Invited Review

Is anthelmintic resistance a concern for the control of human soil-transmitted helminths?

Jozef Vercruysse a,*, Marco Albonico b, Jerzy M. Behnke c, Andrew C. Kotze d, Roger K. Prichard e, James S. McCarthy f, Antonio Montresor g, Bruno Levecke a

a Ghent University, Faculty of Veterinary Medicine, Department of Virology, Parasitology and Immunology, Laboratory of Parasitology, Merelbeke, Belgium
b Ivo de Carneri Foundation, Milan, Italy
c School of Biology, University of Nottingham, Nottingham, United Kingdom
d Division of Livestock Industries, Commonwealth Scientific and Industrial Research Organisation, Brisbane, Australia
e Institute of Parasitology, McGill University, Montreal, Canada
g Queensland Institute for Medical Research, University of Queensland, Herston, Australia
h Department of Neglected Tropical Diseases, World Health Organization, Geneva, Switzerland

ARTICLE INFO

Article history:
Received 30 July 2011
Received in revised form 13 September 2011
Accepted 16 September 2011
Available online 14 October 2011

Keywords:
Soil-transmitted helminths
Anthelmintics
Anthelmintic resistance
Guidelines

ABSTRACT

The major human soil-transmitted helminths (STH), Ascaris lumbricoides, hookworms (Necator americanus and Ancylostoma duodenale) and Trichuris trichiura have a marked impact on human health in many parts of the world. Current efforts to control these parasites rely predominantly on periodic mass administration of anthelmintic drugs to school age children and other at-risk groups. After many years of use of these same drugs for controlling roundworms in livestock, high levels of resistance have developed, threatening the sustainability of these livestock industries in some locations. Hence, the question arises as to whether this is likely to also occur in the human STH, thereby threatening our ability to control these parasites. This is particularly important because of the recent increase in mass control programmes, relying almost exclusively on benzimidazole anthelmintics. It will be important to ensure that resistance is detected as it emerges in order to allow the implementation of mitigation strategies, such as use of drug combinations, to ensure that the effectiveness of the few existing anthelmintic drugs is preserved. In this review we address these issues by firstly examining the efficacy of anthelmintics against the human STH, and assessing whether there are any indications to date that resistance has emerged. We then consider the factors that influence the effect of current drug-use patterns in selecting for resistant parasite populations. We describe the tools currently available for resistance monitoring (field-based coprological methods), and those under development (in vitro bioassays and molecular tests), and highlight confounding factors that need to be taken into account when interpreting such resistance-monitoring data. We then highlight means to ensure that the currently available tools are used correctly, particularly with regard to study design, and we set appropriate drug-efficacy thresholds. Finally, we make recommendations for monitoring drug efficacy in the field, as components of control programmes, in order to maximise the ability to detect drug resistance, and if it arises to change control strategy and prevent the spread of resistance.

© 2011 Australian Society for Parasitology Inc. Published by Elsevier Ltd. Open access under CC BY-NC-ND license.

1. Introduction/background

The three major human major soil-transmitted helminths (STH), Ascaris lumbricoides (roundworm), Necator americanus/Ancylostoma duodenale (the hookworms) and Trichuris trichiura (whipworm) are amongst the most prevalent parasites worldwide. It is estimated that there are more than one billion cases worldwide, of which 450 million have significant morbidity attributable to their infection, the majority of whom are children. An additional 44 million infected pregnant women suffer significant morbidity with hookworm-associated anaemia likely contributing to maternal mortality. It has been estimated that approximately 135,000 deaths occur each year due to STH infections (Awasthi and Bundy, 2007), mainly due to infections with the hookworms through anaemia, the roundworm A. lumbricoides through intestinal or biliary obstruction (Wani et al., 2010) and the whipworm T. trichiura through anaemia and chronic dysentery (Stephenson et al., 2000). When measured in disability-adjusted life years (DALYs) lost, that is the number of healthy years lost to premature death or disability, STH infections are as important as malaria or tuberculosis (Brooker, 2010).
The principal intervention available for controlling STH infections is the periodic administration of one of the four anthelmintics recommended by the World Health Organization (WHO): mebendazole (MEB), albendazole (ALB), levamisole (LEV) or pyrantel (PYR). Of these, benzimidazoles (BZ) – MEB and ALB – are the most frequently used anthelmintics for treatment of STH infections. The costs estimated for drug administration are US$ 0.012–0.91 per subject (Brooker et al., 2008; Montresor et al., 2010). While ivermectin (IVM) is not recommended for the treatment of human STH, except for strongyloidiasis, it has good activity against ascariasis, but much less against hookworms and trichuriasis than the four drugs listed above (Richards et al., 1995). However, in the new setting where so-called “preventive chemotherapy” is advocated (World Health Organization, 2006), IVM merits consideration, as it is the drug of choice in mass drug administration programmes for controlling onchocerciasis and as a component of combination treatment for lymphatic filariasis. In addition, co-administration of IVM significantly improves the activity of ALB against T. trichiura (Ismail and Jayakody, 1999; Belizario et al., 2003). Therefore, in these settings IVM will have collateral beneficial effects for some but not all STH co-infections.

Given the paucity of suitable alternative anthelmintics it is imperative that monitoring programmes are introduced, both to assess progress, and to detect any changes in therapeutic efficacy that may arise from the selection of worms carrying mutations conferring drug resistance (Albonico et al., 2004a). The scale up of chemotherapy programmes currently underway in various parts of Africa, Asia and South America, particularly targeting school children, is likely to exert increasing drug pressure on parasite populations. This has the potential to select for parasite genotypes that can resist anthelmintics. More than 314 million children were treated for soil-transmitted helminthiasis in 2009 (205 million were treated in 2008). Although the number of children treated periodically increased by one third in 2009, 31% coverage is still far below the global target of 75% set by the WHO for 2010. Global coverage is now similar for preschool-aged children (approximately 34%) and for school-aged children (approximately 30%) (World Health Organization, 2011).

In this review we discuss the problems associated with monitoring the efficacy of anthelmintics, in order to answer the question as to whether anthelmintic resistance (AR) is a concern for the control of human STH, requiring urgent remedial action. Firstly, we review the studies describing the efficacy of anthelmintics and the problems relating to the indicators of efficacy. Secondly, we discuss the different potential confounding factors that may affect the assessment of efficacy of anthelmintics and that should be taken account of, or even excluded, before assuming AR in a particular setting. Thirdly, we consider whether the limited number of studies on reduced efficacy in hookworms represent genuine cases of AR, based on the selection of resistance alleles in the parasites, or merely reduced efficacy due to the impact of confounding factors. Fourthly, factors contributing to AR are reviewed, and fifthly recent advances in understanding the molecular basis of AR in worms of livestock are discussed, since these are likely to be highly relevant for human helminths. Sixthly, the in vivo and in vitro diagnostic methods for monitoring the emergence of AR are presented. Finally, the basis for the currently recommended guidelines for monitoring drug efficacy (World Health Organization, 1999) is discussed and recommendations for future monitoring are presented.

2. Expected efficacies of anthelmintics

The two indicators that are used to determine the efficacy of an anthelmintic in human medicine, are the cure rate (CR) and the egg reduction rate (ERR) (Box 1). Although there are significant numbers of studies investigating the CRs and ERRs of anthelmintics against human STH, many of these cannot be directly compared because, as pointed out in the systematic review and meta-analysis of Keiser and Utzinger (2008), many of these studies are confounded by methodological considerations e.g. different diagnostic methods, treatment regimens, drug dose regimens, geographical location. This review further highlighted the paucity of high quality studies that are crucial for guiding clinical decisions about which anthelmintic drug to use in a particular epidemiological setting. It is important at the outset to recognize that none of the single dose anthelmintic regimens are recognized as curative therapy for an individual patient, with current clinical recommendations advocating multiple doses over one to three days.

Bennett and Guyatt (2000) reported drug efficacy data as CR and ERR for 400 mg single-dose ALB, 500 mg single-dose MEB and multiple-dose MEB (100 mg twice daily for three days) in treating A. lumbricoides, hookworm and T. trichiura infection. They concluded that marked variation is mainly observed when the CR is the analysed endpoint. Their major concern in the control of STH infections was the moderate and low efficacy of single dose MEB for hookworm and T. trichiura infections, suggesting that such dose regimens may be sub-optimal. Assuming that the drugs used are of high quality, any change in the efficacy of BZ for A. lumbricoides infection would be clearly apparent because expected efficacy levels are so high. In contrast monitoring changes in drug efficacy against hookworms and T. trichiura is more unreliable because efficacy levels are relatively low and variable.

In their systematic review and meta-analysis of the efficacy of single-dose oral ALB, MEB, LEV, and PYR pamoate against STH infections Keiser and Utzinger (2008) could only use CR as the primary outcome measure, as there were insufficient valid studies, in the public domain, reporting the ERR. Single-dose oral ALB, MEB, and PYR pamoate in A. lumbricoides infection resulted in CR of 88%, 95% and 88%, respectively. The efficacy of single-dose oral ALB, MEB, and PYR pamoate against hookworm infections was 72%, 15% and 31%, respectively. The CRs for infection with T. trichiura following treatment with single-dose oral ALB and MEB were 28% and 36%, respectively. No pooled relative risks could be calculated for PYR pamoate against T. trichiura and LEV for any of the parasites investigated.

Recently, Vercruysse et al. (2011) evaluated the efficacy of single dose 400 mg ALB against STH in children in 7 trials located in countries across Africa, Asia and South America, using a standardized protocol. Overall, the highest CRs were observed for A. lumbricoides (98%) followed by hookworms (88%) and T. trichiura (47%). The variability in the CR for the three parasites could be largely explained by country (trial), age or the pre-intervention (pre-treatment) faecal egg counts (FEC). The latter is probably the most important factor as indeed had a considerable effect on the CR of all three STH infections. This finding highlights the proposition that the CR should not be used as an indicator for drug efficacy, as it is sensitive to variation in the intensity of infection before treatment (see also Montresor, 2011). Hence, comparisons between populations (countries, villages, schools, etc.) differing in pre-intervention FEC are bound to provide different conclusions about drug efficacy. Therapeutic efficacy, as reported by the ERR, was very high for A. lumbricoides (99%) and hookworms (95%), but significantly lower for T. trichiura (51%). In addition, there was considerable variation in ERR for T. trichiura across the different study sites, which could be largely explained by the mean baseline FEC (Levecke et al., in press).

3. Confounding factors affecting anthelmintic efficacy

It is very important to make a distinction between “reduced efficacy” and “anthelmintic resistance”, though in practice it is
not at all easy to do so. Many potential confounding factors may affect the efficacy of an anthelmintic, and should first be excluded before AR can be assumed. Such issues have been extensively investigated in veterinary intestinal nematode infections, but less so among species infecting humans.

An important host-related factor is the significant variation in the pharmacokinetics of anthelmintics. A greater understanding of the pharmacokinetics of anthelmintic drugs such as the BZs in livestock in the last few years has contributed significantly to improving parasite control in livestock (Ali and Hennessy, 1995). In contrast there is a paucity of pharmacokinetic and -dynamic data for anthelmintics in humans (Geary et al., 2010). Since the broad-spectrum anthelmintic activity of BZ compounds relies on the extended presence of effective drug concentrations at the location of the parasite in its precise niche in the host (Lacey, 1990), it implies that in humans increased drug concentration at this site and extending the exposure period of the parasite to the drug should result in enhanced clinical efficacy (Prichard et al., 1978; Ali and Hennessy, 1995). If so, manipulation of the formulation and dose regimen may result in an improved pharmacokinetic profile, thereby improving drug efficacy. For example, reduction of feed intake resulted in increased plasma availability of ALB in animals (Hennessy et al., 1995; Lifschitz et al., 1997). However, in humans the opposite effect was observed, with a five-fold increase in absorption of ALB if administered with a fatty meal (Lange et al., 1988). In this respect it is important to recognize that there is an incomplete understanding of the clinical pharmacology of these drugs in humans. It is often assumed that it is the intraluminal drug concentration that is primarily responsible for exposing the parasite to the drug. However, many of these drugs are absorbed and metabolized in a variable fashion, and drug metabolites may or may not be active against the parasite. Furthermore, the parasite may be exposed to the drug while feeding on blood or tissue fluid or by drug recirculating through the enterhepatic circulation. Without a comprehensive understanding of the interactions between host, drug and parasites, optimization of drug formulation and dose regimens may be difficult. Other confounding host factors include variations in intestinal transit time, (which is accelerated by diarrhoea), and diet which alters the rate of intestinal transit, thereby reducing the duration of exposure of parasites to anthelmintic, and thus reducing efficacy (Sanchez et al., 2006).

Pharmacogenetic variation in drug handling, age-related changes in drug distribution, drug interactions due to concomitant therapy (anti-inflammatory drugs or antibiotics) and co-morbidities (e.g. gastrointestinal diseases, malnutrition and immunodeficiency) may also affect anthelmintic efficacy. In addition, many drugs require a competent immune system to achieve optimal efficacy. Some food types and drugs such as grapefruit and the antacid cimetidine have an effect on cytochrome P450-mediated drug handling at the intestinal luminal interface thereby modifying ALB pharmacokinetics (Nagy et al., 2002).

Parasite-related confounding factors include infection intensity and worm fecundity. High infection intensity (parasite burden) may interfere with the pharmacokinetics (=reduced bioavailability) of an anthelmintic. Individual studies that have segregated the drug efficacy data, by class of intensity, have demonstrated a reduction in CR and ERR at high intensities of T. trichiura and hookworm infections (Bennett and Guyatt, 2000). Combining the ERR results of five recently conducted efficacy trials revealed a similar trend, yet only for T. trichiura (Vercruysse et al., 2011; Levecke et al., in press). Across these trials, the ERR for T. trichiura dropped linearly as a function of the mean baseline FEC. From this analysis, it was deduced that a single dose of ALB, provided satisfactory efficacy (ERR >90%) results for T. trichiura only when the mean baseline FEC were <275 EPG. Secondly, this study highlighted the concern that the current thresholds used for assessing anthelmintic efficacy in T. trichiura (ERR <50%) may lead to incorrect conclusions about anthelmintic resistance, particularly when the mean baseline FECs are high (>800 EPG). It is important to state that, in contrast to the meta-analysis of Bennett and Guyatt (2000), these trials were standardized in terms of drug (single dose ALB 400 mg from GSK), follow-up (14 days) and diagnosis (McMaster egg counting method).

Anthelmintics may also have a paradoxical effect on worm fecundity by changing infection intensity thereby confounding estimates of ERR. In a study by Kopp et al. (2008) a significant increase in egg production by hookworms was observed following drug treatment in dogs. The authors suggested that this was most likely a consequence of the relaxation of density dependent constraints on the egg output of the adult worms that survived the drug treatment, as female fecundity is inversely related to adult worm density. It is most likely that this phenomenon also applies widely to the human STH (Kotze and Kopp, 2008), and if so, drug efficacy changes based on ERR may be underestimated against a background of dynamic female worm reproductive biology.

Finally, additional confounding factors include those related to anthelmintic formulations and especially the quality of the drug. Differences between batches in the quantity of the active ingredient in a drug formulation, its bioavailability, and/or degradation during storage/transport may result in variable efficacy. For luminal-acting anthelmintics, particle size is very important (fine particular size is far superior to coarse); this is likely to be affected by passage along the intestine (Kelly et al., 1975; Wesche and Barnish, 1994). In addition to the quality and concentration of active ingredients, anthelmintics must meet the International Pharmacopoeia standard of dissolution and disintegration times that may affect drug efficacy (Albonico et al., 2007). Although little is known about the quality of anthelmintics sold for human use, several publications have reported variability in the quality of generic anthelmintics used in veterinary medicine. The concentrations of nine anthelmintic products (LEV and MEB) purchased in pharmacies and from agricultural merchants in Kenya varied from 0% to 118% of their claimed composition (Monteiro et al., 1998). Efficacy studies of seven brands of ALB against gastrointestinal nematodes in sheep in Ethiopia showed that the efficacy of only five of the seven brands was satisfactory (Kumsa et al., 2010). Lifschitz et al. (2004) compared the plasma concentration profiles of four randomly chosen generic IVM formulations and observed major differences in drug kinetic behaviour. Therefore it appears that quality control in generic veterinary formulations of anthelmintics is often inadequate (van Wyk et al., 1997). As underdosing selects for the development of AR in target parasites this may represent a major concern in human medicine as well.

4. Evidence of reduced efficacy

The possible development of AR to currently available anthelmintics is a subject of considerable interest. In livestock, resistance to all three of the major anthelmintic classes used (BZ, imidathiazoles-tetrahydropyrimidines and macrocyclic lactones (ML)) is widespread, and has been extensively studied (Wolstenholme et al., 2004; Sutherland and Leathwick, 2011). While it is possible that a similar situation to that observed among veterinary STH might develop in human STH in the future and putative resistance alleles have been found in human STH, currently there are no conclusive data, possibly due to the lack of tools and investigation, demonstrating that putative resistance alleles have increased in frequency in human STH following anthelmintic treatment, and none that these alleles have spread in the parasite populations.

Geer's and Gryseels (2001) reviewed the reports on drug resistance in human helminths and concluded that two studies (De Clercq et al., 1997; Reynolds et al., 1997) provided data suggestive
for the development of AR in hookworms, but actually fell short of providing conclusive evidence. De Clercq et al. (1997) observed no reduction in the number of Nectator eggs and a CR of only 22.9% after treatment with a single dose of MEB 500 mg. A standard egg hatch assay indicated that the Mali strain of N. americanus was almost twice as resistant to BZ as the laboratory reference strain (ED$_{50}$ of 0.117 compared to 0.069). However, the significance of this difference in ED$_{50}$ is difficult to assess given the very different genetic backgrounds of the isolates. A similar study in the same region (Sacko et al., 1999) with a single dose of ALB 400 mg resulted in a CR of 51.4% and a ERR of 77.6%, both significantly higher than reported earlier by De Clercq et al. (1997). Thus the studies fell short of providing conclusive evidence of AR against BZ in Mali.

The most likely reason for the failure of PYR (10 mg/kg) in the treatment of Ancylostoma (ERR of 46% and CR of 13%) according to Reynoldson et al. (1997) was the development of resistance as a consequence of the drug’s frequent use in the community over many years. However, the low number of patients ($n = 15$) and the high pre-treatment egg counts (mean baseline FEC = 869 EPG) may have affected the estimate of efficacy.

In Vietnam, a single dose MEB was found to have disappointing efficacy against hookworm infections (reduction of mean EPG relative to placebo was only 31%) (Floh et al., 2007). However, AR seemed unlikely as repeated dosing with MEB during 3 consecutive days in a second study resulted in a reduction of the mean EPG relative to placebo by 63%. It should also be noted that the second study was conducted among adults (in contrast to children in the first study) with higher pre-treatment burdens compared to the first study (2210 EPG vs. 263 EPG), potentially confounding comparison and/or leading to bias.

On Pemba Island, Zanzibar, the efficacy of MEB against hookworms in school children appeared to have fallen over a period of 5 years, during which time the children were regularly treated with MEB (ERR fell from 82.4% to 52.1%). This suggested the possibility of emergence of MEB-resistant hookworms on Pemba Island (Albonico et al., 2003). In support, an in vitro study using an egg hatch assay showed a lower thiabendazole ED$_{50}$ compared to the veterinary resistance-threshold for Haemonchus contortus. However, this comparison between livestock and human worm species is far from ideal, and more studies are required before a specific efficacy/resistance threshold for human hookworms can be unequivocally agreed (Albonico et al., 2005) (see Section 7.2). Molecular studies were conducted on the hookworm population of Pemba Island but no evidence was found for the $\beta$-tubulin mutation at amino acid residue 200 (Phe/Tyr) (Albonico et al., 2004b; Schwenkenbecher et al., 2007).

No studies have been published on AR in A. lumbricoides. The presence of the Phe200Tyr SNP (single nucleotide polymorphism) in $\beta$-tubulin was detected in low frequency in T. trichiura from non-treated people from Kenya and at high frequency in T. trichiura from treated people from Panama (Diawara et al., 2009). However, these SNP frequencies could not be linked to AR as sample sizes were small, anthelmintic efficacy was not assessed, and drug treated and non-treated samples were from different locations.

To date all reports indicating disappointing efficacy among the BZs have concerned MEB, and even in studies where MEB had low efficacy (Sacko et al., 1999), ALB was used as the positive control and showed extremely high efficacy against hookworms (See also Vercruysse et al., 2011). It is therefore of some concern that a recent report from Ghana has indicated a high failure rate for ALB in the treatment of hookworm infections (Humphries et al., 2011). Although the ERR was 82% following treatment of 102 infected subjects, the CR was only 61%. Even more worrying was the observation that within the group that did not respond to ALB treatment, FECs did not drop after treatment. While this study was not primarily designed as an anthelmintic trial, the work was thoroughly conducted and analyzed, and the ALB tablets were confirmed as actually having been swallowed by all study subjects.

To conclude, the number of studies assessing AR in human STH is currently very limited, and the studies to date suffer from a number of serious flaws: almost none deal with potential confounding factors (e.g. quality of the drugs), many are based on only low numbers of subjects and are significantly underpowered, the absence of standardised protocols and diagnostic techniques, and a lack of any agreed thresholds for defining AR. Therefore, it remains uncertain still whether the reports on reduced efficacy in hookworms represent genuine cases of AR, based on the selection of resistance alleles in the parasites, or rather, whether they simply reflect reduced efficacy.

5. Factors contributing to the development of anthelmintic resistance

Four factors have been identified as contributing to the development of AR: (1) initial resistance allele frequency, (2) treatment frequency, (3) refugia and (4) possibly underdosing.

5.1. Initial resistance allele frequency

At this stage, there is almost no information on the frequency of putative resistance alleles, based on veterinary parasites, in human STH, although the tools are now available for assessing this for BZ in treatment naive and treated populations infected with STH.

5.2. Treatment frequency

This is an important determinant of the speed of selection of AR: the greater the drug pressure (related to treatment frequency, relative fitness of resistant compared with susceptible worms, dose regimens, survival of free-living stages, refugia and other factors), the faster the selection of resistant nematode strains. Treatment frequencies of 5 or more times a year (up to 10 treatments/year) are not uncommon in livestock (Dorny et al., 1994). In humans, the frequency of treatments is usually limited to 1–3 per year for A. lumbricoides/hookworms/T. trichiura (Warren et al., 1993; Renganathan et al., 1995). However, even at these lower treatment frequencies, selection of AR has been repeatedly reported in sheep and goat nematodes (Coles, 1995; Boudsocq et al., 1999). This is especially the case when the same drug has been used over prolonged periods, as is the case with BZ in the control of human STH, and this combined with lower treatment frequencies might be enough to select for resistance. This has been clearly shown in nematodes of livestock, where farmers tend to use a single drug until it fails (Reinemeyer et al., 1992).

5.3. Refugia

The phenomenon of refugia, i.e. the proportion of the parasite population that is not exposed to drugs and thus escapes selection for resistance, is a very important factor in the selection pressure for development of resistance, and one whose impact on the development of AR is too often overlooked (van Wyk, 2001). The size of refugia will be mainly determined by (1) the fraction of the population treated (i.e. mass treatments versus selective or targeted selective treatments) and (2) the proportion of the worm population present in the environment where it is not subject to drug action (e.g. in the soil). This is influenced in turn by a range of factors including climate, resilience of the transmission stages in the face of environmental stressors and longevity of the free-living stages. For most of the STH of human importance there is still a poor understanding of the exact role of these important parameters,
and this in turn hinders the construction of reliable predictive mathematical models. Mathematical modelling in veterinary parasitology has shown that it is possible to delay the development of AR by not treating part of the herd or flock. Whether this is ethically acceptable in human chemotherapy programmes is a highly debatable/controversial topic, and in the current climate is unlikely to be given approval by ethical review committees.

Dobson et al. (2011) showed that leaving some sheep untreated worked best in situations where animals were already grazing or were moved onto pastures with low populations of infective larvae. In those cases, AR was delayed and nematode control was maintained when 1–4% of adult stock remained untreated.

In control programmes of human STH, treatment is often directed at specific target groups typically school-aged children (they often remain in school for 5 years and therefore most children can be guaranteed to have received a set number of treatments) and other population groups at greatest risk of morbidity. This effectively provides refugia, especially for hookworm as adults tend to harbour greater worm burdens than children. Thus targeting school children likely reduces the selection pressure on hookworms. This effect is further magnified because coverage and compliance is often less than 80% (Chitsulo et al., 2000). On the other hand, mass treatment with ALB is now a standard component of the Global Programme to Eliminate Lymphatic Filariasis (GPELF) and ALB is being distributed to all communities as part of the Schistosomiasis Control Initiative (SCI) activities.

The size of refugia is also largely determined by factors such as the timing of the treatment and the climatic conditions immediately prior to treatment as both will influence selection pressure on the parasites. Treatment of humans under very dry climatic conditions may strongly select for an increase in the proportion of resistant worms because the climatic conditions will kill previously deposited eggs and larvae on the soil surface. However, in wetter environments, pre-parasitic stages of susceptible worms might survive in the environment and dilute the resistant genes in subsequent generations. However, the relevance of the timing of treatment on the refugia for the human STH is yet to be thoroughly studied. Treating at times of the year when a high proportion of parasites are in refugia in the environment, should delay the spread of resistance. However, the purposeful timing of MDA programmes may present logistic problems and be less effective in controlling the overall community worm burden. Villagers are usually more easily accessible in the dry season, so that entire communities can be treated simultaneously, whereas in the wet season people may leave villages early to work on their agricultural plots and thus be more difficult to reach for treatment, so a compromise has to be achieved.

5.4. Underdosing

Underdosing may constitute another risk factor for the development of AR. As was shown in the models developed by Smith et al. (1999), the impact depends on the initial (before exposure to a given anthelmintic) and the resultant (after treatment) frequency of resistance alleles in the helminth population. Depending on their ability to kill all or part of the susceptible homozygote, heterozygote and/or homozygote resistant helminths, and on the initial frequency of resistant alleles, specific dose regimens may select for AR in different ways. Assuming that resistance is determined by a single major gene comprising two alleles at a single autosomal locus and low initial frequency of the allele for resistance, the most dangerous dose is the one that kills all susceptible homozygotes but none of the heterozygous or homozygous resistant genotypes. In contrast, when the initial frequency of the allele for resistance is high, the dose that most strongly promotes resistance is the one that kills all susceptible homozygotes and all heterozygotes, but none of the resistant homozygotes (Smith et al., 1999).

Sub-optimal regimens are the rule in human treatment: anthelmintics are administered in single doses that never achieve 100% efficacy, a practice that is widely advocated and implemented in public health helminth control programmes. While the operational objectives are laudable: to ensure compliance of the target population where the objective is only morbidity control (Warren et al., 1993), this may have unwanted consequences when applied widely and over a significant period. Taking into account the limited efficacy of single dose anthelmintic treatments, the current approach might select for resistance under certain conditions. The efficacy of most anthelmintics at the dose rates used, is only moderate for the human STH (compared to veterinary STH with ERR of >98%). It is apparent that the currently recommended dose regimens are underdosing parasites in humans. However, data beyond the scope of this paper indicate that increasing the size of the single dose of BZ will not always result in increased efficacy because the nature of their antiparasitic action depends on prolongation of contact time (Vercruysse, 2005). While repeated dosing is more effective, this approach is difficult to implement in public health programmes, because coverage falls with time when it is necessary to ask subjects to return for repeated treatment or to take daily unsupervised therapy.

Thus currently recommended regimens could constitute a significant contributing factor to the development of AR in STH. On the other hand, suboptimal dose regimens may exert less selection pressure for AR development than would dose regimens that initially produced virtually 100% efficacy. These alternative possibilities need to be investigated experimentally as well as through mathematical modelling. The subject of anthelmintic dosing and the pharmacodynamics and pharmacokinetics of anthelmintics in humans have been addressed in a recent review (Geary et al., 2010).

6. Advances on the molecular basis of AR

Knowledge of the genetics of AR in helminths is still incomplete. Nevertheless a number of important principles derived from the study of livestock helminths are likely to be relevant for human helminths. The rate at which AR develops and spreads in helminth populations will depend on a range of factors, including population structure, gene migration and the prevalence of resistance alleles in the initial untreated population. To date very little information is known about population structure in human STH, or about gene migration, although the latter could be expected to be high in human STH given the levels of human migration. Studies in veterinary nematodes (e.g. Skuce et al., 2010) suggest that resistance mutations may arise independently in different populations and then be transported by host migration. The prevalence of resistance alleles is usually considered to be low. However, in untreated H. contortus populations, the initial prevalence of RFLP polymorphisms associated with resistance to BZ at the isotype 1 and 2 β-tubulin loci were 46% and 12%, respectively; surprisingly high figures (Beech et al., 1994). Similarly high (10–20%) prevalence of genetic polymorphisms associated with IVM resistance were reported in unexposed H. contortus (Anderson et al., 1998). However, these were not actual alleles, but restriction length polymorphisms (RFLP), which may have contained more than a single allele, so that the prevalence of a resistance causing SNP may have been lower than this figure.

The best understood type of AR is BZ resistance. In veterinary nematodes changes at Phe200Tyr or Phe167Tyr (both TTC→TAC), or rarely Glu198Ala (GAA→GCA) in β-tubulin confer BZ-resistance (Kwa et al., 1994; Ghisi et al., 2007; Mottier and Prichard, 2008).
In this respect, Elard et al. (1999) determined the actual prevalence of the 200Tyr codon in β-tubulin, in five different populations (each with between 38 and 101 individual parasites) of BZ susceptible *Teladorsagia circumcincta*, and found for the total population of 281 diploid organisms genotyped, that the overall prevalence of alleles containing the 200Tyr codon was 2.7%. However, the range of prevalence of the resistance-associated codon 200 was between 0% and 12.0% in the five different susceptible populations. The rapid development of resistance to IVM and BZ in helminths of livestock may be explained, in part, by an initial high prevalence of resistance alleles in some populations of nematodes. Data on the prevalence of putative AR alleles in STH of humans are only now being collected (see e.g. Schwenkenbecher et al., 2007; Diawara et al., 2009) and need to be investigated as a priority.

The number of genes involved in resistance and their mode of inheritance (dominant or recessive) are additional factors with an important influence on the rate at which AR spreads. Polymorphism in β-tubulin isotype 1 seems to be most important for BZ resistance in *H. contortus* (Mottier and Prichard, 2008). Recently, our understanding of LEV resistance has greatly advanced with Boulin et al. (2011) reporting that a number of changes in several genes, which together encode the acetylcholine receptor, result in loss of the LEV-sensitive acetylcholine receptor and thereby may cause resistance in veterinary nematodes. These data together with the findings of Kopp et al. (2005) suggest that resistance to the nicotinic drugs is based on changes in the expression of the subunits that together form the receptor. Although monepantel, a recently developed veterinary anthelmintic for sheep, has not been used to date to treat human parasites, resistance can be relatively easily selected in *H. contortus* in the laboratory. Under such in vitro selection pressure a number of possible genetic changes in the *Hcnptl*-1 gene result in the loss of a functional monepantel-sensitive acetylcholine receptor (Rufener et al., 2009). IVM resistance is thought to be polygenic, and while it is thought that β-glycoproteins are involved in IVM resistance, a detailed understanding of the genetic basis of resistance is still lacking (Prichard, 2007).

Although there is ongoing debate about the number of genes involved in resistance to these different anthelmintics, of major importance to human medicine is the observation that, at least for BZ and ML resistance, reversion to susceptibility is rare once AR has developed in helminths of livestock, even when other drugs with completely different mechanisms of action are used for protracted periods. This is an observation of major importance and is supported by a significant body of field and experimental data (reviewed by Conder and Campbell, 1995).

Although the findings from veterinary helminths should not be extrapolated directly to humans, the likelihood that similar phenomena apply in STH of medical importance should not be overlooked.

7. Monitoring the emergence of AR

7.1. Coprological methods

To date, the clearance of eggs from stools (CR) and the egg reduction rate (ERR), remain the most widely applied metrics for monitoring drug efficacy. Of these two indicators of drug efficacy, the CR has been the most frequently reported (Keiser and Utzinger, 2008), yet it is less sensitive to changes in drug efficacy (Box 1). This is because drugs with an equal CR may nevertheless show significant differences in ERR. Although expert consensus holds that the ERR is the best parasitological measure for future surveillance of changes in drug efficacy, the level of understanding of the effects of the factors inherent both in study design (diagnostic method, sample size and statistical analysis) and host–parasite interactions (level and variation in egg excretion between and within subjects) is still too poor to support agreement for an acceptable standardized protocol.

7.1.1. Diagnostic methods

Currently, three diagnostic methods are used to assess drug efficacy in human trials. These include the WHO recommended and widely used Kato-Katz thick smear (World Health Organization, 1991, 1999), the more recently developed FLOTAC method (Cringoli et al., 2010; Knopp et al., 2011) and the much older McMaster egg counting method (Ministry of Agriculture Fisheries and Food, 1986; Vercruysse et al., 2011). These three diagnostic methods differ considerably in terms of their diagnostic performance, their sensitivity, accuracy (degree of deviation from ‘true’ FEC or bias) and precision (degree of repeatability of FEC results performed on the same stool sample) of assessing FEC; and in terms of their ease of use under field conditions (need for financial, human and technical resources). Although justifiably considerable emphasis has been placed in the past on the sensitivity of diagnostic methods and, this is highly relevant for a qualitative indicator, such as CR, the impact of sensitivity on ERR is likely to be less important. This became apparent in a recent study; where differences in sensitivity both between diagnostic methods (Kato-Katz thick smear and McMaster egg counting technique) and between laboratories across Africa, Asia and Latin America resulted in less accurate CR results, but not for ERR (Levecke et al., 2011a). Therefore, the ERR based on quantitative FEC, because of its greater accuracy and precision, is a more reliable method for both diagnosis and assessment of drug efficacy. Comparative studies of quantitative diagnostic techniques have been conducted in veterinary sciences, and have indicated recently that the FLOTAC method provides an even more accurate and precise FEC and while the McMaster method is accurate, it nevertheless lacks precision (Cringoli et al., 2010, 2011; Rinaldi et al., 2011). These differences in precision between both methods can be largely explained by the mass of stool examined, up to 1 g for FLOTAC (1 observed egg being equivalent to 1 egg per gram stool (EPG)) and only 20 mg for McMaster method (1 egg observed being equivalent to 50 EPG). Similar in-depth studies of the assays deployed in public health settings are not available, yet a few studies have demonstrated that the Kato-Katz method results in significant higher FEC than the FLOTAC method (Knopp et al., 2009a,b) and the McMaster egg counting method (Levecke et al., 2011a), suggesting it is the least accurate method. Based on the mass of stool examined in a Kato-Katz thick smear (41.7 mg, and 1 egg observed is equivalent to 24 EPG), the precision of this method should be lower than the FLOTAC method. Whether in practice it is more precise than the McMaster egg counting method, however, still remains unclear. This is because, despite the use of the standard template, there is considerable variation in the actual mass of stool examined in a Kato-Katz thick smear (Engels et al., 1997; Levecke et al., 2011a), and not unexpectedly this contributes greatly to increasing the variability in FEC. Of the three diagnostic methods currently used, it is the authors’ view that the McMaster method is the most feasible and most adaptable for fieldwork. Its advantage include its simplicity and the fact that it does not require any apparatus, other than a microscope and calibrated McMaster slides (vs. FLOTAC which also requires a centrifuge and a specialized cell) and allows simultaneous enumeration of the eggs of all STHs (vs. Kato-Katz thick smear which tends to be unreliable with hookworm eggs). Based on previous studies assessing the cost of these diagnostic methods, it is estimated that the average time for preparing, reading and examining one stool sample is ~5 min for the McMaster method, 10 min for the Kato-Katz smear and 15 min for FLOTAC (Levecke et al., 2009; Speich et al., 2010; Cringoli et al., 2010).
7.1.2. Sample size and variation in egg excretion

The required sample size will not only depend on the accuracy and precision of the diagnostic method (e.g., FLOTAC will require smaller sample sizes compared to McMaster method), but this will also be affected by the level and variation in egg excretion between subjects. In veterinary sciences, there is already empirical evidence that low FECs may confound interpretation of ERR results, particularly when sample sizes are small or diagnostic methods are not precise (El-Abdellati et al., 2010; Levecke et al., 2011b), and this is supported by evidence from statistical modelling (Vidyashankar et al., 2007). As a consequence, performing an ERR where all three STHs are being evaluated concurrently in the same population, requires a different study design, especially because of the differences in fecundity between the helminth species: A. lumbricoides >> hookworm > T. trichiura (Bethony et al., 2006). In addition, there is considerable variation in egg excretion between subjects (Maizens et al., 1993), with the majority of subjects either excreting no eggs or low numbers of eggs. As a consequence of this, small sample sizes or less precise diagnostic methods are likely to impede accurate estimation of ERR.

In a first attempt to gain more insights into the ERR, Levecke et al. (unpublished results) performed a statistical simulation in which the ERR was carried out under varying conditions of sample size, precision of diagnostic methods, level and variation of egg excretion. This analysis suggested that a sample size of between 50 and 200 subjects is required for a reliable estimation of ERR, irrespective of the infection status (negative/positive). Moreover, the results indicated that for these sample sizes, greater precision of the diagnostic method did not significantly improve the detection of reduced efficacy, highlighting that in practice the feasibility of using a particular method in a given situation in the field is perhaps a more important determinant of the choice of diagnostic method. For smaller sample sizes, variation in egg excretion impeded a successful interpretation of the ERR, even when precise diagnostic methods were applied.

7.1.3. Statistical analysis

Wherever possible, statistical analysis should include assessment of the effect of treatment with age, sex and any other relevant factors, simultaneously in multi-factorial tests. With the development of fast, powerful and relatively cheap laptop computers, large and robust statistical programmes, that control for these confounding variables, and most importantly pre-intervention intensity can now be used relatively easily in the field. However, it should be noted that statistical tests are best applied to data accrued at the individual subject level (individual-based formulae, see below), where the value of FEC or change in FEC is the dependent variable. As discussed below, an important recommendation of this review is that ERR should be calculated at the group level and the outcomes of these can only be compared to established and agreed threshold of efficacy (Box 1). Tests based on randomised sub-sampling of the data acquired from large trials can be used to establish 95% confidence limits (bootstrap analysis) and the degree to which these overlap with agreed threshold for efficacy.

An often neglected factor that is now known to have a profound effect on efficacy of treatment is the intensity of infection prior to treatment. Montresor et al. (1998) recommended the following ranges of egg counts corresponding to low, moderate and high intensities of infection for A. lumbricoides (1–4999 EPG, 5000–49,999 EPG and ≥50,000 EPG), for hookworms (1–1999 EPG, 2000–3999 EPG and ≥4000 EPG) and for T. trichiura (1–999 EPG, 1000–9999 EPG and ≥10,000 EPG), respectively, and these can be included in analyses based on individual-based formulae (see below) as levels within a factor accounting for pre-intervention FEC. Alternatively for treatment group-based formulae, ERR can be calculated separately for each class of pre-intervention intensity. However, it is not always possible to find sufficient numbers of subjects that fit these different intensity levels, especially the moderate and high intensity categories.

Important differences in ERR have been observed depending on the formula used to calculate the reduction in FEC. Individual-based formulae are likely to be less accurate (underestimation of the ERR) and less precise (large 95% confidence intervals) compared to group-based formulae (Cabaret and Berrag, 2004; Levecke et al., 2011b; Verbruyse et al., 2011). In addition to this, Dobson et al. (2009) also reported that that ERRs calculated using geometric means are more prone to bias than ERs calculated using arithmetic means. Therefore, group-based ERR calculated using arithmetic means are the best available indicator of drug efficacy, and should be adopted in future monitoring and evaluation studies of large-scale anthelmintic treatment programmes, and as indicated above, presented for each pre-intervention intensity category separately as well as for the all subjects combined.

7.2. In vitro biological assays

A number of tests are available for the measurement of drug sensitivity in livestock nematodes using free-living life stages. These include the egg hatch test (EHT; Dobson et al., 1986; Coles et al., 2006) larval development assay (LDA; Hubert and Kerboeuf, 1992; Gill et al., 1995), motility and/or migration assays (Gill et al., 1991; Kotze et al., 2006), and a recently described larval arrested morphology assay (LAMA; Kopp et al., 2008). The most suitable test for use with the human STH will depend on several factors, including the drug group(s) of most interest, practical factors associated with the ease of use of the various assays in the field, and, most importantly, which tests are shown to be most effective at indicating the presence of drug resistance with the human STH. A significant issue in the potential for application of these tests to human STH is that, in their current form, they are only applicable to hookworms. The assays require either the isolation of eggs or infective larval stages, and whilst this is possible with hookworms as their eggs hatch rapidly after isolation from faeces, it is not relevant for A. lumbricoides and T. trichiura which develop to the infective stages within the egg and do not hatch until ingested by the host. Further work will be required to develop drug sensitivity assays for these species. One possibility may be to modify the LDA for these species to examine larval development within the eggshell (Eriksen, 1990) rather than following hatch as with the current livestock LDA.

The World Health Organization (1999) has promoted the evaluation and standardization of the EHT in human helminthology, and the test has been used a number of times to describe BZ sensitivity in human hookworms (De Clercq et al., 1997; Albionico et al., 2005; Kotze et al., 2005). These latter two studies showed that the EHT is a reproducible test under field conditions, and that it was effective in quantifying drug sensitivity in drug-susceptible hookworms. Kotze et al. (2009) subsequently examined the responses of canine and human hookworms using a 96-well assay format as a possible means for standardising the use of EHTs at different field sites. The assay format they described has subsequently been used to effectively define hookworm BZ sensitivity at field sites in Brazil (Kotze et al., unpublished data) and China (Kotze et al., 2011). This latter study highlighted logistical difficulties that may arise in sample collection and processing procedures of human faeces compared to livestock, and a number of measures for the use of EHTs were recommended for improving study outcomes in future field-based human studies. These measures included the need to monitor the degree of egg embryonation in egg samples where the local sampling methods required the use of an overnight period for children to produce a stool, the need to account for the impact of this on...
IC₅₀, IC₅₅, and IC₉₀ values, and the pooling of stool samples from groups of 5–10 individuals within a study population.

Although, in studies carried out to date using the EHT it has been possible to define a measure of BZ sensitivity in the field, a great deal of work remains to be done to validate the test as a resistance diagnostic tool for human hookworms because there are no known drug resistant isolates available to study correlations between the results of EHT and drug efficacy data as determined by FEC reduction. In addition, more studies are needed to further standardise the technique so that data from different field sites, or from different time points, can be compared with confidence. If the test proves able to discriminate between susceptible and resistant populations, there is a need to define concentration thresholds that indicate the presence of resistance. This has been done for the livestock species, thereby allowing the adoption of EC₅₀ and EC₉₀ values as being indicative of resistance (Coles et al., 1992, 2006). More baseline data from different human hookworm populations, as well as some assay data using known resistant populations, are required in order to define such thresholds. Even though the EHT is fairly standardised in veterinary medicine, evidence of variability between laboratories exists, and several factors need to be standardised/agreed upon in order to minimise variation (von Samson-Himmelstjerna et al., 2009). Preparation of assay reagents in a ‘quality controlled environment’ by a limited number of laboratories, and their distribution to field sites may help overcome some of these inconsistencies (Kotze et al., 2009).

The LDA has also been used for resistance diagnosis in the live-stock industries (Hubert and Kerboeuf, 1992; Coles et al., 1992; Gill et al., 1995). An LDA kit was sold commercially in the 1990s and early 2000s in Australia as the DrenchRite™ test for detection of resistance to BZ, nicotinic agonist, and ML drugs in sheep nematodes (Lacey et al., 1990). Although it was effective in detecting resistance to the first two of these drug groups in all the important nematode species, and detecting ML resistance in H. contortus, its use declined greatly as it was not effective in identifying ML-resistant Osterotagia spp. (Lloyd, 1998; Palmer et al., 1998) However, given its usefulness with the livestock species and the two drug groups of most importance for human hookworms, that is, the BZ and nicotinic agonists, there is potential in its use in the human field. An advantage of using an LDA would be that BZ and nicotinic agonists could be examined in a single assay, whereas the EHT is only applicable to BZ. Although IVM sensitivity could also be measured simultaneously, as mentioned above, the LDA was unable to detect ML resistance in Ostertagia spp.

Recently, a motility assay and a novel assay (LAMA) were developed to measure pyrantel sensitivity in canine hookworms (Kopp et al., 2008). These assays were subsequently examined with human and rodent-adapted human hookworms, allowing for a standard assay format of drug concentrations to be described (Kotze et al., 2009). However, as described above for the EHT and LDA, the ability of the assays to diagnose PYR resistance with human hookworms will remain unknown until they can be tested against isolates of human hookworm known to be resistant to the nicotinic agonist drugs. An important note here is that unlike the EHT, these motility/LAMA tests have been shown to distinguish between susceptible and resistant canine hookworms due to the availability of such isolates with this species (Kopp et al., 2008). In contrast, no BZ-resistant canine hookworm isolates have been reported in the literature, and these would be useful for providing a degree of validation for the EHT. An advantage of these motility/LAMA tests is that the collection of the biological material, namely infective stage larvae after migration from faecal cultures, is much less laborious than the purification of eggs from faeces as is required for the EHT and LDA. It is apparent that development of a high throughput, robust quantitative assay for determining drug response would represent a major advance. In this respect, a number of recent reports of prototype assays have been published (Smout et al., 2010; Chen et al., 2011). Whether these assay platforms are suitable for field deployment remains to be established.

It is clear that further work is required to develop and validate the use of in vitro biological assays for diagnosis of resistance with human STH. These assays represent an intermediate stage in sensitivity between the ERR and molecular tests. However, they are more immediately applicable than the molecular tests which, as described above, will take some time before they can relate allele frequencies to drug efficacy. The in vitro diagnostic tests therefore represent a means for detecting the emergence of resistance in hookworms in the crucial period before molecular tests become available. Assessment of BZ sensitivity with hookworms using the EHT is the most advanced test and could be used more widely in the field now. Although questions remain as to its ability to diagnose resistance in human hookworms, its utility with livestock nematodes (Coles et al., 1992, 2006) should generate confidence in its usefulness with the human hookworm species. The priorities should be further standardisation of methods, the collection of baseline data, and the use of the assays with known, or suspected, resistant isolates when these become available through either identification in current efficacy trials (for example, Vercruysse et al., 2011), or through laboratory-based selection. The second area of interest should be the examination of the assays suitable for the nicotinic agonist drug group. Of these, the combination of a motility and arrested morphology assay may offer the most promise (Kopp et al., 2008; Kotze et al., 2009). As mentioned above, a further priority should be the development of suitable in vitro biological tests for T. trichiura and A. lumbricoides.

8. Recommendations for monitoring drug efficacy

The Guidelines on how to monitor drug efficacy in public health were published by the WHO in the late 1990s (World Health Organization, 1999). These provided recommendations on sample size (~200 infected subjects equally representing low, moderate and high levels of infection intensity), stool sampling (2 stool samples of 2 different days both before and after administration of drugs), the FEC method (Kato-Katz thick smear with a detection limit of 24 epg) and thresholds defining reduced efficacy (ERR <70% for A. lumbricoides and <50% for T. trichiura and hookworm). However, as discussed above these guidelines need to be revised in the light of the information that has come to hand since then and the data critically reviewed in this paper. Another important implication of the current guidelines is the additional amount of technical and financial resources that are required to undertake studies to monitor anthelmintic efficacy under these WHO guidelines. Based on the cost assessment of the Kato-Katz thick smear for STH diagnosis by Speich et al. (2010), it can be deduced that such stool sampling would require US$ 3.46 per subject.

In the light of evidence accrued since the publication of these guidelines for monitoring drug efficacy, agreement on the most appropriate and valid monitoring of chemotherapy-based helmint control programmes is urgently required. The approach should be standardised and consistent throughout so that changes in drug efficacy with time can be readily identified and that the outcomes from trials in different years and locations can be legitimately compared.

For this purpose we propose the following algorithm for monitoring drug efficacy in future trials (Fig. 1). The monitoring of trials should be based on a target random minimal sample of 200 subjects, independent of their infection status (negative/positive), level and variation in egg excretion. The number of subjects to recruit, however, will depend largely on the proportion of the subjects meeting the inclusion criteria (between 6 and 12 years,
Fig. 1. Flow chart on how to monitor and to interpret drug efficacy during mass drug administration programmes. *N is based on the formula below

\[ N = \frac{(100\% - \text{expected proportion of subjects not meeting the inclusion criteria in } \%) \times (100\% - \text{expected dropout in } \%)}{\left(\frac{0.95}{0.95}\right)^2} = 222. \]

For example, for 5% of the subjects not meeting the inclusion criteria and an expected dropout of 5% recruitment of 222 subjects is recommended. **At present, there is no inference available to calculate 95% confidence interval for the ERR formula. Bootstrap analysis, therefore is recommended. ***Baseline FEC of all subjects examined are lowly aggregated when \( k > 0.01 \). \( k = \frac{\text{arithmetic mean of baseline FEC}^2}{\text{variance baseline FEC} - \text{arithmetic mean of FEC}} \)****Based on baseline FEC of infected subjects only.
absence of a severe incurrent medical condition or diarrhoea) and the compliance rate at follow-up. Two stool samples per subject should be examined (1 at baseline and 1 at follow-up). The recommendation of a reduction in stool samples per subject (previously 2 at baseline and 2 at follow-up according to WHO guidelines) may seem surprising but it is justified, because any variation in FEC across time (some individual FEC will be overestimated) will have no or very little impact on the ERR based on the arithmetic mean FEC at the group level. The above approach therefore represents a compromise between the collection of unnecessarily detailed data with associated cost and time, and achieving what is practical in a field setting with expectation of good compliance from the subjects involved.

On account of their simplicity, the McMaster egg counting ([Levecke et al., 2009] and the Kato-Katz thick smear methods ([World Health Organization, 1991]) are the FEC methods of choice in large-scale surveys. Although the Kato Katz is more sensitive and detects higher FEC for A. lumbricoides infections ([Levecke et al., 2011a]), it is not a method of choice where hookworms are highly prevalent and the major subjects of the trial. Yet, both methods provide comparable ERR results for all STH (unpublished results, Albonico et al.). The FEC of the subjects at follow-up should be assessed at 7–14 days after the administration of drugs. The ERR should be calculated subsequently based on the ratio of the arithmetic mean of FEC at follow-up and the arithmetic mean of baseline FEC ([Vercruysse et al., 2011]). Subsequently, 95% confidence intervals should be calculated based on bootstrap analysis (10,000 iterations).

Vercruysse et al. (2011) recommended that in future monitoring of single-dose ALB-dependent control programmes a minimum ERR (based on arithmetic means) of 95% for A. lumbricoides and 90% for hookworms should be used as current thresholds, independent of the baseline FEC. For T. trichiura, the thresholds depend on the mean baseline FEC of the infected subjects, being 90%, 70% and 50% for trials where the mean baseline FEC is less than 275 EPG, 550 and 800 EPG, respectively ([Levecke et al., in press]). Any reduction from these proposed thresholds (upper limit of 95% CI < threshold) should trigger the need for further investigation. However, it is important to note, that ERR can only be interpreted when the baseline FEC across the 200 subjects shows a low degree of aggregation (k of negative binomial distribution >0.01, Fig. 1) and/or the group comprises a minimum of at least 30 infected children. Given the above, and for the sake of consistency in future reporting, trial investigators should report the ERR in all three FEC intensity classes for all species, as well as for the combined dataset.

In the first instance, programme managers and investigators should investigate the performance of key aspects of the control programme, including (1) coverage of the programme (the proportion of target children that received drugs), (2) compliance of the children (the proportion of children that received the tablet and actually swallowed it) and (3) the source of and storage conditions of the anthelmintic (allowing the presence and quantity of active ingredient, dissolution and disintegration times to be subsequently tested in the event of apparent low efficacy). If these potential confounding factors can be ruled out, confirmation of AR by either in vitro biological or molecular assays is strongly recommended.

9. Conclusions, research priorities and recommendation for SOPs

The limited number of studies in the public domain that have reported reduced efficacy of anthelmintics does not yet provide conclusive evidence of drug resistance among human STH, but the warning signs are clear. Moreover, research is critically needed to understand the genetic and molecular basis of anthelmintic resistance in humans; while putative resistance alleles have been found in human STH, so far there are no studies investigating changes in the frequency of resistance mutations following drug pressure (Box 2).

The few studies of AR in STH carried out so far have been confounded by methodological flaws e.g. type of diagnostic methods, treatment regimens, drug dose rates, geographical location. There is a lack of adequate guidelines through which managers of control programmes can be informed how best to react should treatment failures occur in public health settings. Surveys for monitoring drug efficacy need to be undertaken regularly implementing standard operating procedures (SOPs) that will ensure consistency and uniformity.

As the recommended monitoring indicator highlighted in this review (that is, ERR) is dependent on the intensity of pre-treatment STH infection, especially in the case of T. trichiura, we believe that this confounding effect is best avoided by reporting ERR separately for each intensity class, as well as overall, and we strongly recommend that this be adopted as a SOP for all species (not just T. trichiura) to avoid inconsistencies in reporting. A clear definition of drug resistance with evidence-based thresholds for each STH is one of the most pressing needs, and the values we have provided here for different baseline FECs in the case of T. trichiura and overall in the case of A. lumbricoides and hookworms, against which the obtained ERR, are a useful starting point. As more data become available, once SOPs have been implemented and become established and universally used, it should be possible in time to refine the actual values, but for the present those given here should be implemented in all trials (SOP). In order to generate confidence in the robustness of the ERR in each intensity class, calculations should only be reported for intensity classes where at least 30 subjects were treated (SOP). The recommendations suggested in this review should be translated now into practical algorithms for application in the field and should serve as the platform of an eventual more comprehensive document that the Department of Neglected Tropical Diseases at the WHO is requesting the WHO collaborating centres to produce in order to address this pressing yet neglected issue.

Given the focus of this review on identifying AR in human STH, it is important to emphasise that should AR be detected and confirmed as such, there are is still no agreed strategy in the public domain, as to how to deal with it. This now requires urgent consideration and agreement. It is pertinent that only a limited number of anthelmintic classes are available for treatment of human STH. Therefore, there is an urgent need to develop new anthelmintics that work through novel modes of action. Already

<table>
<thead>
<tr>
<th>Subject</th>
<th>FEC at baseline (EPG)</th>
<th>FEC at follow-up (EPG)</th>
<th>ERR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>10000</td>
<td>9.2</td>
<td>1000</td>
</tr>
<tr>
<td>B</td>
<td>1000</td>
<td>6.9</td>
<td>100</td>
</tr>
<tr>
<td>C</td>
<td>100</td>
<td>4.6</td>
<td>10</td>
</tr>
<tr>
<td>D</td>
<td>10</td>
<td>2.4</td>
<td>15</td>
</tr>
<tr>
<td>E</td>
<td>10</td>
<td>2.4</td>
<td>1</td>
</tr>
<tr>
<td>AR mean</td>
<td>2224.0</td>
<td>5.1</td>
<td>225.2</td>
</tr>
<tr>
<td>GEO mean</td>
<td>164.0</td>
<td>31.4</td>
<td></td>
</tr>
</tbody>
</table>

CR = 100% × (1–AR/C2) = 0.0%.
ERR[1] = 100% × (1–225.2/2224.0) = 89.9%.
ERR[2] = 100% × (1–31.4/164.0) = 80.9%.
ERR[3] = 62.0%.
Box 1
Definition and formulae for calculating the cure rate (CR) and egg reduction rate (ERR)

CR and ERR are the most applied metrics for assessing anthelmintic efficacy. The CR equals the percentage of subjects who no longer pass eggs after treatment, whereas the ERR equals the reduction in the number of eggs excreted. Formulas for calculating these metrics are provided below. In contrast to the CR, various formulae have been used for calculating the ERR. These formulae differ with respect to the mean (arithmetic vs. geometric) and statistical unit under consideration (group vs. subject). ERR[1] and ERR[2] provide ERR at the level of the group, whereas ERR[3] provides ERR at the level of the subject (= mean of individual ERR).

\[
CR = 100\% \times \left(1 - \frac{\text{number of subjects excreting eggs at follow-up}}{\text{number of subjects excreting eggs at baseline}} \right)
\]

ERR[1] = 100\% \times \left(1 - \frac{\text{arithmetic mean of group FEC at follow-up}}{\text{arithmetic mean of group FEC at baseline}} \right)

= 100\% \times \left(1 - \frac{\sum_{i=1}^{n} \text{FEC at follow-up}_i}{\sum_{i=1}^{n} \text{FEC at baseline}_i} \right)

ERR[2] = 100\% \times \left(1 - \frac{\text{geometric mean of group FEC at follow-up}}{\text{geometric mean of group FEC at baseline}} \right)

= 100\% \times \left(1 - \frac{\prod_{i=1}^{n} \text{FEC at follow-up}_i^{1/n}}{\prod_{i=1}^{n} \text{FEC at baseline}_i^{1/n}} \right)

ERR[3] = 100\% \times \left(1 - \frac{\text{arithmetic mean of individual FEC at follow-up}}{\text{arithmetic mean of individual FEC at baseline}} \right)

= 100\% \times \left(1 - \frac{\sum_{i=1}^{n} \text{FEC at follow-up}_i}{\sum_{i=1}^{n} \text{FEC at baseline}_i} \right)

An example is provided in Table 1, to illustrate in more detail the different metrics/formulae and their consequencs. In this example, the efficacy of a drug against T. trichiura was evaluated in five subjects (A–E). At baseline, all subjects excreted eggs, with the FEC ranging from 10 to 10,000 EPG (arithmetic mean = 2224.0 EPG, geometric mean = 164.0 EPG). At follow-up, eggs of T. trichiura were found in all subjects, with the FEC ranging from 15 to 1,000 EPG (arithmetic mean = 225.2 EPG, geometric mean = 31.4 EPG). The CR in the present example was 0.0%, the ERR ranged from 62.0% to 89.9%, depending on the ERR formula used. ERR[3] yielded the lowest ERR result, which can be explained by the specific ERR of subject D. For this subject, the FEC at follow-up was higher than those at baseline resulting in a negative ERR, which is in sharp contrast to the remaining four subjects for whom high ERR were observed. In this ERR formula, however, each of the subjects included provides an equally weighted contribution to the outcome, and hence interpretation is highly sensitive to extreme values (see also Vercruysse et al., 2011). Group based formulas are more robust with respect to extreme values, yet the difference between ERR[1] and ERR[2] is still remarkable. A detailed study explaining these differences was conducted by Dobson et al. (2009), highlighting that ERR[2] is more prone to bias than ERR[1].

Box 2
Research priorities to mitigate and monitor anthelmintic resistance for human soil-transmitted helminths

- Standardization of methods for determining ERR, including:
  - Quantitative parasitologic technique
  - Quantity of stool analysed
  - Statistical methodology
- Standardizing trial design and monitoring strategy to detect anthelmintic resistance, including:
  - Population size
  - Infection intensity strata
  - Thresholds for defining suboptimal response
- Comprehensive survey of resistance-determining alleles in STH
  - E.g., the codon 200 polymorphism in β-tubulin
  - Development of “user-friendly” and low cost molecular assays for detection of resistance allele
- Mathematical modelling of the effect of various factors on the evolution of anthelmintic resistance, including:
  - Dose regimen and clinical efficacy
  - Background resistance allele frequency
  - Refugia
- Development and standardization of in vitro assays of AHR
  - Assay methodology
  - Thresholds for defining suboptimal response
  - Developing new tests to measure drug resistance phenotype in STH that do not hatch in the environment (A. lumbricoides and T. trichiura)
- Using modern techniques to define the pharmacokinetics in humans of relevant anthelmintics
- Improvements to the current armoury of anthelmintics
  - Optimization of dosages, dose rates and treatment schedules with current anthelmintics
- Assessment of combination therapies for efficacy against human STH
- Development of novel anthelmintics that are unlikely to share resistance mechanisms with current anthelmintics (e.g. operate through different modes of action)
- Collection of resistant lines of human STH
  - For use as positive controls in in vitro assessments of field samples
  - For investigation of the molecular and genetic basis of AR

in the veterinary field a limited number of such candidates, meeting some of the criteria for selection, have been developed in recent years but none are ideal for treatment of STH e.g. the two veterinary products, emodepside (currently very expensive) and monepantel (activity against STH is currently being assessed, personal communication), and tribendimidine and nitazoxanide, which have been approved for human use but still have limited utility (Xiao et al., 2005; Olliaro et al., 2011). Should AR arise imminently and spread, at a time when our anthelmintic armoury is still so limited, we may have to resort to traditional methods of control based on improvements in sanitation, education to change human behaviour and to ensure that footwear is more widely adopted in the case of hookworm control. The worrying aspect of this, of course, is that we have known about the relevance of these latter control measures for well over a century, but STH have persisted nevertheless.
References


Montresor, A., 2011. Cure rate is not a valid indicator for assessing drug efficacy and
Monteiro, A.M., Wanyangu, S.W., Kariuki, D.P., Bain, R., Jackson, F., McKellar, Q.A.,
Levecke, B., Rinaldi, L., Charlier, J., Maurelli, M.P., Morgoglione, M.E., Vercruysse, J.,
detection of resistance to macrocytic lactones in Haemonchus contortus, and


