugandan sample tested with six drugs. IC50s are given in nanomolar concentrations. Growth values for the zero-drug wells are plotted at the left-hand ends of the graphs (at 2 log units below the lowest drug concentration used) for display purposes.
coefficients for MQ-QN and CQ-DQ did not, on the whole, improve with these more-stringent criteria.

IVART was not designed with assays based on microscopic assessment in mind, since the use of microscopy-based methods to assess growth is decreasing. For all the drugs described here, growth inhibition should be complete at high drug concentrations, so the issues of determining maximal inhibition and the use of the growth ratio (usually infinity in schizont maturation experiments) to define reliable assays under certain circumstances do not apply. Nevertheless, the tool may be useful to laboratories continuing to use this methodology provided these issues are appreciated.

Distinctive properties of artemisinin derivatives. The high-throughput nature of IVART provided a unique opportunity to undertake a systematic examination of growth characteristics across a range of drugs and readout methods. As well as informing the design of IVART itself, this process provided additional biological insights informative for the future design and interpretation of in vitro antimalarial susceptibility studies. There was clear evidence that artemisinin derivatives show higher efficacy (i.e., inhibition of growth) than ACT partner drugs; for example, DHA inhibited a significantly greater proportion of growth than MQ in all sets of assays. This finding is consistent with the earlier onset of action of artemisinins, at the ring stage of parasite development (39–41), but the fact that this property is substantially greater in ELISA-based readouts had not been documented previously. The most likely reason for this is that both LDH and HRP2 are produced in significant quantities by ring-stage parasites (42), while hypoxanthine and SYBR green signals accumulate only at more-mature stages of asexual parasite development. There also appeared to be an effect of site, possibly reflecting the critical role of the timing of drug exposure in relation to the parasite stage. In locations with substantial delays between the removal of the sample from the patient and the setup of in vitro culture, parasites are more likely to first encounter the drug at mature stages, when they are susceptible to a wider range of compounds.

These observations prompt a reevaluation of how resistance is measured for different classes of antimalarial drugs. In the ACT era, assessment may require different approaches for artemisinins, which act rapidly against both the ring and mature stages, and ACT partner drugs, which act only against the more-mature stages of parasite development. The timing and duration of parasite-drug contact have been identified as important determinants of antimalarial susceptibility in the laboratory (2, 43), and specific methodologies and analyses for different applications are likely to provide more relevant information than a single method alone. Both the microscopic and hypoxanthine methods were developed in the era of slower-acting antimalarials with longer half-lives (CQ and MQ) (44); in these assays, ring-stage parasites contribute little signal (indeed, ring-stage growth is not assessed at all if hypoxanthine is added only after 24 h of incubation). In contrast, forms of artemisinin resistance reported from Southeast Asia (27, 45, 46) have been proposed to be confined to ring-stage parasites (47) and would not be predicted to influence susceptibility at the trophozoite or schizont stage. Short pulses of a drug during ring-stage growth have lasting growth-inhibitory effects (39–41, 48–52), and ring-stage pulse assays (in which relatively high concentrations of artemisinins are applied for relatively short periods) have been described recently (53, 54), providing the first clear view of ring-stage artemisinin resistance in parasites from western Cambodia (53). The need to remove a drug or to quantify ring-stage growth using a specific marker may present a challenge for widespread field use of this technique.

Summary and future work. IVART provides high-throughput, rapid, single pass analysis of in vitro sensitivity data, avoiding a variety of manual and potentially subjective processing steps currently in use. The tool can be applied to data sets obtained by a variety of methods and defines a subgroup of core IC50s of greater reliability for pooled analyses and association studies. Its advantages, therefore, relate to consistency of approach and convenience.

The criteria suggested for accurate identification of a reliable subset of assays for use in association studies appear to be well suited to the signal-to-background properties of data sets from ELISA and SYBR green assays (methods increasingly used by in vitro testing laboratories); this is evidenced by substantially improved correlation scores for interdrug comparisons upon the application of these criteria. Nevertheless, ongoing monitoring of the tool’s operation will be important in order to confirm prospectively that the approaches described are appropriate for further data sets from a variety of laboratories and methods. The handling of assays with sparse data around the IC50, and consequently steep falls in growth between two drug concentrations, is a particular challenge for high-throughput approaches. Future incorporation of an additional algorithm that is better able to fit 2-parameter models to such data may provide a further advance, although this will require careful validation on a similarly representational data set. Large data sets of the type described here may also be used to develop mixed-effects modeling approaches to the analysis of concentration-inhibition data involving, a Bayesian framework for assessing whether individual P. falciparum isolates are resistant to a given drug.

ACKNOWLEDGMENTS

WWARN is supported by a Bill and Melinda Gates Foundation grant. This work was supported in part by the Intramural Research Program of NIAID, NIH.

We thank Carole Mackosso, Jeff Smith, Delia Bethell, and Ball Ekapi rat for assistance with data acquisition and management, Sean Collins and Alberto Olliaro (WWARN) for design and programming of the IVART interface, Kasia Stepniewska (WWARN) and Sue Lee for statistical advice, and Bill Watkins and Ric Price for discussions.

REFERENCES


