Measuring the impact of antimicrobial stewardship programs

Jan-Willem H. Dik, Ron Hendrix, Randy Poelman, Hubert G Niesters, Maarten J. Postma
Bhanu Sinha, Alex W Friedrich

Expert Reviews of Anti-Infective Therapy 2016; 14(6):569-575
Abstract

Antimicrobial Stewardship Programs (ASPs) are being implemented worldwide to optimize antimicrobial therapy, and thereby improve patient safety and quality of care. Additionally, this should counteract resistance development. It is, however, vital that correct and timely diagnostics are performed in parallel, and that an institution runs a well-organized infection prevention program. Currently, there is no clear consensus on which interventions an ASP should comprise. Indeed this depends on the institution, the region, and the patient population that is served. Different interventions will lead to different effects. Therefore, adequate evaluations, both clinically and financially, are crucial.

Here, we provide a general overview of, and perspective on different intervention strategies and methods to evaluate these ASP programs, covering before mentioned topics. This should lead to a more consistent approach in evaluating these programs, making it easier to compare different interventions and studies with each other and ultimately improve infection and patient management.
Introduction

Antimicrobial resistance is a growing problem. One of the major drivers of this disconcerting development is inadequate use of antimicrobials, both in healthcare centers and out-patient settings (Goossens, 2009), as well as in livestock (Marshall and Levy, 2011). The so-called One Health approach targets resistance development on all the before-mentioned levels. Such a broad approach is considered crucial, in order to effectively minimize the worldwide healthcare antimicrobial resistance threat (Renwick, et al., 2016). Part of this approach is improving the usage of antimicrobials in healthcare centers and out-patient settings, which in turn helps reducing resistance development (World Health Organization, 2012). Antimicrobial Stewardship Programs (ASPs) are being hailed as a solution to improve antimicrobial therapies and thus results in a better patient outcome and safety. Different national and international guidelines are available for hospitals, long-term care facilities and general practitioners (de With, et al., 2016; Dellit, et al., 2007; SWAB, 2012). There is, however, no clear consensus on the impact of different interventions (Davey, et al., 2013; Schuts, et al., 2016). Effects (clinical and financial) in specific settings or patient populations are difficult to compare and/or evaluate. Some interventions might even be redundant or counterproductive, although in general, published results are often favorable (Davey, et al., 2013; Schuts, et al., 2016). This inconclusiveness, necessitates performing scientifically sound (cost-)effectiveness studies on ASP interventions (McGowan, 2012). There are multitudes of methods to evaluate several interventions, but in general, they lack uniformity (Evans, et al., 2015; McGowan, 2012; Morris, 2014). In this review, we will discuss stewardship in general, its (pre)requisites and the main interventions and their approaches for evaluation, thereby giving a general and up-to-date overview.

Importance of a broad stewardship program

The term “antimicrobial stewardship” has been coined roughly 20 years ago (McGowan and Gerding, 1996). Stewardship programs are now being implemented worldwide and hundreds of articles are published yearly (Howard, et al., 2015). As it became clear that inadequate antimicrobial use (prophylactic and therapeutic) contributes to resistance development, improving antimicrobial usage became a focus for many healthcare institutions, using a subset of different interventions (Davey, et al., 2013; Dellit, et al., 2007). However, it is often be overlooked that reducing resistance rates should not be the primary goal. The ultimate goal is to with improve clinical outcome and patient safety by providing optimal patient care. Patients should be the main focus and they have important questions related to infectious problems: i) How can I be protected from a (resistant) infection? ii) Do I have an infection, and if yes, what is causing it? iii) What is the optimal treatment to cure it? Answering these questions requires a broader approach than just an ASP and also entails that certain requirements are met. Besides implementing an ASP, an Infection prevention Stewardship Program (ISP) should be present to ensure that other patients do not get infected by pertinent (resistant) pathogens, which are
often easy transmissible. Furthermore, optimal and timely diagnostics that can adequately and rapidly diagnose the patient’s problems are vital (Diagnostic Stewardship Program [DSP]). Only if all three aspects are covered - optimal treatment, prevention and diagnostics (an integrated, Antimicrobial, Infection prevention & Diagnostic [AID] stewardship program) - and all involved stakeholders have the necessary meta-competence (meaning a broad understanding of all relevant above mentioned aspects), healthcare centers can optimally treat infectious patients and tackle the development of antimicrobial resistance (Dik, et al., 2016b; Lammie and Hughes, 2016). Because patient transfers between institutions are also pathogen transfers (Donker, et al., 2010), these three aspects should not only be covered within one local center, but in a (regional) healthcare network. This entails close collaboration of all healthcare facilities (i.e. hospitals, but also general practices and long-term care facilities) within a clearly defined region (Ciccolini, et al., 2013). Harmonization of guidelines and practices can be a first start regarding this aspect (Müller, et al., 2015). In the near future, such an integrated AID-approach should lead to a more personalized treatment plan, which is optimally adapted to the specifics of each single patient.

**Importance of diagnostics**

It is thus important that adequate diagnostics are performed on time, and provide rapid results to have impact on patient care (Caliendo, et al., 2013). Ideally, results, including resistance patterns, should be available before the patient is started on antimicrobial treatment. Three parameters influence the value of diagnostic tests: quality, cost and time. Overall, the sensitivity and specificity of new commercial and often multiplex-based molecular, Point-Of-Care (POC) assays approach the quality of Laboratory Developed Tests (LDTs). In this situation, lower costs and/or shorter turnaround times become the main drivers for increased value. From a managerial point of view, we introduced the euro-hour (€hr) concept, comparable with kilowatt hour to easily visualize the impact of both parameters (Dik, et al., 2016b). In this concept, the costs of a test are multiplied by the turnaround time and therefore represent the impact of implementing a POC test. POC tests can only have an impact on antimicrobial therapy and patient management if results are timely available, interpreted and followed-up by a medical specialist (Buehler, et al., 2016; Rogers, et al., 2015). When implemented, it increases the probability of a correct (preliminary) diagnosis – including the reduced need for further diagnostics, streamline antimicrobial treatment sooner if needed (thereby also minimizing the risk for toxicity), and improve infection prevention measurements. In the field of oncology this so-called theragnostic approach is under continuous development during recent years and it would be a powerful tool in personalized infection management as well (Dik, et al., 2016b; Lammie and Hughes, 2016). Examples of POC tests or rapid diagnostics (e.g. multiplex PCR and MALDI-TOF MS) already implemented for ASPs are: MRSA screening and testing (Mather, et al., 2016; Tacconelli, et al., 2009b; Wassenberg, et al., 2012); resistance screening (Evans, et al., 2016); the use in septic/bacteremic patients (Banerjee, et al., 2015; Bauer, et al., 2010a; Clerc, et al., 2014); use of biomarkers (of which procalcitonin probably shows the most
promising results) (Bouadma, et al., 2010; de Jong, et al., 2016; Nobre, et al., 2008; Schuetz, et al., 2012); and with viral infections, such as for respiratory illness (Müller, et al., 2015; Popowitch, et al., 2013).

**Basics of ASPs**

Often ASP interventions are subdivided into three groups: the front-end approach, the back-end approach and supplemental interventions. Front-end interventions focus on the start of empirical therapy such as pre-analytic consultations and guidelines for (empiric) therapy. Back-end interventions focus on optimization of therapy after e.g. two or three days. For example, an intravenous to oral switch promotion; de-escalation; or timely stop of therapy, as appropriate. Finally, there are interventions that supplement an ASP such as using resistance data to keep local guidelines up to date and the availability of educational programs (Dellit, et al., 2007). The evaluation of the program is also an important aspect of the latter group. Such a different array of interventions implies that the timeline of impact of an ASP varies broad, with effects on the short, middle and long term (see Figure 4.1 for a schematic overview).

An ASP should focus on improving patient care and safety, by increasing appropriateness of all antimicrobial use (i.e. prophylactics, empiric therapy and directed therapy). When looking at prophylactic antimicrobials, there are special cases like Selective Digestive (or Intestinal) Decontamination (SDD) and Selective Oropharyngeal Decontamination (SOD). SDD and SOD are generally implemented at ICUs, and thus are often not part of a standard ASP. They are, however, important forms of antimicrobial use that can influence resistance development, and will impact overall policy of antimicrobial usage. It is thus imperative to take these interventions into account when implementing (and evaluating) antimicrobial use and/or interventions strategies (Daneman, et al., 2013; Plantinga and Bonten, 2015).

**Unresolved issues with evaluating ASPs**

The recent systematic Cochrane review on interventions to improve antimicrobial therapy systematically looked at all studies describing one or more interventions, and evaluated their quality and strength of evidence (Davey, et al., 2013). This review mainly focuses on the clinical effects. One of the main conclusions that can be drawn from the review is the lack of quality of evaluations reported. This is exemplified by the fact that the majority of the studies could not be included (Davey, et al., 2013). A recent highly comprehensive systematic review and meta-analysis of 14 different antimicrobial stewardship objectives found similar results, albeit generally with a low quality of evidence (Schuts, et al., 2016).
Financial effects are equally important. In 2015 two reviews on financial evaluations of ASP studies were published. Both conclude that ASPs are evaluated inconsistently and often even poorly, making it almost impossible to compare studies with another (Coulter, et al., 2015; Dik, et al., 2015c). Keeping these results in mind, we will provide a general overview of the different methods to evaluate ASP interventions both clinically and financially (leaving out structural and process-focused aspects). Furthermore, we will mention some pros and cons for each method.

Different methods of evaluating ASPs

Randomized controlled trials (RCTs) are considered as the gold standard and most preferable type of study. However, they are often less suitable for antimicrobial intervention studies, due to logistics, ethics and costs. Nevertheless, in recent years there were a couple examples looking at ASP interventions in a randomized controlled manner (Banerjee, et al., 2015; Fleet, et al., 2014; Lesprit, et al., 2013). The large majority of published evaluations are however observational studies (e.g. case-control studies, interrupted time series analyses [ITS], etc.) (Davey, et al., 2013; Howard, et al., 2015; Schuts, et al., 2016). For these studies, comparable cohorts of patients are a major source for bias. This can be even more influenced by changes over time, because the control period is usually several years prior to the intervention period.

With regard to economic evaluations, the preferred method would be to do a cost-effectiveness/utility study from a societal perspective (Drummmond, et al., 2005). Generally speaking, the level of expertise and the time required to do such an analysis are often too scarce to be practically accessible. In practice, this has led to the fact that economic evaluations performed on ASPs are often cost-minimization analyses (Dik, et al., 2015c). Finally, of relevance is the fact that ASPs are implemented on specific wards and/or for specific patient
groups (e.g. ICUs, long-term care facilities, septic patients, pediatric patients, and MRSA infections). This makes comparability difficult and it is therefore essential to mention in detail the patient characteristics, as well as the setting of implementation.

**Clinical outcome measures**

The most important goal for an ASP should be to improve quality of patient care. A number of different measures are used that describe some aspects of clinical outcomes. Most important are mortality rates. In general, most studies that evaluate mortality conclude it is not compromised, and that an ASP is thus a safe intervention (a non-inferiority analysis). Especially in ASPs targeted at the more severe patient groups (e.g. septic patients), mortality can be an important outcome. For less severe infections (e.g. urinary tract infections), the use of mortality as an outcome measure might be less informative. Length of stay (LOS), (secondary) infection rates, and re-admission rates are also often measured and evaluated (Davey, et al., 2013; Schuts, et al., 2016). Other less frequently studied outcomes are toxicity and possible side-effects (e.g. IV catheter-related problems or phlebitis) (see Table 4.1). LOS is one of the more accessible variables to obtain. It is, however, important to take possible secular trends towards earlier discharge into account when evaluating an ASP in a case-cohort setting, especially if the time-period spans multiple years (e.g. did the institution in general saw a drop in LOS over time). Besides the overall hospital LOS, ward specific LOS (such as ICU stay) is an option. The latter is of course most interesting if the program is also ward specifically implemented. If treatment improves and infections are cured more effectively, the re-lapse rates may decrease and re-admission rates consequently will go down. However, this outcome measure is biased if there are other hospitals in the vicinity where patients might be readmitted, for example, in clusters of academic centers and surrounding general hospitals. As a more indirect effect, the infection rate for *Clostridium difficile* can be taken as an outcome measure (see e.g. Nathwani et al, for a successful program (Nathwani, et al., 2012)). In some studies, a direct correlation of *C. difficile* infection with antimicrobial use is suspected, especially regarding cephalosporins and clindamycin (Slimings and Riley, 2014). This rate might consequently be used as an indirect indicator for antimicrobial use and therefore as an ASP outcome measure.

**Microbiological outcome measures**

Besides the important clinical outcomes (that directly impact patient care and safety), a secondary important goal is the reduction of resistance levels (see Table 4.1). There are multiple ways to evaluate this goal. Resistance levels can be measured as percentage of patients (or cultures) with microorganisms “resistant” for a certain antibiotic compared to the number with “susceptible” microorganisms. This parameter can be measured in infected patients or colonized patients. Furthermore, the rate of infections with a resistant microorganism can be
taken as a measure (preferably as a percentage of patients infected with the susceptible variant). Difficulties arise due to the longer time that is required before a change in resistance levels can be observed. In addition, reliability of certain trends in the data is difficult to estimate when looking at small numbers. The long time-frame also implies that the influence of possible confounders becomes greater. These slow, subtle changes make that antibiograms have been shown to be inconclusive as separate outcome measures and the application of an interrupted time series analysis is therefore a better and preferred method for resistance measures (Schulz, et al., 2012). Furthermore, the baseline level of resistance is also of importance: countries with high resistance levels will most likely see larger effects in a shorter time, provided there is no major influx of resistant microorganisms. A final complicating factor is that “resistant” bacteria reside in the community and in neighboring healthcare centers. If an ASP is not implemented regionally, positive results might not be achievable, namely at referral centers. In this setting a majority of the patients carrying “resistant” bacteria, will come from the surrounding healthcare region (Donker, et al., 2010; Donker, et al., 2015). The quality of evidence of ASP effects on resistance is still low (Davey, et al., 2013; Tacconelli, et al., 2016). To improve studies specifically looking at correlating antimicrobial use and resistance development, a Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) tool was developed (STROBE-AMS) (Tacconelli, et al., 2016). Concerning SDD/SOD, there is still discussion regarding possible resistance development and it is therefore highly advisable to monitor possible resistance development when implementing such an intervention (Health Council of the Netherlands, 2015).

Antimicrobial consumption outcome measures

ASPs mainly focus on accomplishing changes in broad-spectrum IV prescriptions, because broad spectrum antimicrobials are more likely to promote resistance development and IV treatment is more likely to cause secondary infections/complications. The programs focus on either support of the prescribing physician at the start during decision making regarding therapy, or after a few days to support the evaluation of diagnostics and subsequent possible adjustments of therapy. An IV to oral switch program is one of the most frequently implemented interventions (Nathwani, et al., 2015). To measure the effect on therapy different outcome measures can be used (see Table 4.1). Often it is chosen to quantify the antimicrobial therapy as Defined Daily Doses (DDDs, as defined by the WHO), with a denominator correcting for clinical activity such as bed days or admissions (http://www.whocc.no/ate_ddd_index). This can be done either by looking at dispensing data or at purchasing data, which are strongly correlated with each other (Tan, et al., 2016). From a national point of view, a broader denominator such as inhabitants of the healthcare region is also valuable. Because, as mentioned before, people are transferred through different healthcare centers within a region, transferring bacteria with them as well. IDSA/SHEA advocates DDDs per 1000 patient days as universal outcome measure for ASP programs (Dellit, et al., 2007). This is, however, not always suitable as it is a highly generalizing method
that does not take into account patient specifics (such as complicated infections which require high dose therapy as for e.g. endocarditis), and is known to overestimate the use (de With, et al., 2009). With respect to pediatric populations, these outcomes are far from optimal, because they are based on adult dosages. This should be taken into account, although up to now it is unclear what measure should preferably be used instead of DDDs (Fortin, et al., 2014). It should thus be noted that a change in DDDs is not entirely suitable for drawing conclusions on the success of an ASP. Optimal antibiotic therapy can also mean that undertreated patients should receive more antibiotics (for example deep-seated infections or overweight patients).

**Table 4.1: Overview of different outcome measures and some general remarks.**

<table>
<thead>
<tr>
<th>Outcome measures</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clinical</strong></td>
<td></td>
</tr>
<tr>
<td>Mortality</td>
<td>Important, but less suitable for mild infection (e.g. uncomplicated UTI)</td>
</tr>
<tr>
<td>LOS</td>
<td>General or ward-specific (e.g. ICU stay); easy to obtain, but highly sensitive to biases</td>
</tr>
<tr>
<td>Complications</td>
<td>Complications such as IV catheter-related problems and phlebitis</td>
</tr>
<tr>
<td><em>Clostridium difficile</em></td>
<td>Also indirect measure for antimicrobial use</td>
</tr>
<tr>
<td>Re-admission rates</td>
<td>Due to relapse, but important to consider effect of neighbouring institutions</td>
</tr>
<tr>
<td>Other complications</td>
<td>For example, IV catheter-related thrombosis and non-infective phlebitis</td>
</tr>
<tr>
<td><strong>Microbiological</strong></td>
<td></td>
</tr>
<tr>
<td>Resistance levels</td>
<td>Difficult to measure due to generally long time frame (months to years)</td>
</tr>
<tr>
<td><strong>Antimicrobial consumption</strong></td>
<td></td>
</tr>
<tr>
<td>Total use</td>
<td>Often measured in DDDs however, discrepancies are known due to generalization</td>
</tr>
<tr>
<td>IV/PO ratio</td>
<td>Of interest with an active IV-to-PO switch program</td>
</tr>
<tr>
<td>Broad/narrow ratio</td>
<td>Potentially relevant with regard to resistance development</td>
</tr>
<tr>
<td>Appropriateness of therapy</td>
<td>Labour intensive and possibly subjective, but of importance</td>
</tr>
<tr>
<td><strong>Financial</strong></td>
<td></td>
</tr>
<tr>
<td>Cost-benefit ratio</td>
<td>Preferably done as cost-effectiveness study, including all costs and benefits (at least at hospital level, but preferably at societal level)</td>
</tr>
</tbody>
</table>

UTI: urinary tract infection; LOS: length of stay; ICU: intensive care unit; IV: intravenous; DDD: daily defined dosage; PO: per os
Personalized therapy measures such as Prescribed Daily Doses (PDDs) or Recommended Daily Doses (RDD) are a more patient specific approach to quantify antimicrobial treatment and might give more suitable results (de With, et al., 2009; Gagliotti, et al., 2014). Furthermore, the length of the therapy (in days) can be evaluated (duration of therapy, DOT). A discrepancy when compared to DDDs is known, due to the difference between administered dose and the WHO DDD values (Polk, et al., 2007). Because an ASP focuses on optimizing therapy, often by promoting narrow spectrum oral medication, it can be worthwhile evaluating effects of these interventions specifically. This can be done by looking at the percentage of IV medication versus oral (in DDDs/PDDs/RDDs or DOTs) and/or the percentage broad versus narrow spectrum antibiotics (in DDDs/PDDs/RDDs or DOTs). If it is expected that improvements on a ward will be more systemic of nature, it is also worthwhile to evaluate the percentage of patients receiving antimicrobial therapy. Finally, appropriateness of therapy can be evaluated. This is a more labor intensive method, while it often requires reviewing single patient’s files, making it also less objective than ‘hard numbers’ such as DDDs (DePestel, et al., 2014). However, because the goal of an ASP is optimal therapy (according to protocol), appropriateness of therapy is an outcome measure that directly evaluates the main goal of the intervention. An example of this, is the analysis with regard to the appropriateness therapy of urinary tract infections (Stewardson, et al., 2014).

**Financial outcome measures**

With regard to financial evaluations of an ASP, there is much room for improvement (Coulter, et al., 2015; Davey, et al., 2013; Dik, et al., 2015c). Most notable issue is that not all costs are taken into account, but just a subset of costs and benefits chosen based on data availability, potentially leading to other cost-effectiveness results. Obviously, for correct interpretation of costs and benefits of a stewardship program, it is important to take into account all costs (and benefits) besides the obvious ones (e.g. antimicrobial costs) (Drummond, et al., 2005). Preferable, this collection of costs should be done prospectively with up front agreed-upon variables and parameters. This minimizes the chances that certain cost types are neglected or cannot be informed. Often various types appear not to be collected when an evaluation is performed retrospectively. Although highly desirable, it is not always necessary or feasible to be cost saving. It is however important, to know if the intervention is the most cost-effective way to reach the preferred outcome(s), compared to other potential interventions or the baseline situation. If indeed it is not cost saving, it is worthwhile to take into account a certain threshold of maximum cost per outcome (i.e. cost or willingness to pay per quality-adjusted life year [QALY], life-year saved, or other chosen outcome) to enhance optimal allocations of budgets.

An obvious start for integrative costing is to consider all costs that had to be made to implement the program or intervention. This definitely includes time spent by the staff involved, both specifically hired and those already working in the institution. In the latter case
this formally concerns so-called opportunity costs. Furthermore, the required infrastructure (for example costs for the introduction of an IT program) and consumables (for example extra or new diagnostics) should be considered. In short, all resources and costs of running the program should be included, consistently measured by opportunity costing which reflects the alternative next-best application of these resources and costs (applying to the people involved in the intervention, but also maintenance contracts for IT programs or depreciation costs of laboratory equipment).

If implementation and daily execution costs are known, one can relate these possible savings or benefits, and draw conclusions on the cost-efficiency or cost-effectiveness. Preferably, all outcome measures that were evaluated clinically are quantified and transformed into monetary and/or utility values. In general, this will include: LOS, antimicrobial use, other procedures done to treat patients (including nursing time), changes in re-admissions, infections and other complications. Quantification into monetary values can be open to interpretation and, for example for LOS high variations in willingness to pay were shown (Stewardson, et al., 2014), meaning that proper justification is important, inclusive inspection of guidelines for pharmacoeconomic research (http://www.ispor.org).

For ASPs, costs are usually calculated from a hospital perspective regarding monetary outcomes and related to survival and quality-of-life as humanities’ outcomes. When taking a societal perspective other outcomes should also be included, for example costs due to reduced labor productivity (Dik, et al., 2015c). Unfortunately, ASP evaluations that include financial outcomes often only include direct antimicrobial costs within a very limited perspective, making it nearly impossible to draw conclusions from current literature on full cost-effectiveness of ASPs and hampering comparison with cost-effectiveness of other interventions in healthcare that are often done from this societal perspective (for example, drugs) (Coulter, et al., 2015; Davey, et al., 2013; Dik, et al., 2015c).

Conclusions and recommendations

Antimicrobial stewardship programs are an important topic with hundreds of publications appearing yearly. In the last few years, besides numerous studies focusing on antibiotics, many studies were published on antifungals with a comparable set-up as antibiotic stewardship studies and thus also comparable quality issues (Brüggemann and Aarnoutse, 2015; Muñoz, et al., 2015; Valerio, et al., 2015). Because ASPs are consisting of multiple interventions and not every healthcare center is implementing the same interventions, outcome evaluations are also highly diverse. In this respect, a maturity model can for example help to establish the current status of an ASP (van Limburg, et al., 2014). This complexity is further increased by the method of evaluation (e.g. RCT, ITS or case-control study), and different outcome measures used. Finally there are multiple challenges to obtain high quality data on effectiveness of an ASP, for example a lack of data of presumptive diagnosis at time of prescribing (or not
prescribing) antimicrobials (and subsequent evaluation of appropriateness), as well as exact timing of diagnostics versus start of therapy. Comparing and interpreting different ASP studies is therefore extremely challenging. Clearly, there is a need for appropriate and well-standardized definitions of interventions of an ASP, of the preferred method of evaluation, and of the preferred outcome measures, inclusive those from the financial-economic perspective. Until then, authors should clearly explain and discuss their methods of evaluation in order to make the field of ASPs more transparent.

**Five year view**

Within the foreseeable future, more tools will be available within daily practice to guide antimicrobial therapy in the best possible way. Faster diagnostics, genomic data, and smarter clinical decision support systems are some of these examples, as well as the growing importance of regional healthcare networks and integrative, interdisciplinary collaboration between specialists. This entails that ASPs will continue to develop and that interventions are expected to become easier to implement. Such developments also impact the evaluations of ASPs. It is therefore even more important that ASP evaluations will be performed in a transparent and comparable manner to help streamline the development process.