Improving community-acquired pneumonia treatment by antimicrobial stewardship

Valentijn Schweitzer

Improving community-acquired pneumonia treatment by antimicrobial stewardship

Het verbeteren van de behandeling van thuis-opgelopen longontsteking door middel van antimicrobial stewardship

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

dinsdag 3 november 2020 des middags te 4.15 uur

door

Valentijn Andries Schweitzer

geboren op 27 juni 1987
te Beverwijk

Promotor:

Prof. dr. M.J.M. Bonten

Copromotoren:

Dr. C.H.E. Boel

Dr. J.J. Oosterheert

TABLE OF CONTENTS

1

GENERAL INTRODUCTION

GENERAL INTRODUCTION

PART I: ANTIMICROBIAL STEWARDSHIP

Over the past decades, antimicrobial resistance has been steadily increasing. As a consequence, the number of antimicrobials to which bacteria are susceptible is diminishing and the demand for new antimicrobials grows. Unfortunately, the development of new antimicrobials is stagnating, with a 90% decrease in newly approved antibiotics over the last 30 years[1]. In addition, many of the new antimicrobials that have been approved over the past years share similar targets and no new antibiotic classes have been discovered[2]. One of the driving forces in the emergence and the selection of antimicrobial resistance is antibiotic use[3]. Therefore, an obvious way to counteract, or at least slow down, the rise of antimicrobial resistance is by using available antimicrobials optimally. This effort of optimizing antimicrobial use is called antimicrobial stewardship. The primary goal of antimicrobial stewardship is to optimize clinical outcomes while minimizing unintended consequences of antimicrobial use, including toxicity, the selection of pathogenic organisms, and the emergence of resistance[4]. A comprehensive systematic review and meta-analysis about the effectiveness of antimicrobial stewardship interventions identified many different kind of interventions: audit and feedback, education, reminders, structural changes in the working environment, and antibiotic restriction[5]. Interventions that included enablement (reminders, audit and feedback, review and recommend change) and restriction were associated with a stronger intervention effect. On average, antimicrobial stewardship interventions reduced unnecessary prescribing by 15% (95% confidence interval 14%-16%) and the duration by 1.95 days (95% confidence interval 2.22-1.67)[5]. Importantly, this reduction was not associated with an increase in mortality. Even though many antimicrobial stewardship intervention studies have been published, the evidence base for which stewardship interventions are most effective and their effect on clinical outcomes is still weak[6]. Many stewardship studies are of low methodological quality, with uncontrolled before-after studies being the most prevalent study design[6]. Therefore, there is a great need to improve the methodological quality of antimicrobial stewardship studies and for antimicrobial stewardship studies using appropriate research designs[7].

PART II: COMMUNITY-ACQUIRED PNEUMONIA

Of all the antibiotics that are prescribed in hospitals, the majority is for pneumonia (19.2%), followed by urinary tract infections (10.7%), and skin and soft tissue infections (9%)[8]. Consequently, pneumonia is an appealing target for antimicrobial stewardship interventions. Pneumonia can be classified according to where it was acquired, either in the community (community-acquired pneumonia, CAP) or in the hospital (hospital-acquired pneumonia, HAP). This classification is important for clinical practice because

their aetiology and management differ. Community-acquired pneumonia (CAP) is a lower respiratory tract infection with high morbidity and mortality, which predominately affects elderly patients[9]. The most common pathogens causing CAP are *Streptococcus pneumoniae* (15.9%) and *Haemophilus influenzae* (6.8%), but in the majority of patients no pathogen is identified (63,4%)[9]. Besides bacterial pathogens, viral pathogens may be identified in up to 23% of patients with CAP[10]. It is difficult to reliably determine the causative pathogen of CAP at the time of presentation. Therefore, empirical treatment is based on the concept that the most severely ill patients require immediate broad-spectrum treatment covering the great majority of possible pathogens while in non-severe patients there is room to start with narrow-spectrum antibiotics and wait for clinical response and or diagnostic test results. Hence, CAP is classified into mild, moderate-severe, and severe CAP based on one of the available scoring systems: the CURB-65 score, the pneumonia severity index (PSI) or the pragmatic classification[11]. For patients with moderate-severe CAP (admitted to a non-ICU ward; CURB-65 score of 2; or PSI-score 3-4) different empirical treatment options are recommended by international guidelines. The American and British guidelines recommend beta-lactam macrolide combination therapy or fluoroquinolone monotherapy[12,13], while Swedish, Danish and Dutch guidelines recommend narrow-spectrum beta-lactam monotherapy[11,14]. The rationale to recommend narrow-spectrum beta-lactam monotherapy is based on *Streptococcus pneumoniae* being the most common causative pathogen in CAP and the disease severity of moderate-severe CAP patients is low enough for a "wait-and-see" policy. Thus far, two randomised controlled trials (RCTs) have investigated whether empirical coverage for atypical pathogens (atypical because they can not be treated with the typical beta-lactam antibiotics: *Mycoplasma pneumoniae*, *Chlamydia pneumoniae*, and *Legionella pneumophila*) improves clinical outcomes[9,15]. One was a multicenter non-inferiority RCT comparing beta-lactam monotherapy with beta-lactam macrolide combination therapy which failed to demonstrate non-inferiority regarding the primary outcome of clinical stability (risk difference of 7,6%, 1 sided 90% CI of 13% overlapping the non-inferiority margin of 8%). The other was a multicenter cluster randomised cross-over trial comparing a strategy with preferred empirical treatment of beta-lactam monotherapy with beta-lactam macrolide combination therapy and fluoroquinolone monotherapy. In this study, beta-lactam monotherapy was non-inferior to the broader regimens regarding 90-day mortality (2 sided 90% CI not overlapping the non-inferiority margin of 3%). In addition, two randomised controlled trials have evaluated the efficacy of narrow-spectrum beta-lactams in moderate-severe CAP patients[16,17]. One RCT compared moxifloxacin to amoxicillin in patients with mild-to-moderate suspected pneumococcal CAP[16]. The clinical success rate was similar in the moxifloxacin group (173/200, 86.5%) compared to the amoxicillin group (171/208, 82.2%). The other RCT showed no difference in success rate for sparfloxacin (133/159,

83.6%) compared to amoxicillin (144/170, 84.7%) in patients with moderate-severe CAP[17]. Yet, in clinical practice the adherence to antibiotic recommendations in the CAP guideline are low (range 30.5%-62.9%, depending on severity classification used)[18]. Therefore, antimicrobial stewardship to improve the adherence to CAP treatment guidelines is urgently needed. In addition, it is currently unknown which moderate-severe CAP patients can be treated with narrow-spectrum antibiotics and which patients might benefit from broader treatment.

AIMS OF THIS THESIS

The aims of this thesis are to explore the methodology of studies evaluating antimicrobial stewardship interventions. In addition, we aim to investigate whether antimicrobial stewardship is effective in reducing broad-spectrum antibiotic use in patient with moderate-severe CAP without compromising patient outcomes. Lastly, we aim to explore optimal antibiotic treatment strategies in adult patients with moderate-severe CAP.

THESIS OUTLINE

In the first part of this thesis the focus will be on the methodology of antimicrobial stewardship. We start by describing the methodology that is currently used in antimicrobial stewardship in a comprehensive systematic literature review (**chapter 2**). In this chapter we describe various design quality features of antimicrobial stewardship studies with the goal to identify features which limit validity and translation of research findings into clinical practice. In addition, we identify factors associated with certain quality features. In **chapter 3** the results of the systematic review are used as a basis for a consensus document with recommendations for the optimal study design of antimicrobial stewardship intervention studies. This consensus document was developed by a multidisciplinary and multinational working group using a consensus procedure and consisted of a theoretical framework and specific recommendations to improve study design. In **chapter 4** a novel methodology called the Response Adjusted for Days of Antibiotic Risk (RADAR) is evaluated. RADAR was proposed as a possible new and more efficient outcome measure for antimicrobial stewardship non-inferiority trials. We investigated the efficiency and possible disadvantages of RADAR in a post-hoc analysis of a non-inferiority trial in which different antibiotic strategies were compared in moderate-severe CAP patients.

In the second part of this thesis the focus will be on implementing an antimicrobial stewardship intervention and the optimal antibiotic treatment of moderate-severe community-acquired pneumonia. First, in **chapter 5**, we evaluate the possible influence of confounding by indication in observational studies investigating the safety of de-escalating antibiotic therapy. We investigated the impact of confounding by indication by performing a simulation study embedded in a prospective cohort of CAP patients. In **chapter 6** we explore whether patients with Healthcare-associated pneumonia (HCAP), a specific subset of CAP patients that were in recent contact with healthcare, should be treated as CAP patients with narrow-spectrum beta-lactam monotherapy (amoxicillin) or with broader spectrum antibiotics. In **chapter 7** we investigate whether we can identify clinical variables that can predict increased treatment efficacy of either beta-lactam monotherapy, beta-lactam macrolide combination therapy, or fluoroquinolone based strategies. In **chapter 8** we present a narrative review on inappropriate antibiotic therapy for lower respiratory tract infections in elderly patients. We discuss possible antimicrobial stewardship interventions and novel rapid diagnostic techniques that may optimize antimicrobial use in these patients. Lastly, in **chapter 9** we present the results of a multicentre stepped-wedge cluster randomised antimicrobial stewardship intervention trial. In this trial we implement an evidence-based multifaceted intervention bundle based on education and audit and feedback with the goal of reducing broad-spectrum use and demonstrate that the reduction is safe for patients.

REFERENCES

1. Spellberg B, Bartlett J, Wunderink R, Gilbert DN. Novel approaches are needed to develop tomorrow's antibacterial therapies. *Am J Respir Crit Care Med*. 2015. doi:10.1164/rccm.201410-1894OE

2. Nielsen TB, Brass EP, Gilbert DN, Bartlett JG, Spellberg B. Sustainable discovery and development of antibiotics - Is a nonprofit approach the future? *N Engl J Med*. 2019. doi:10.1056/NEJMp1905589

3. Costelloe C, Metcalfe C, Lovering A, Mant D, Hay AD. Effect of antibiotic prescribing in primary care on antimicrobial resistance in individual patients: Systematic review and meta-analysis. *BMJ*. 2010. doi:10.1136/bmj.c2096

4. Dyar OJ, Huttner B, Schouten J, Pulcini C. What is antimicrobial stewardship? *Clin Microbiol Infect*. 2017. doi:10.1016/j.cmi.2017.08.026

5. Davey P, Marwick CA, Scott CL, et al. Interventions to improve antibiotic prescribing practices for hospital inpatients. *Cochrane Database Syst Rev*. 2017;2017(2). doi:10.1002/14651858.CD003543.pub4

6. Schuts EC, Hulscher MEJL, Mouton JW, et al. Current evidence on hospital antimicrobial stewardship objectives: A systematic review and meta-analysis. *Lancet Infect Dis*. 2016;16(7):847-856. doi:10.1016/S1473-3099(16)00065-7

7. Hulscher MEJL, Prins JM. Antibiotic stewardship: does it work in hospital practice? A review of the evidence base. *Clin Microbiol Infect*. 2017. doi:10.1016/j.cmi.2017.07.017

8. Versporten A, Zarb P, Caniaux I, et al. Antimicrobial consumption and resistance in adult hospital inpatients in 53 countries: results of an internet-based global point prevalence survey. *Lancet Glob Heal*. 2018. doi:10.1016/S2214-109X(18)30186-4

9. Postma DF, van Werkhoven CH, van Elden LJR, et al. Antibiotic Treatment Strategies for Community-Acquired Pneumonia in Adults. *N Engl J Med*. 2015;372(14):1312-1323. doi:10.1056/NEJMoa1406330

10. Jain S, Self WH, Wunderink RG, et al. Community-Acquired Pneumonia Requiring Hospitalization among U.S. Adults. *N Engl J Med*. 2015;373(5):415-427. doi:10.1056/NEJMoa1500245

11. Wiersinga WJ, Bonten MJ, Boersma WG, et al. Management of community-acquired pneumonia in adults: 2016 guideline update from the dutch working party on antibiotic policy (SWAB) and dutch association of chest physicians (NVALT). *Neth J Med*. 2018.

12. Metlay JP, Waterer GW, Long AC, et al. Diagnosis and Treatment of Adults with Community-acquired Pneumonia. An Official Clinical Practice Guideline of the American Thoracic Society and Infectious Diseases Society of America. *Am J Respir Crit Care Med*. 2019. doi:10.1164/rccm.201908-1581st

13. Lim WS, Baudouin S V, George RC, et al. BTS guidelines for the management of community acquired pneumonia in adults: update 2009. *Thorax*. 2009;64(Suppl 3):iii1-iii55. doi:10.1136/thx.2009.121434

14. Athlin S, Lidman C, Lundqvist A, et al. Management of community-acquired pneumonia in immunocompetent adults: updated Swedish guidelines 2017. *Infect Dis (Auckl)*. 2018. doi:10.1080/23744235.2017.1399316

15. Garin N, Genné D, Carballo S, et al. -Lactam Monotherapy vs -Lactam-Macrolide Combination Treatment in Moderately Severe Community-Acquired Pneumonia: A Randomized Noninferiority Trial. *JAMA Intern Med*. 2014;174(12):1894-1901. doi:10.1001/jamainternmed.2014.4887

16. Petitpretz P, Arvis P, Marel M, et al. Oral moxifloxacin vs high-dosage amoxicillin in the treatment of mild-to-moderate, community-acquired, suspected pneumococcal pneumonia in adults. *Chest*. 2001;119(1):185-195. doi:10.1378/chest.119.1.185

17. Aubier M, Verster R, Regamey C, Geslin P, Vercken JB. Once-daily sparfloxacin versus high-dosage amoxicillin in the treatment of community-acquired, suspected pneumococcal pneumonia in adults. *Clin Infect Dis*. 1312;26(6):1312-1323.

18. Huijts SM, Van Werkhoven CH, Boersma WG, et al. Guideline adherence for empirical treatment of pneumonia and patient outcome treating pneumonia in the netherlands. *Neth J Med*. 2013.

PART

2

THE QUALITY OF STUDIES EVALUATING ANTIMICROBIAL STEWARDSHIP INTERVENTIONS: A SYSTEMATIC REVIEW

Valentijn A. Schweitzer, Inger van Heijl, Cornelis H. van Werkhoven, Jasmin Islam, Kelly D. Hendriks-Spoor, Julia Bielicki, Marc J.M. Bonten, A. Sarah Walker, Martin J. Llewelyn, Consensus on Antimicrobial Stewardship Evaluations (CASE) study group

ABSTRACT

BACKGROUND

Antimicrobial stewardship aims to optimise antibiotic use and minimise selection of antimicrobial resistance. The methodological quality of published studies in this field is unknown.

OBJECTIVES

Our objective was to perform a comprehensive systematic review of antimicrobial stewardship research design and identify features which limit validity and translation of research findings into clinical practice.

DATA SOURCES

The following online database was searched: PubMed.

STUDY ELIGIBILITY CRITERIA

Studies published between January 1950 and January 2017, evaluating any antimicrobial stewardship intervention in the community or hospital setting, without restriction on study design or outcome.

METHODS

We extracted data on pre-specified design quality features and factors that may influence design choices including: (1) clinical setting, (2) age group studied, (3) when the study was conducted, (4) geographical region and (5) financial support received.

RESULTS

The initial search yielded 17,382 articles; 1,008 were selected for full-text screening, of which 825 were included. Most studies (675/825, 82%) were non-experimental and 104 (15%) used interrupted time series analysis, 41 (6%) used external controls and 19 (3%) used both. Studies in the community setting fulfilled a median of 5/10 quality features (IQR 3-7) and 3 (IQR 2-4) in the hospital setting. Community setting studies (25%, 205/825) were significantly more likely to use randomisation (OR 5.9 (95%CI 3.8-9.2)), external controls (OR 5.6 (95%CI 3.6-8.5)) and multiple centres (OR 10.5 (95%CI 7.1-15.7)). From all studies, only 48% (398/825) reported clinical and 23% (190/825) reported microbiological outcomes. Quality did not improve over time.

CONCLUSIONS

Overall quality of antimicrobial stewardship studies is low and has not improved over time. Most studies do not report clinical and microbiological outcome data. Studies conducted in the community setting were associated with better quality. These limitations should inform the design of future stewardship evaluations so that a robust evidence base can be built to guide clinical practice.

INTRODUCTION

Antimicrobial resistance (AMR) is increasing globally and is a substantial threat to human health[1]. There is a clear relationship between antibiotic exposure and AMR both in populations[2] and individual patients[3]. An estimated 30% of human antibiotic use may be unnecessary and healthcare systems around the world are aiming to achieve substantial reductions in unnecessary antibiotic prescribing. The term 'antimicrobial stewardship' is used to describe use of antibiotics which balances the need for effective individual treatment against the longer-term, societal impact of antibiotic use on antibiotic resistance[4]. Interventions to improve antimicrobial stewardship are usually multifaceted and include periodic or individual patient audit and feedback, decision support, education (educational meetings, educational materials), and antimicrobial formulary restriction[5].

Recognition of the threat posed by AMR and the need to optimise antibiotic prescribing has driven an exponential increase in the publication of studies evaluating antimicrobial stewardship interventions over the last 20 years[6]. Previous systematic reviews have synthesised this evidence with the aim of making recommendations for practice[5,7-12]. These have, appropriately, considered studies with the lowest possible risk of bias but have excluded >50% of published studies in which methodological quality falls below Effective Practice and Organisation of Care (EPOC) criteria[13]. Because the minority of studies are of sufficient quality, many areas of practice rely on a weak evidence base[8]; conducting studies which do not inform practice is a waste of time and valuable resources[5,12].

Journals are beginning to report the minimum standards for antimicrobial stewardship studies to be published[14]; however, there remains a need for consensus on how to design, analyse and report studies evaluating interventions to improve antibiotic prescribing. This would optimise use of valuable resources and strengthen the evidence base in this field. Currently, no overview exists of how antimicrobial stewardship evaluations are designed. We conducted a systematic review of antimicrobial stewardship evaluations with the aim of identifying areas in stewardship evaluation most in need of improvement, to increase validity and translation of research findings into clinical practice.

METHODS

SEARCH STRATEGY

We searched PubMed for studies evaluating antimicrobial stewardship interventions between January 1950 and January 2017. The search strategy (appendix 1) was designed to be as broad as possible. Inclusion criteria were: any study evaluating an antimicrobial stewardship intervention, without restriction on the type of intervention studied and what outcomes were evaluated. Studies were excluded if they were (1) not in English, (2) case-reports, (3) focused mainly on HIV or (4) narrative or systematic reviews. All studies were screened by one author (VAS, CHvW, JI, KH or IvH). In case of duplicate publications, only the original article was included. A random selection of 700 (~4% of total) studies were assessed by a second author (VAS, CHvW, JI, KH or IvH). Uncertainties about the inclusion of studies was resolved by discussion.

DATA EXTRACTION

Studies fulfilling the inclusion criteria were identified by screening on title and abstract. All selected studies then underwent full-text evaluation by one author (VAS, CHvW, JI, KH or IvH) against the inclusion and exclusion criteria and data was subsequently extracted using a standard data extraction template (appendix 2). Data were extracted on study characteristics (i.e. title, authors, year of publication), design quality features, and factors possibly associated with methodological quality. If no funding was reported it was assumed that studies received no financial support. Authors were not contacted in case data were missing or incomplete. A random selection of 10% of the studies were extracted by a second author. We followed the PRISMA criteria for the reporting of systematic reviews (appendix 3)[15].

SELECTION OF QUALITY FEATURES, AND FACTORS ASSOCIATED WITH QUALITY

In February 2017 we established an international Consensus Working Group funded by the Joint Programming Initiative on Antimicrobial Resistance (JPIAMRWG-010) to develop recommendations on the design, analysis and reporting of antimicrobial stewardship evaluations. The working group coordinators (VAS, CHvW, ML, ASW, MB) invited members to the consensus group based on their expertise on antimicrobial stewardship and/or trial methodology, ensuring that all key clinical areas (primary care, secondary care, intensive care medicine and paediatrics) were represented. The consensus group agreed that a review to identify areas in stewardship evaluation most in need of improvement to increase validity and translation of research findings into clinical practice was required. For this purpose, existing quality scores were not applicable because these focus solely on methodological quality. The group selected quality features and factors likely to underlie design quality and features based on plenary group discussion during the consensus meeting. Ten quality features were selected for assessment: (1) randomized research design used, (2) external control group assessed, (3) multiple centres used, (4) sustainability of the intervention sufficiently assessed (≥12 months), (5) sample size calculation reported, (6) prospective data collection, (7) correction for confounding factors, (8) primary outcome defined and reported, (9) clinical outcome reported, and (10) microbiological outcome reported.. Selected factors likely to underlie design quality features were (a) the clinical setting (community versus hospital), (b) age group studied (studies including children versus adults, or both), (c) year when study was conducted (newer versus older studies, categorised at approximate quintiles: 1977-2004, 2005-2010, 2011-2013, 2014-2015, 2016-2017) (d) geographical region, and (e) financial support. The quality features and corresponding categorizations are shown in table S1.

STATISTICAL ANALYSIS

Descriptive statistics were used to describe the quality features of included studies. Differences stratified by subgroup were displayed using spider graphs (Microsoft Excel, version 2010). To assess the independent relationship between factors and quality features we performed multivariable logistic regression models with backward stepwise selection (exit p-value >0.10), presenting odds ratios and 95% confidence intervals in a heat map displaying the strength of the association (Microsoft Excel, version 2010). Analyses were performed using the Statistical Package for the Social Sciences for Windows (Version SPSS 21.0.0.0).

RESULTS

The initial search yielded 17,382 articles. After title and abstract screening, 1,008 articles were selected for full-text screening (Figure 1). Of these, 183 were excluded leaving 825 articles for full assessment (appendix 4). Among 700 randomly selected articles that were screened by a second author, 640/700 (91%) were excluded by both authors, 23/700 (3.3%) were selected for inclusion by both authors, 13/700 (1.8%) were selected for inclusion by only the first author, and 24/700 (3.4%) were selected for inclusion by only the second author, resulting in a percentage agreement after title/abstract screening of 95%, with a moderate interobserver agreement (Cohen's kappa: 0.53). After discussion of the full text articles, consensus about inclusion was reached in 99.5% (696/700) of the selected articles. From the 3.4% (24/700) studies that were newly identified by a second author, 7/700 (1.0%) were considered correct inclusions after discussion. Therefore, the low proportion of missed papers justified not screening in duplicate. Among the 83 articles selected for double data-extraction by a second

author, the percentage agreement per variable ranged from 91%-100%, with all the quality features showing an agreement of >95%.
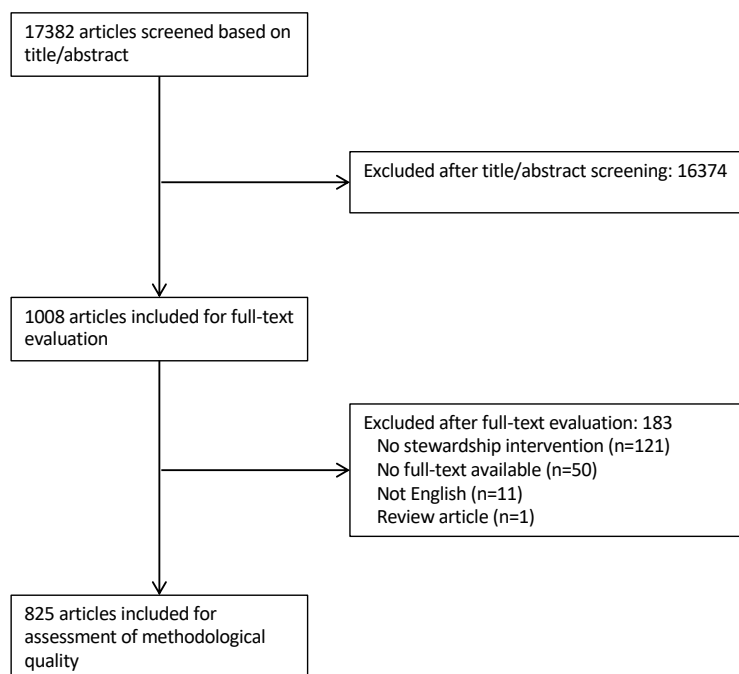

Figure 1. Flowchart of the studies included in the systematic literature review

CHARACTERISTICS OF INCLUDED STUDIES

A total of 825 studies were included, In the community setting, most studies were multicentre (72%, 148/205) and the commonest study designs were before-after studies without an interrupted time-series (ITS) analysis (23%, 48/205), randomised controlled trials (14%, 28/205), and parallel cluster randomised trials (15%, 30/205)(Table 1). Among the included studies in the hospital setting, most were single centre (84%, 519/620) and the commonest study designs were before-after studies without an ITS analysis (59%, 365/620), before-after studies with an ITS analysis (13%, 82/620), and cohort studies without a control group (12%, 75/620). Among the before-after studies without an ITS analysis, the majority were single centre (86%, 352/411), and retrospective (58%, 239/411). In the 86 (10%) cluster randomised studies a median of 28 clusters were randomised, with 57% (49/86) randomising >20 clusters, 28% (24/86) randomising <15 clusters, 21% (18/86) randomising <10 clusters, and 9% (8/86) randomising <5 clusters. In the hospital setting, 13% (2/13) randomised >20 clusters, while in the community setting 65% (64/71) randomised >20 clusters. A minority of studies were conducted in children (12%,

101/825). Both in the community and the hospital setting, most studies did not target a specific disease or syndrome (31%, 63/205, 55% 324/620, respectively) or specific antibiotic class (86%, 177/205, 69% 427/620, respectively). The majority of interventions were bundles (57%, 470/825). Commonest interventions in the community setting included education (73%, 149/205), audit with periodic feedback (29%, 59/205), and clinical decision support 15% (31/205). In the hospital setting, commonest interventions included education 42% (260/620), audit and feedback on an individual patient level (40%, 245/620), restriction (18%, 113/620), and clinical decision support (18%, 112/620). Both in the community and hospital setting, virtually all included studies reported process measure outcomes (99%, 818/825) (Table 2). Both in the community setting and the hospital the most commonly reported process measures included the proportion of patients treated with antibiotics (59%, 121/205, 21%, 131/620), costs/cost-effectiveness (18%, 36/205, 32%, 200/620, respectively), appropriateness (17%, 34/205, 29%, 178/620, respectively), and defined daily doses (17%, 34/205, 25%, 156/620, respectively). In the community setting, commonest reported clinical outcomes were revisits (11%, 22/205), clinical cure (6%,12/205), and infection (5%, 10/205), while in the hospital setting these were mortality (32%, 302/620), length of stay (32%, 201/620), and hospital readmissions (12%, 76/620).

Table 1. Characteristics of included studies stratified by studies performed in the community and the hospital setting

Study characteristics	Community (n=205) n (%)	Hospital (n=620) n (%)
Number of patients included (median, IQR)	1255 (278-11230)	423 (186-1398)
Number of centres involved (median, IQR)	27 (8-90)	1 (1-1)
Age		
Adults	64 (31)	285 (46)
Children	33 (16)	68 (11)
Both	108 (53)	267 (43)
Specific disease targeted		
No specific disease targeted	63 (31)	342 (55)
Upper respiratory tract infections	106 (52)	25 (4)
Lower respiratory tract infections	57 (28)	70 (11)
Bacteraemia	0 (0)	42 (7)
Urinary tract infections	18 (9)	27 (4)
Prophylaxis	2 (1)	39 (6)
Sepsis	0 (0)	17 (3)
Skin and soft tissue infections	1 (1)	8 (1)
Abdominal infections	3 (1)	10 (2)
Other	16 (8)	62 (10)
Antibiotic class targeted		
No specific antibiotic class targeted	177 (86)	427 (69)
Cephalosporins	4 (2)	66 (11)
Fluoroquinolones	10 (5)	54 (9)
Carbapenems	0 (0)	48 (8)
Vancomycin	0 (0)	31 (5)
Aminoglycosides	0 (0)	28 (5)
Penicillins	10 (5)	15 (2)

Table 1 continued.

Study characteristics	Community (n=205) n (%)	Hospital (n=620) n (%)
Antibiotic class targeted		
Macrolides	8 (4)	5 (1)
Other	17 (8)	89 (14)
Antimicrobial stewardship interventions		
Education	149 (73)	260 (42)
Audit and feedback − individual patient	12 (6)	245 (40)
Audit and feedback − periodic	59 (29)	82 (13)
Restriction	13 (6)	113 (18)
Clinical decision support	31 (15)	112 (18)
Rapid diagnostic testing	24 (12)	68 (11)
Therapeutic drug monitoring	0 (0)	15 (2)
Guideline implementation	20 (10)	95 (15)
Delayed prescribing	9 (4)	0 (0)
Other	42 (20)	78 (13)
Research designs		
Before-after study	48 (23)	365 (59)
Before-after study (ITS*)	22 (11)	82 (13)
Cohort without control group	6 (3)	75 (12)
Controlled before-after study	22 (11)	19 (3)
Randomised controlled trial (RCT)	28 (14)	40 (6)
Parallel cluster randomised trial	30 (15)	6 (1)
Parallel cluster randomised trial with baseline period	26 (13)	2 (1)
Controlled before-after study (ITS)	7 (3)	12 (2)
Non-randomised parallel cluster study	5 (2)	8 (1)
Factorial cluster randomised trial	7 (3)	0 (0)
Cluster randomised cross-over trial	2 (1)	3 (1)
Stepped wedge cluster randomised trial	0 (0)	1 (1)
Non-randomised cluster cross-over study	0 (0)	3 (1)
Factorial randomised controlled trial (RCT)	1 (1)	2 (1)
Non-randomised stepped wedge study	0 (0)	2 (1)
Adaptive RCT	1 (1)	0 (0)

IQR: interquartile range, *ITS: interrupted time series

QUALITY FEATURES

The percentage of studies including each quality feature is shown in Table 3. Studies in the community setting fulfilled a median of 5 quality features (IQR 3-7), while studies in the hospital setting fulfilled 3 (IQR 2-4). None fulfilled all 10 quality features. In the community setting 2% (4/205) fulfilled 9, 16% (33/205) fulfilled at least 8, and 35% (72/205) fulfilled at least 7 quality indicators. In the hospital setting 1% (4/620) fulfilled 9, 3% (19/620) fulfilled 8, and 6% (37/620) fulfilled 7 quality indicators. Of note, there were substantial differences between studies which did and did not use randomised designs to the extent to which other quality features were present. Among the 150 randomised studies, all used an external control group, 71% (107/150) included multiple centres, 64% (97/150) reported a sample size calculation and 96% (144/150) collected data prospectively. In contrast, among the non-randomised studies, 12% (78/675) used an external control group, 21% (142/675) included multiple centres, 11% (76/675) reported a sample size calculation and 43% (288/675) collected data prospectively.

Table 2. Outcomes reported in the included antimicrobial stewardship studies stratified by studies performed in the community and the hospital setting

Process measure outcomes	Community (n=205) n (%)	Hospital (n=620) n (%)
Costs/cost-effectiveness	36 (18)	200 (32)
Appropriateness	34 (17)	178 (29)
Defined daily doses (DDD)	34 (17)	156 (25)
Proportion treated with antibiotics	121 (59)	131 (21)
Recommendation acceptance	6 (3)	114 (18)
Guideline adherence	27 (13)	100 (16)
Duration of treatment	5 (2)	93 (15)
Days on therapy (DOT)	6 (3)	62 (10)
Time to appropriate therapy	3 (1)	71 (11)
Antibiotic knowledge	17 (8)	14 (2)
None	2 (1)	5 (1)
Other	58 (28)	166 (27)
Clinical outcome measures		
None	144 (70)	283 (46)
Mortality	9 (4)	203 (33)
Length of stay	5 (2)	201 (32)
Infection	10 (5)	75 (12)
Hospital readmission	0 (0)	76 (12)
Adverse effects	5 (2)	52 (8)
Revisits	22 (11)	0 (0)
Clinical cure	12 (6)	27 (4)
Intensive care unit admission	0 (0)	24 (4)
Hospital admission	9 (4)	0 (0)
Time to clinical stability	3 (1)	6 (1)
Other	16 (8)	49 (8)
Microbiological outcome measures		
None	188 (92)	447 (72)
Colonization/infection resistant pathogens	17 (8)	146 (24)
Clostridium difficile infections	5 (2)	62 (10)
Other	2 (1)	12 (2)

Table 3. Design quality features of the included studies stratified by studies performed in the community and the hospital setting

Quality feature	Community (n=205) n (%)	Hospital (n=620) n (%)
Randomised research design	95 (46)	55 (9)
External control group	129 (63)	99 (16)
Multicentre	148 (72)	101 (16)
Sample size calculation reported	77 (38)	96 (15)
Prospective data collection	144 (70)	288 (46)
Correction for confounding factors	113 (55)	157 (25)
Primary outcome defined	116 (57)	272 (44)
Clinical outcome reported	61 (30)	337 (54)
Microbiological outcome reported	17 (8)	173 (28)
Sustainability assessed (≥12 months)	115 (56)	347 (56)

FACTORS ASSOCIATED WITH DESIGN QUALITY

Design quality was considerably better in almost all quality aspects of community versus hospital setting studies (Figure S1A), with more use of randomised designs (46% vs. 9%),

external controls (63% vs. 16%), sample size calculations (38% vs. 15%), prospective data collection (70% vs. 46%), correction for confounding (55% vs. 25%), use of a defined primary outcome (57% vs. 44%), and involvement of multiple centres (72% vs. 16%). However, community setting studies less often reported clinical (30% vs. 54%) and microbiological (8% vs. 28%) outcomes. Community setting remained significantly associated with all these factors in multivariable models (Table 4).

Table 4. Results of stepwise backward selection of multivariable model containing all the factors with the different design quality indicators as outcome. The colours indicate either a strong negative association (OR<1.0) in red, or a strong positive association (OR>1.0) in green. Numbers indicate odds ratio's with 95% confidence intervals

	Factors	Model 1 Randomised design	Model 2 External control	Model 3 Sample size calculation	Model 4 Prospective data	Model 5 Confounding correction	Model 6 Primary outcome	Model 7 Clinical outcome	Model 8 Microbiological outcome	Model 9 Multicentre	Model 10 Sustainability assessed
Clinical setting	Community (n=205)	5.9 (3.8 - 9.2)	5.6 (3.6 - 8.5)	2.9 (1.9 - 4.5)	1.7 (1.2 - 2.6)	3.0 (2.1 - 4.3)	1.4 (1.0 - 2.0)	0.3 (0.2 - 0.5)	0.2 (0.1 - 0.4)	10.5 (7.1 - 15.7)	
Financial support	Yes (n=385)	4.7 (2.9 - 7.5)	4.1 (2.6 - 6.3)	3.1 (2.1 - 4.7)	2.2 (1.6 - 3.0)	2.5 (1.8 - 3.4)	1.8 (1.4 - 2.5)	1.9 (1.4 - 2.6)		3.0 (2.0 - 4.3)	
Age group studied**	Children (n=101)	1.1 (0.6 - 2.2)	0.8 (0.4 - 1.6)	0.8 (0.5 - 1.4)	1.1 (0.7 - 1.9)		1.0 (0.6 - 1.6)	0.6 (0.4 - 1.0)			1.4 (0.9 - 2.2)
	Both (n=375)	0.6 (0.4 - 1.0)	0.6 (0.4 - 0.9)	0.4 (0.2 - 0.6)	0.5 (0.4 - 0.7)		0.5 (0.4 - 0.7)	0.4 (0.3 - 0.5)			1.6 (1.2 - 2.2)
Year study conducted***	2005-2010 (n=160)	1.0 (0.5 - 1.9)	1.1 (0.6 - 1.9)	2.5 (1.3 - 4.8)	0.4 (0.2 - 0.6)	0.6 (0.4 - 1.0)		0.9 (0.6 - 1.5)	2.1 (1.2 - 3.7)		1.1 (0.7 - 1.8)
	2011-2013 (n=156)	0.5 (0.3 - 1.1)	0.5 (0.2 - 0.9)	1.8 (0.9 - 3.5)	0.2 (0.1 - 0.3)	0.7 (0.4 - 1.2)		1.0 (0.6 - 1.6)	1.7 (1.0 - 3.0)		1.6 (1.0 - 2.5)
	2014-2015 (n=173)	0.4 (0.2 - 0.8)	0.5 (0.2 - 0.9)	2.5 (1.3 - 4.9)	0.1 (0.1 - 0.2)	1.1 (0.7 - 1.8)		1.3 (0.8 - 2.1)	1.2 (0.7 - 2.1)		1.8 (1.1 - 2.8)
	2016-2017 (n=176)	0.8 (0.4 - 1.5)	0.9 (0.5 - 1.7)	2.5 (1.3 - 4.9)	0.1 (0.1 - 0.2)	1.2 (0.7 - 1.9)		1.7 (1.0 - 2.7)	1.4 (0.8 - 2.4)		1.2 (0.8 - 1.8)
Geographical region****	Europe (n=260)	2.5 (1.5 - 4.2)	2.2 (1.4 - 3.5)	1.6 (1.1 - 2.5)	2.8 (1.9 - 4.1)		1.1 (0.8 - 1.5)	1.1 (0.8 - 1.6)	0.9 (0.6 - 1.4)	2.0 (1.3 - 3.1)	
	Asia (n=128)	2.6 (1.3 - 5.0)	2.4 (1.3 - 4.4)	1.1 (0.6 - 2.0)	1.7 (1.1 - 2.7)		0.4 (0.3 - 0.7)	1.8 (1.1 - 2.8)	1.2 (0.8 - 2.0)	1.0 (0.6 - 1.8)	
	Oceania (n=36)	1.6 (0.6 - 4.2)	1.5 (0.6 - 3.7)	0.5 (0.1 - 1.4)	1.8 (0.8 - 3.9)		0.5 (0.2 - 1.0)	0.5 (0.2 - 1.0)	0.2 (0.1 - 0.8)	0.9 (0.4 - 2.2)	
	Africa (n=29)	4.0 (1.1 - 14.5)	3.3 (0.9 - 11.8)	1.8 (0.5 - 5.8)	2.0 (0.6 - 6.2)		0.9 (0.3 - 2.5)	1.8 (0.6 - 5.3)	0.2 (0.1 - 1.7)	7.2 (2.0 - 25.1)	
	South America (n=17)	1.8 (0.4 - 7.4)	2.3 (0.6 - 8.4)	2.0 (0.6 - 6.5)	2.2 (0.8 - 6.4)		1.1 (0.4 - 2.9)	1.1 (0.4 - 3.0)	1.3 (0.4 - 3.8)		

*Reference category: studies performed in the hospital setting (n=620), **Reference category: studies performed in adults (n=349), ***Reference category: studies performed before 2005 (n=160),

****Reference category: studies performed in North America (n=368)

Strong negative association   Weak association   Strong positive association

From the studies that reported financial support, 20% (53/264) were industry funded, and 84% (221/264) were publicly funded. Studies with financial support were of higher methodological quality than studies without (Figure S1B), as they more frequently used randomised designs (31% vs. 7%), external controls (34% vs. 8%), sample size calculations (33% vs. 11%), prospective data collection (63% vs 43%), correction for confounding (46% vs. 21%), a defined primary outcome (56% vs. 39%), and involved multiple centres (46% vs. 16%). Financial support remained significantly associated with these factors in multivariable models (Table 4). In addition, financial support increased the frequency of reporting clinical outcomes in multivariable models. There was little change in design quality over time, other than a decrease in the proportion of studies with prospective data collection (77% in 1977-2004, 67% in 2005-2010, 42% in 2011-2013, 40% in 2014-2015, 39% in 2016-2017) and an increase in studies reporting a clinical outcome (39% in 1977-2004, 44% in 2005-2010, 44% in 2011-2013, 53% in 2014-2015, 59% in 2016-

2017) (Figure S1D). These outcomes were significantly associated with calendar time in multivariable models (Table 4), and sample size calculations were independently reported more in later studies. The decrease in studies with prospective data collection is most prominent in studies in the hospital setting (Table S2). There were no large differences between studies performed in children versus adults (Figure S1C). Geographical region was independently associated with randomised designs, using an external control, prospective data collection, performing sample size calculations, reporting a primary, clinical or microbiological outcome, and being multicentre (Table 4).

DISCUSSION

In previous systematic reviews of antimicrobial stewardship evaluations, many studies have been excluded due to not fulfilling minimal methodological quality criteria. We have undertaken the first comprehensive systematic review focusing on describing quality, rather than excluding studies based on quality, to facilitate formulating recommendations for improvement. In addition, we evaluated quality features required for validity and translation into practice instead of focusing solely on methodological quality. Our systematic review revealed that the design quality of antimicrobial stewardship evaluations is low, with only a minority of studies reporting clinical and microbiological outcome data. Design quality is considerably better in studies performed in the community setting.

We find published evaluations provide a striking lack of evidence for the clinical and microbiological impacts of antimicrobial stewardship interventions. The majority of studies focus exclusively on process measures. While it is clearly essential to establish whether an intervention is effective in changing antibiotic use, reporting clinical outcomes is crucial to assess the safety of antimicrobial stewardship interventions [16,17]. The clinical outcomes reported often utilise routinely collected data, which may explain the differences between the community and hospital setting. In particular, in the hospital setting, commonly used clinical outcomes are mortality and length of hospital stay[18,19]. As indicated by the observed time trends, these outcomes are being used with increasing frequency, probably because extraction of relevant data from electronic health records is becoming more feasible. Such routinely available data are not the most sensitive and patient-relevant outcomes. In the hospital setting, markers of early treatment response such as clinical stability may be preferable. In the community setting, repeat consultations, relapse of infection, and hospital admissions may be more relevant; yet data on these outcomes are not routinely collected[20–22].

Very few stewardship studies report microbiological outcomes. This is surprising given that reducing antimicrobial resistance is the ultimate goal of antimicrobial stewardship. However, this is consistent with a meta-analysis on the effect of stewardship interventions on infection and colonisation with antibiotic-resistant bacteria and *Clostridium difficile* infections that showed the literature on this topic is sparse and dominated by low quality research[23]. Some authorities have called for stewardship evaluations to routinely include consideration of the impact of stewardship on resistance[24] but studies generally lack power to determine this. Relationships between antimicrobial exposure and resistance may be more efficiently established through specific mechanistic studies rather than within stewardship evaluations.

Our analysis demonstrates that factors that would be expected to effect study design do, whiles others do not. The contrast between the community and hospital setting is striking in terms of the greater use of multicentre, randomised controlled designs. One explanation could be that clusters required for cluster randomisation are more readily available in the community setting. In contrast, clinical outcomes and microbiological data are less readily available in the community setting. Retrospective study designs are therefore less feasible in the community setting.

Financial support was associated with better design quality. In addition to the costs inherent to conducting multicentre, prospective studies with longer follow-up, the process of securing funding may drives careful consideration of study validity. Less than half of the stewardship studies reviewed reported external funding. However, our finding of an association between external funding and improved design quality underscore the necessity of external funding to support appropriate implementation and robust evaluation of antimicrobial stewardship programmes[25].

Our results show that there is no improvement of design quality over time, which is in contrast to previous reports[5,23]. This may be explained by our evaluation and inclusion of all studies without a pre-selection on study design, while previous reviews only included adequate studies with interpretable data[5].

In keeping with previous reports we find that the majority of antimicrobial stewardship studies used non-randomised research designs, with before-after studies being the most prevalent. This quasi-experimental research design is commonly used for quality improvement projects. However, such studies are at risk from multiple forms of bias[16,24,26] and the Effective Practice and Organisation of Care (EPOC) criteria strongly discourages the inclusion of before-after studies without an ITS analysis in systematic reviews[13]. Incorporation of properly conducted ITS analysis into before-after studies

has the potential to robustly control for time-dependent bias[27] but only 20% (104/515) of before-after studies we identified used ITS analysis. Moreover, it has been shown that contemporary ITS analyses are often performed with an insufficient number of data points[11].

This systematic review has several strengths. First, our comprehensive search strategy gives a unique overview of the quality of studies evaluating antimicrobials stewardship interventions. Second, we used the PRISMA reporting guide for systematic reviews[15]. Third, the quality indicators and candidate factors were selected in a consensus procedure as part of a Joint Programming Initiative on Antimicrobial Resistance (JPIAMR) funded consensus group, which was selected to both include experts on the field of antimicrobial stewardship and trial methodology. Finally, this is the first comprehensive systematic review to determine the extent to which published antimicrobial stewardship evaluations include quality features required for validity and translation into practice .

The limitations of our review were firstly, we only searched PubMed and excluded non-English studies, which makes it possible that antimicrobial stewardship studies indexed elsewhere and non-English studies were missed. However, if we compare the studies identified by our searching strategy with the largest community and hospital stewardship systematic reviews, only 6 studies were missed due to being indexed elsewhere and 11 due to being non-English. Therefore, this is likely to have had a minimal impact on the total results[5,12]. Secondly, the screening, inclusion, and data extraction was performed by only one investigator that could have resulted in studies being missed, wrongly included or misclassification of the extracted data. To estimate the amount of studies that might be missed, a proportion of the studies were screened and data was extracted by a second author. In this second round, we showed that the percentage of agreement was high, with a moderate interobserver agreement. Assuming that every disagreement in included studies would have inadvertently excluded a study (1% of 700 studies reviewed twice), we may have missed a maximum of 104 inclusions in the other studies not screened twice. In a systematic review with meta-analysis the consequence of missing or wrongly including a single study could have a large impact on the pooled effect estimate. However, as we did not focus on the outcome of individual studies but rather on the design quality of many studies, given the large number of studies included, it is unlikely that the missed studies would have changed the conclusions. And lastly, the quality indicator definition of a primary outcome was only based on what was described in the manuscript. It is possible that a primary outcome was defined retrospectively based on the observed data in a proportion of the studies.

Concerns about the methodological quality of antimicrobial stewardship studies have been raised before, with several publications making recommendations to improve their scientific methods[16,24,26]. However, have found no improvement in methodological quality over time except for more frequent inclusion of sample size calculation and clinical outcomes. Therefore, there is still a need for clear recommendations to improve antimicrobial stewardship design quality. Recommendations for improvement should especially consider: (1) emphasizing the importance of choosing appropriate clinical and microbiological outcomes, (2) focusing on robust methods to evaluate stewardship interventions in the hospital setting. Implementing these recommendations in future antimicrobial stewardship studies will help in the optimal use of resources to determine which stewardship interventions are most effective to change clinical practice. Building on the work from the systematic review we established a working group of expert investigators in this field. This systematic review identifies the limitations in design features that are most important for validity and translation into clinical practice. The results will be used to formulate recommendations in a white paper that will support investigators with key design decisions, support funders assessing proposals for stewardship studies and enhance the quality and impact of research in this crucial area.

SUPPORT STATEMENT

ACKNOWLEDGMENTS

CONFLICT OF INTEREST
The authors declare no conflict of interest.

REFERENCES

1. O'Neill J. Review on Antimicrobial Resistance. Tackling Drug-Resistant Infect Glob Final Rep Recomm 2016. https://amr-review.org/sites/default/files/160518_Final paper_with cover.pdf.

2. Goossens H, Ferech M, Vander Stichele R, Elseviers M, ESAC Project Group H, Sprenger M, et al. Outpatient antibiotic use in Europe and association with resistance: a cross-national database study. Lancet (London, England) 1998;365:579–87. doi:10.1016/S0140-6736(05)17907-0.

3. Costelloe C, Metcalfe C, Lovering A, Mant D, Hay AD. Effect of antibiotic prescribing in primary care on antimicrobial resistance in individual patients: systematic review and meta-analysis. Bmj 2010;340:c2096. doi:10.1136/bmj.c2096.

4. Dyar OJ, Huttner B, Schouten J, Pulcini C, ESGAP (ESCMID Study Group for Antimicrobial stewardshiP). What is antimicrobial stewardship? Clin Microbiol Infect 2017;0. doi:10.1016/j.cmi.2017.08.026.

5. Davey P, Marwick CA, Scott CL, Charani E, Mcneil K, Brown E, et al. Interventions to improve antibiotic prescribing practices for hospital inpatients. Cochrane Database Syst Rev 2017;2017. doi:10.1002/14651858. CD003543.pub4.

6. Molina J, Cisneros JM. A chance to change the paradigm of outcome assessment of antimicrobial stewardship programs. Clin Infect Dis 2015;61:807–8. doi:10.1093/cid/civ496.

7. Wagner B, Filice GA, Drekonja D, Greer N, MacDonald R, Rutks I, et al. Antimicrobial stewardship programs in inpatient hospital settings: a systematic review. Infect Control Hosp Epidemiol 2014;35:1209–28. doi:10.1086/678057.

8. Schuts EC, Hulscher MEJL, Mouton JW, Verduin CM, Stuart JWTC, Overdiek HWPM, et al. Current evidence on hospital antimicrobial stewardship objectives: A systematic review and meta-analysis. Lancet Infect Dis 2016;16:847–56. doi:10.1016/S1473-3099(16)00065-7.

9. Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, et al. Audit and feedback : effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev 2012;6:Art. No.: CD000259. DOI: 10.1002/14651858.CD000259. doi:10.1002/14651858.CD000259.pub3.

10. Charani E, Edwards R, Sevdalis N, Alexandrou B, Sibley E, Mullett D, et al. Behavior change strategies to influence antimicrobial prescribing in acute care: A systematic review. Clin Infect Dis 2011;53:651–62. doi:10.1093/cid/cir445.

11. Ramsay C, Brown E, Hartman G, Davey P. Room for improvement: A systematic review of the quality of evaluations of interventions to improve hospital antibiotic prescribing. J Antimicrob Chemother 2003;52:764–71. doi:10.1093/jac/dkg460.

12. Drekonja DM, Filice GA, Greer N, Olson A, MacDonald R, Rutks I, et al. Antimicrobial stewardship in outpatient settings: A systematic review. Infect Control Hosp Epidemiol 2015;36:142–52. doi:10.1017/ice.2014.41.

13. Effective Practice and Organisation of Care. No Title. EPOC Resour Rev Authors 2015:http://epoc.cochrane. org/resources/epoc-resources-.

14. Pulcini C, Huttner A. CMI policy on antimicrobial stewardship research. Clin Microbiol Infect 2017. doi:10.1016/j.cmi.2017.11.006.

15. Moher D, Liberati A, Tetzlaff J ADPG. Preferred reporting items for systematic reviews and meta-analyses:

the PRISMA statement. Bmj 2009.

16. McGowan JE. Antimicrobial Stewardship—the State of the Art in 2011 Focus on Outcome and Methods. Infect Control Hosp Epidemiol 2012;33:331–7. doi:10.1086/664755.

17. Gillespie D. Use of co-primary outcomes for trials of antimicrobial stewardship interventions. n.d.

18. Horikoshi Y, Higuchi H, Suwa J, Isogai M, Shoji T, Ito K. Impact of computerized pre-authorization of broad spectrum antibiotics in Pseudomonas aeruginosa at a children's hospital in Japan. J Infect Chemother 2016;22:532–5. doi:10.1016/j.jiac.2016.05.001.

19. Hohn A, Heising B, Hertel S, Baumgarten G, Hochreiter M, Schroeder S. Antibiotic consumption after implementation of a procalcitonin-guided antimicrobial stewardship programme in surgical patients admitted to an intensive care unit: a retrospective before-and-after analysis. Infection 2015;43:405–12. doi:10.1007/s15010-014-0718-x.

20. Vellinga A, Galvin S, Duane S, Callan A, Bennett K, Cormican M, et al. Intervention to improve the quality of antimicrobial prescribing for urinary tract infection: a cluster randomized trial. CMAJ 2016;188:108–15. doi:10.1503/cmaj.150601.

21. Butler CC, Simpson SA, Dunstan F, Rollnick S, Cohen D, Gillespie D, et al. Effectiveness of multifaceted educational programme to reduce antibiotic dispensing in primary care: Practice based randomised controlled trial. BMJ 2012;344. doi:10.1136/bmj.d8173.

22. Persell SD, Doctor JN, Friedberg MW, Meeker D, Friesema E, Cooper A, et al. Behavioral interventions to reduce inappropriate antibiotic prescribing: A randomized pilot trial. BMC Infect Dis 2016;16. doi:10.1186/s12879-016-1715-8.

23. Baur D, Gladstone BP, Burkert F, Carrara E, Foschi F, Döbele S, et al. Effect of antibiotic stewardship on the incidence of infection and colonisation with antibiotic-resistant bacteria and Clostridium difficile infection: a systematic review and meta-analysis. Lancet Infect Dis 2017;17:990–1001. doi:10.1016/S1473-3099(17)30325-0.

24. de Kraker MEA, Abbas M, Huttner B, Harbarth S. Good epidemiological practice: a narrative review of appropriate scientific methods to evaluate the impact of antimicrobial stewardship interventions. Clin Microbiol Infect 2017;23:819–25. doi:10.1016/j.cmi.2017.05.019.

25. Pulcini C, Morel CM, Tacconelli E, Beovic B, de With K, Goossens H, et al. Human resources estimates and funding for antibiotic stewardship teams are urgently needed. Clin Microbiol Infect 2017;23:785–7. doi:10.1016/j.cmi.2017.07.013.

26. McGregor JC, Furuno JP. Optimizing research methods used for the evaluation of antimicrobial stewardship programs. Clin Infect Dis 2014;59:S185–92. doi:10.1093/cid/ciu540.

27. Cantey JB, Wozniak PS, Pruszynski JE, Sánchez PJ. Reducing unnecessary antibiotic use in the neonatal intensive care unit (SCOUT): a prospective interrupted time-series study. Lancet Infect Dis 2016;16:1178–84. doi:10.1016/S1473-3099(16)30205-5.

SUPPLEMENT

Table S1. Definition of the design quality determinants used and corresponding categorisation if applicable

Design quality indicator	Definition	Categorisation (if applicable)
Randomised research design	Allocation of the antimicrobial stewardship intervention or comparator was random. Randomisation is defined as control by means of random allocation at any level (individual or cluster).	Randomised: RCT, c-RCT, cx-RCT, sw-CRT Non-randomised: BA(with or without ITS analysis), cBA, cohort studies, non-randomised trials
External control group	The outcome was also assessed in an external control group without antimicrobial stewardship intervention. External indicates that a contemporaneous cluster or group was included in which the intervention under study is not implemented.	Parallel control group: cBA, RCT, non-randomised trials, c-RCT, cx-RCT, sw-CRT No parallel control group: BA, cohort studies
Number of centres	Amount of centres involved in the study, either as control or intervention.	Single centre: 0-1 centre Multicentre: >1 centre
Sustainability of the intervention sufficiently assessed	The duration of follow-up of the cluster after the intervention was introduced to assess the sustainability of the intervention	Yes: Duration of follow-up ≥12 months No: Duration of follow-up <12 months
Sample size calculation reported	A sample size calculation was performed to ensure sufficient power for the primary outcome.	-
Prospective data collection	The data was prospectively collected. If not reported we assumed the data collection to be retrospective.	-
Confounding correction	The intervention effect was corrected for confounding bias, either by randomisation, matching, stratification or correction.	-
Primary outcome defined	A primary outcome was clearly defined.	-
Clinical outcome reported	Any clinical outcome was reported. Clinical outcomes include mortality, length of stay, readmissions, revisits, etc.	-
Microbiological outcome reported	Any microbiological outcome was reported. Microbiological outcomes include CDI, colonisation or infection with antimicrobial resistant bacteria	-

RCT: randomised controlled trial, c-RCT: parallel cluster randomised controlled trial, cx-RCT: cluster cross-over randomised controlled trial, sw-CRT: stepped wedge cluster randomised trial, BA: before-after study, cBA: controlled before-after study, ITS: interrupted time-series, CDI: *Clostridium difficile* infection

Table S2. Design quality features changes over time of the included studies stratified by studies performed in the community and the hospital setting

Hospital	1977-2004	2005-2010	2011-2013	2014-2015	2016-2017
Randomised design	10%	17%	6%	6%	8%
External control	12%	20%	6%	8%	11%
Sample size calculation	8%	21%	13%	18%	17%
Multicentre	9%	14%	18%	20%	18%
Prospective data	71%	68%	37%	36%	31%
Confounding correction	24%	22%	19%	29%	30%
Primary outcome	38%	50%	37%	45%	48%
Clinical outcome	43%	51%	50%	58%	66%
Microbiological outcome	21%	38%	32%	23%	27%
Sustainability assessed	53%	50%	55%	65%	54%
Community					
Randomised design	43%	47%	41%	41%	61%
External control	45%	51%	41%	45%	65%
Sample size calculation	19%	44%	35%	48%	52%
Multicentre	68%	75%	78%	69%	71%
Prospective data	89%	65%	57%	59%	74%
Confounding correction	58%	42%	54%	66%	65%
Primary outcome	53%	53%	59%	66%	58%
Clinical outcome	32%	29%	27%	31%	29%
Microbiological outcome	6%	5%	8%	17%	10%
Sustainability assessed	43%	56%	78%	55%	52%

Figure S1. Design quality indicators stratified by factors: **(A)** community versus hospital setting, **(B)** financial support versus no financial support, **(C)** age setting: children, adults or both, **(D)** old versus new studies, **(E)** geographical region

C

Adults (n=349)
Children (n=101)
Both (n=375)

E

North America (n=368)
Europe (n=260)
Asia (n=128)
Oceania (n=36)
Africa (n=29)
South America (n=17)

Figure S1 continued.

D

1977-2004 (n=160)
2005-2010 (n=160)
2011-2013 (n=156)
2014-2015 (n=173)
2016-2017 (n=176)

Figure S1 continued.

3

OPTIMISING DESIGN OF RESEARCH TO EVALUATE ANTIBIOTIC STEWARDSHIP INTERVENTIONS; CONSENSUS RECOMMENDATIONS OF A MULTINATIONAL WORKING GROUP

Valentijn A. Schweitzer, Cornelis H. van Werkhoven, Jesús Rodríguez Baño, Julia Bielicki, Stephan Harbarth, Marlies Hulscher, Benedikt Huttner, Jasmin Islam, Paul Little, Celine Pulcini, Alessia Savoldi, Evelina Tacconelli, Jean-Francois Timsit Maarten van Smeden. Martin Wolkewitz, Marc J.M. Bonten, A. Sarah Walker, and Martin J. Llewelyn, Joint Programming Initiative on Antimicrobial Resistance (JPIAMR) Working Group on Design of Antimicrobial Stewardship Evaluations

PART I | CHAPTER 3

OPTIMISING DESIGN OF RESEARCH TO EVALUATE ANTIBIOTIC STEWARDSHIP INTERVENTIONS;
CONSENSUS RECOMMENDATIONS OF A MULTINATIONAL WORKING GROUP

ABSTRACT

SCOPE

Antimicrobial stewardship interventions and programmes aim to ensure effective treatment while minimising antimicrobial-associated harms including resistance. Practice in this vital area is undermined by the poor quality of research addressing both what specific antimicrobial use interventions are effective and how antimicrobial use improvement strategies can be implemented into practice. In 2016 we established a working party to identify the key design features which limit translation of existing research into practice and then to make recommendations for how future studies in this field should be optimally designed. The first part of this work has been published as a systematic review. Here we present the working group's final recommendations.

METHODS

An international working group for design of antimicrobial stewardship intervention evaluations was convened in response to the fourth call for leading expert network proposals by the Joint Programming Initiative on Antimicrobial Resistance (JPIAMR). The group comprised clinical and academic specialists in antimicrobial stewardship and clinical trial design from six European countries. Group members completed a structured questionnaire to establish the scope of work and key issues to develop ahead of a first face-to-face meeting which 1) identified the need for a comprehensive systematic review of study designs in the literature and 2) prioritised key areas where research design considerations restrict translation of findings into practice. The working group's initial outputs were reviewed by independent advisors and additional expertise was sought in specific clinical areas. At a second face-to-face meeting the working group developed a theoretical framework and specific recommendations to support optimal study design. These were finalised by the working group co-ordinators and agreed by all working group members

RECOMMENDATIONS

We propose a theoretical framework in which consideration of the **intervention rationale** the **intervention setting, intervention features** and the **intervention aims** inform selection and prioritization of outcome measures, whether the research sets out to determine **superiority** or **non-inferiority** of the intervention measured by its primary outcome(s), the most appropriate **study design** (e.g. experimental or quasi- experimental) and the **detailed design features**. We make eighteen specific recommendation in three domains: outcomes, objectives and study design.

CONCLUSIONS

Researchers, funders and practitioners will be able to draw on our recommendations to most efficiently evaluate antimicrobial stewardship interventions.

BACKGROUND AND CONTEXT

Antimicrobial resistance is a rapidly growing and major threat to human health[1]. Overuse of antimicrobials drives resistance at the individual[2] and population level[3]. The term antimicrobial stewardship refers to interventions and programmes which aim to optimise antimicrobial use; achieving effective treatment while minimising antimicrobial-associated harms including resistance[4].

Despite the large and exponentially increasing number of studies published since the term Antimicrobial Stewardship was coined [5-7], evidence remains remarkably weak both for **what** specific antimicrobial use interventions are effective (in terms of mortality, length of stay, adverse events, resistance rates) and **how** antimicrobial use improvement strategies can be implemented to deliver the desired antimicrobial use in daily clinical practice[8]. A 2016 systematic review of evidence supporting key antimicrobial use interventions (e.g. prescribing according to guidelines, de-escalation of therapy, intravenous to oral switching) identified predominantly low-quality and highly heterogenous supporting evidence[9]. The evidence around improvement strategies is similarly weak, dominated by uncontrolled before-after studies and inadequately performed interrupted time series analyses, mostly performed within single hospitals[10].

We recently reported a broad systematic review of antimicrobial stewardship intervention studies which highlighted key frequent design weaknesses[7]. Studies which aim to assess effectiveness of antimicrobial use interventions are typically under powered and fail to provide evidence on safety or even do not report clinical outcome data at all. Improvement strategy studies are often multifaceted with inadequate process evaluation to allow mediators of impact to be assessed[11]. Generally, the field of antimicrobial stewardship research is dominated by single-centre observational and quasi-experimental studies which fail to deal optimally with risks of different forms of bias and that lack external validity[7, 8].

Building on this work we established a working group of investigators in this field which used a consensus-building iterative process over 12 months to build a conceptual framework and develop specific recommendations for the design of stewardship

PART I | CHAPTER 3

OPTIMISING DESIGN OF RESEARCH TO EVALUATE ANTIBIOTIC STEWARDSHIP INTERVENTIONS;
CONSENSUS RECOMMENDATIONS OF A MULTINATIONAL WORKING GROUP

evaluations, which were then reviewed and amended by an expert advisory committee. This guidance is the final result of that process and aims to support investigators when making key design decisions and funders assessing proposals for studies of antimicrobial stewardship interventions and hopefully enhances the quality and impact of research in this crucial area.

METHODS

An international working group for design of antimicrobial stewardship intervention evaluations was convened in response to the fourth call for leading expert network proposals by the Joint Programming Initiative on Antimicrobial Resistance (JPIAMR). The study sponsor was the UK Medical Research Council. The working group co-ordinators (MJMB, MJL) and co-applicants (VAS, ASW and CHvW) purposively selected an additional eight leading clinical and academic specialists in antimicrobial stewardship and clinical trial design from six European countries (France, Germany, Italy, the Netherlands, Spain, Switzerland and the UK) to contribute. Selection secured input from the diversity of professionals involved in antimicrobial stewardship practice (infection, internal medicine, intensive care medicine) and research (trial design, statistics and qualitative research) disciplines. Consensus was sought through a nominal group process. Group members completed a structured questionnaire to establish the scope of work, key study designs used in antimicrobial stewardship, identify the major limitations on different study designs and key issues to develop ahead of a first face-to-face meeting. The group met in March 2017 and anonymised responses were feedback to the whole group and relevant literature was presented (VAS, CHvW, MJL). This identified the need for a comprehensive systematic review of study designs in the literature. In parallel, in moderated small group work, candidate solutions were proposed to address the limitations identified, and in a final round-table moderated discussion the group prioritised four key areas where research design considerations restrict translation of findings into practice: features of the intervention under evaluation; appropriate selection of outcome measures; demonstration of superiority / non-inferiority of the intervention according to the outcome measures selected and strategies to minimise bias within experimental and quasi-experimental study designs. The working group's initial outputs were reviewed by two independent advisory experts, both senior, clinically active antimicrobial stewardship experts in different European countries. Their input prompted widening the group to bring in additional expertise in the field of implementation research, primary care and paediatrics. A second face-to-face meeting the working group used the findings of the systematic review to develop a theoretical framework through which researchers can address these four key research

design considerations. The group proposed a series of key questions researchers can use to highlight the major issues they need to address to arrive at an optimal design for their specific research project. Final agreement of recommendations presented here by all eighteen members of the working group was achieved by email.

A THEORETICAL FRAMEWORK FOR DESIGNING ANTIMICROBIAL STEWARDSHIP EVALUATIONS

The impact of intervention design

Detailed discussion of how antimicrobial stewardship interventions are designed is beyond the scope of this guidance. However, the design of the scientific evaluation of an intervention depends on how that intervention was designed, and this then may depend on a set of interdependent considerations (Figure 1a). The **intervention rationale** should include its basis in theory and existing evidence. (Table 1 is a glossary of terms used in this guidance). The existing evidence that informed the research question should be clearly explained on an efficacy-effectiveness-implementation spectrum [12], as these considerations will determine how outcomes are selected and prioritized (Figure 1b).
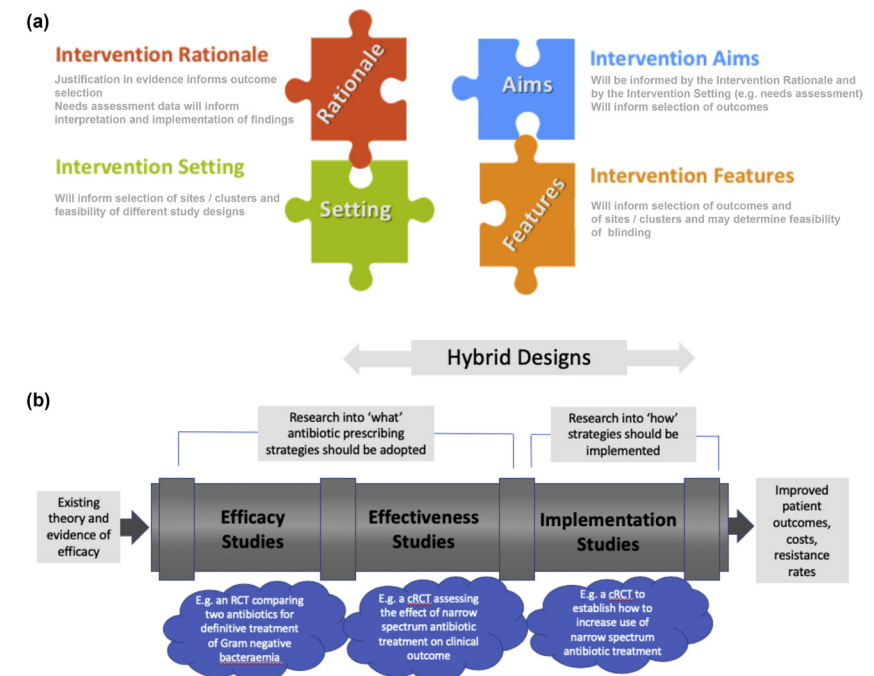
Figure 1. (A) Interacting considerations relating to the intervention to be evaluated and their impact on study design. **(B)** An evaluation pipeline for antimicrobial stewardship intervention. Adapted from [12].

PART I | CHAPTER 3

OPTIMISING DESIGN OF RESEARCH TO EVALUATE ANTIBIOTIC STEWARDSHIP INTERVENTIONS;
CONSENSUS RECOMMENDATIONS OF A MULTINATIONAL WORKING GROUP

Table 1. Glossary of terms

Term	Explanation
Intervention rationale	The theory and evidence behind the stewardship intervention which is to be evaluated encompassing external factors (e.g. behavioural theory, evidence from previous research) and the clinical setting.
Clinical setting	The environment in which the intervention is evaluated, both physical (e.g. ICU, emergency room, hospital type, primary care, long-term care) and practical (e.g. prescribing practice, team structures, staffing, behaviour).
Intervention aim(s)	The improvement being sought (e.g. reduction in inappropriate antimicrobial prescribing, reduction in use of specific antimicrobial classes or reduced *Clostridium difficile* infection)?
Features of the intervention	The different elements which make up a multifaceted intervention (e.g. education, decision support).
Cluster	A unit representing a group of smaller components, at which an intervention is delivered (e.g. a hospital ward representing all the doctors working in it, a group of primary care physicians working in a practice)
Outcomes of interest	The outcomes measured to determine effectiveness, safety and costs of the intervention.
Experimental design studies	Studies which use randomisation to allocate the stewardship intervention and control, either to individual patients/professionnals or clusters of patients/professionals.
Quasi-experimental design studies	Studies which don't use randomisation to allocate the stewardship intervention but rather use as controls different time period(s) and/or site (s), either external (controlled before-after studies) or internal (interrupted time series analyses, before-after studies).
Contamination	Unintended exposure of patients in the control phase or cluster to some or all of the intervention.
Efficacy study	A study which assesses whether an antimicrobial use intervention produces the expected result under ideal and controlled conditions.
Effectiveness study	A study which assesses whether an antimicrobial use intervention produces the expected result under 'real-world' pragmatic conditions.
Implementation Study	A study which assesses the impact of an antimicrobial use improvement strategy in daily practice
Mediator analyses	Techniques to investigate mechanisms through which complex interventions achieve an observed effect
Superiority analysis	An analysis which sets out to determine if the intervention or strategy being assessed is better than comparator
Non-inferiority analysis	An analysis which sets out to determine whether the intervention or strategy being assessed not worse (by a prespecified amount, the non-inferiority margin) than comparator
Process Indicators	Measures of the care that is actually delivered to the patients (e.g., empirical regimen according to guidleine)
Structure indicators	Measures of the organization of the healthcare system (e.g., the availability of a stewardship team)
Ecological assessment (of antimicrobial resistance)	Measurement of burden if antimicrobial resistant organism(s) or gene(s) in the environment or aggregated patient samples

Detailed characterization of the **intervention setting** is required to allow assessment of external validity and to minimize selection bias. Stewardship interventions are typically multifaceted and each **intervention feature** must be specified precisely. The same holds for how the intervention's impact will be determined; this will influence definition and selection of outcomes, selection of clusters/sites and feasibility of blinding. The **intervention aims** will be informed by the rationale and setting and will also be key to selecting the **primary** and **secondary outcomes;** whether these will determine effectiveness and safety or how implementation results change antimicrobial use and what data are required to support translation of study findings into practice. These considerations will inform whether the research sets out to determine **superiority** or **non-inferiority** of the intervention measured by its primary outcome(s) against standard practice and the detectable effect sizes/non-inferiority margins, the most appropriate **study design** (e.g. experimental or quasi- experimental) and the **detailed design features**.

Recommendations regarding selection of outcome measures

When assessing the impact of a stewardship intervention, researchers should aim to consider all intended and potential unintended effects[13-15]. Outcome measures can be helpfully grouped into three domains as clinical (typically to assess safety of an antimicrobial-sparing intervention in terms of patient outcome), microbiological (resistance), and care-related (processes and structures of care, sometimes referred to as quality or performance outcomes)[16] (Table 2). Whether the study is primarily assessing effectiveness, implementation or a combination of both, will determine how outcomes are selected and prioritised, but, in general, appropriate outcome measures should be prospectively defined from each of the three domains. It is essential to recognise that whilst individually randomised efficacy trials aim to avoid selection bias, the inevitably restricted populations that enter such trials can potentially lead to generalisability bias, making extrapolation to wider populations challenging. While stewardship studies typically assess interventions made at the cluster level, assessment of clinical, microbiological and care related outcomes is often possible at an individual patient level and should be included where possible to address this.

Table 2. Outcome measures in antimicrobial stewardship evaluations

Clinical outcome measures	
Examples	**Notes**
Clinical cure, clinical failure, time to clinical response, recurrence rate.	Typically used to determine the safety of the intervention in terms of patient treatment outcome.
Mortality, length of stay, need for escalation of care (e.g. from ward to high dependency or critical care), (re)admission to hospital, revisits	May include microbiological evidence of clinical outcome (e.g. microbiological cure or recurrence). Most are directly relevant to the individual patient.
Patient reported outcomes (e.g. quality of life measures).	Important safety outcomes which are relatively easy to gather at cluster-level, but may only be linked partially to the intervention and may be a long way down the patient pathway
Adverse drug reactions, drug–drug interactions	Gathering relevant data may require individual consent but could be from a subset of patients or use anonymised electronic records.

Microbiological (resistance) outcome measures	
Examples	**Notes**
Colonisation by antimicrobial resistant pathogens (e.g. MRSA or multi-drug resistant (MDR) Enterobacteriaceae)	Valuable as short-term surrogate measures of antimicrobial resistance-related harm but relevance to individual patients is indirect through risk of antimicrobial resistant infection in the future or through transmission. Ecological assessments may be more feasible than individual patient-level measurement.
Infection by specific organisms (C. *difficile*, antimicrobial resistant bacteria)	Outcome directly relevant to the impact of the antimicrobial intervention on the individual patient but uncommon and may require long follow-up beyond that needed for clinical outomes

PART I | CHAPTER 3

OPTIMISING DESIGN OF RESEARCH TO EVALUATE ANTIBIOTIC STEWARDSHIP INTERVENTIONS;
CONSENSUS RECOMMENDATIONS OF A MULTINATIONAL WORKING GROUP

Table 2 continued.

Care provision (quality or performance) outcome measures	
Examples	Notes
Drug use (e.g. Defined daily doses (DDD) or Days of Therapy (DOT) per admission or per bed-day)	Measurement of antimicrobial use (e.g. volume, range of agents) used to determine whether the intervention has potential to
Appropriateness of treatment (e.g. proportion of prescriptions in accordance with guidelines)	have an effect on clinical or microbiological outcomes (if no impact on process, then no clinical/microbiological impact by
Measures of intervention (e.g. recommendations given, use of clinical decision support)	definition)
Resource requirements (e.g. staff time, clinical consultations, diagnostic testing)	Can be selected to measure appropriateness of antimicrobial selection
Costs measures	Important for health-economic analyses and assessment of sustainability
	Important for mediator analyses.

Clinical outcomes are missing from many published stewardship studies. In fact, most of these studies were not sufficiently powered to exclude clinically meaningful harm. Concern that this prevents adoption of antimicrobial reduction strategies into practice has led some to call for routine use of co-primary clinical outcomes in stewardship evaluations[17]. The working group felt that clinical outcome measures should always be pre-specified and reported. Exceptions could be implementation studies of interventions for which concerns over safety will not be a barrier to adoption of their findings.

Microbiological outcomes address the impact of the intervention on antimicrobial resistance and/or rates of *Clostridium difficile* infection. A central rationale for antimicrobial stewardship interventions is that reducing antimicrobial exposure should reduce harm to a patient's microbiome and selection for antibiotic resistance. However, the evidence base remains sparse, and mostly of low quality, with lack of reliable pre-intervention data a particular limitation[9, 18, 19]. Incorporating assessment of colonisation/infection by resistant organisms within a stewardship study can be challenging as event rates are often low and the relationship between antimicrobial exposure and resistance may be temporally distant and complicated by interactions with exposure to resistant pathogens and infection control measures. The working group agreed that while reductions in antimicrobial resistance should not be the primary outcome of stewardship studies, measurement of prevalence or incidence of *C. difficile* infection and of antimicrobial resistance should be included in the design where possible, and it should be clear whether measured resistance is in relation to the infecting pathogen and type of infection or among colonising strains.

Care provision outcome measures (sometimes called quality or performance measures) include process indicators, prescribing behaviours, and antimicrobial use data. These are usually relatively straightforward to obtain and are important to gather and report since clinical outcomes can only be interpreted meaningfully if it is clear that patient management has truly changed. Process indicators may address prescribing

quality (e.g. guideline adherence or documentation practice) and reveal mediators of observed results. They are particularly important in implementation research to assess how the intervention under evaluation was actually delivered across the study (fidelity). This allows distinction between strategies that do and do not change the behaviours they aim to change and identification of those elements of an intervention that are impactful and of barriers for implementation[11]. Gathering appropriate qualitative data (e.g. from service managers, care providers and patients as appropriate) will allow an intervention's impact on cultural aspects of antibiotic use to be evaluated. Process outcomes are needed to assess organisational impact, of both implementation and long-term sustainability. Sustainability assessment is particularly important when an intervention has significant organisational-level impact through diversion of activity or cost[20]. For detailed consideration of these issues researchers should consult current guidance on development and evaluation of complex interventions[21].

TIMING OF OUTCOME MEASUREMENTS

Within each domain of outcome measure, consideration must be given to appropriate timing depending on the nature of the intervention and population (e.g. long and short term mortality, clinical complications during hospitalisation or after discharge). Timing of measurement of microbiological outcomes should be considered to assess impact on resistance including *C. difficile* and timing of process outcome measurements should be considered to assess long-term sustainability.

Establishing superiority or non-inferiority

Where a stewardship study sets out to establish the effectiveness of an intervention, incorporation of appropriate controls is essential if the results are to inform practice, irrespective of whether an experimental or non-experimental design is used (see below). Researchers need to decide whether their primary objective is to determine superiority or non-inferiority of the intervention *vs* control.

Interventions aiming to improve treatment outcome. In some situations, a relevant clinical benefit can be hypothesised for an intervention (e.g. an intervention that focuses on increasing earlier targeted treatment based on test results or preventing under-treatment) and a study assessing the effectiveness of the intervention would seek superiority of the intervention *vs.* control for an appropriate primary clinical outcome.

Intervention aims to reduce antimicrobial exposure. In most situations, stewardship interventions aim to preserve clinical outcome while reducing unnecessary antimicrobial exposure (e.g. less inappropriate initiation of antibiotics, choice of narrower spectrum or shorter duration) and improving quality of prescribing. As a result there is often

PART I | CHAPTER 3

OPTIMISING DESIGN OF RESEARCH TO EVALUATE ANTIBIOTIC STEWARDSHIP INTERVENTIONS;
CONSENSUS RECOMMENDATIONS OF A MULTINATIONAL WORKING GROUP

some degree of real or perceived risk of patient-level harm, which may be specific to the intervention, patient population, setting and disease. Researchers designing effectiveness evaluations should consider what potential for patient harm would prevent adoption of the intervention *even if it were effective in reducing antimicrobial exposure*. Researchers should select appropriate secondary clinical endpoint(s) to address this concern. Ideally in this situation the research should seek both superiority for an appropriate process measure *and* non-inferiority (i.e. not qualitatively worse than control) for a co-primary clinical outcome. The key measure to assess non-inferiority is the non-inferiority margin, being the smallest outcome difference for which the intervention would be considered no worse than control. The size of the non-inferiority margin strongly influences the sample size required to demonstrate non-inferiority with sufficient power. What margin is chosen depends on the outcome selected. The margin needs to be small enough to exclude relevant harm, which would prevent intervention implementation into practice. Researchers should justify the non-inferiority margin chosen with regard to severity and frequency of the outcome in the control group (which may, for example be affected by case-mix[22]).

Naturally, trials designed for demonstrating non-inferiority of clinical outcomes usually require large sample sizes. In such trials an interim analysis of a process outcome could be used to determine futility; if the intervention does not lead to the pursued process change continuing that intervention may not be logical, as non-inferiority will be the inevitable outcome.

Recognising that achieving adequate power to exclude clinically relevant non-inferiority will not always be feasible, the group felt that researchers should at least specify and report point estimates and confidence intervals for a single prespecified lead clinical outcome. Bayesian analyses may be helpful to directly estimate the probability that intervention is more than 2.5%, 5%, 7.5% etc inferior to control[23]. Researchers should also prespecify the clinical outcomes they will use to assess the safety of the intervention, and all available clinical outcome data should be reported, in order to allow future meta-analysis. Unavailability of data should be explained. Unplanned exploratory analyses of clinical outcomes should be reported as such.

In studies addressing how interventions with established efficacy should be implemented, the quantitative outcome measures will be predominantly process measures and comparisons will seek to determine superiority of the intervention over comparator.

Sample size calculations
Studies evaluating effectiveness of an antimicrobial intervention need to be powered

to demonstrate clinically relevant non-inferiority. In a superiority trial, detecting a large effect with high probability is almost always possible at a feasible sample size. Whereas demonstrating superiority only requires the confidence interval for the effect estimate to exclude zero, regardless of its width, determining non-inferiority requires the entire confidence interval to lie below the non-inferiority margin[24]. As a result, much larger participant numbers are usually required to demonstrate non-inferiority within clinically relevant margins which may be very small and difficult to define for outcomes such as mortality[25]. This difference lies in that superiority trials tend to be powered on an expected effect, which is often larger than what would be deemed a clinically relevant effect, whereas non-inferiority trials need to be powered on a clinically relevant effect. One proposed solution to this issue is the Desirability of Outcome Ranking (DOOR)/Response Adjusted for Days of Antibiotic Risk (RADAR) approach which uses investigator ranked composite outcomes. This approach is based on the assumption that the same outcome with less antimicrobial exposure is desirable[26]. Yet, problems with clinical interpretation and sensitivity to the clinical outcomes chosen have been reported[27, 28]. It remains to be determined to what extent the RADAR approach can robustly establish the effectiveness of novel stewardship interventions.

Interrupted time series studies require enough sequential measures before and after the intervention; the study's power will depend on the number of data points, their distribution, variability, the expected strength of the intervention effect and confounding factors such as seasonality[29], and therefore there are no straightforward sample size formulae. Researchers should consider the minimal requirements set out in the Cochrane Effective Practice and Organisation of Care (EPOC) resources[30].

STUDY DESIGN
Stewardship interventions typically target prescribers/professionals rather than individual patients. As a consequence, evaluations involving individual patient randomisation are usually not possible because of contamination. Instead, intervention allocation must be clustered (e.g. hospital, ward, primary care practice, or physician). An important advantage of allocation at the cluster level is that it is more representative of real-life clinical practice. It is therefore more suited to studying both antimicrobial use interventions and antimicrobial improvement strategies rather than efficacy. Whereas in individual patient trials, randomisation can be expected to control for confounding bias and maximise internal validity, with cluster randomised controlled trials (cRCT), researchers need to give careful consideration to how clusters are defined and characterised. Clusters should be defined at the lowest level (e.g. clinical team, ward, practice, hospital) where contamination is unlikely as this will maximise the number of available clusters and hence study power. However, with the small number of clusters

PART I | CHAPTER 3

OPTIMISING DESIGN OF RESEARCH TO EVALUATE ANTIBIOTIC STEWARDSHIP INTERVENTIONS; CONSENSUS RECOMMENDATIONS OF A MULTINATIONAL WORKING GROUP

typically available in stewardship evaluations, randomisation cannot be relied on to avoid imbalance between intervention and control clusters. Therefore baseline imbalances which may influence the intervention's impact (e.g. antimicrobial use, antimicrobial resistance rates, infection control standards, antimicrobial stewardship structures and processes, case-mix of patients) should be specified *a priori* and data on these should be gathered for inclusion in multivariate analyses. Baseline imbalance in factors which a strong association with outcome or that could potentially modify the effect of the intervention can be addressed through stratified randomisation (e.g. putting clusters into similar pairs and allocating one of each pair randomly to intervention vs control), or use of a cross-over design (see below). Cluster characterisation is also essential to understand any observed heterogeneity of the intervention's effect between clusters. It optimises external validity by allowing others to judge the representativeness for their clinical practice and to understand the logistical challenges of implementation.

Experimental study designs (Table 3)

Three main forms of cluster-randomised design may be appropriate depending on the intervention. As above, **parallel cRCTs**, in which each cluster is randomised to either the intervention or control, minimise risk of contamination and maximise independence of the intervention from cluster-level characteristics. In some situations, perceptions of the intervention may influence whether clusters are willing to be randomised to control or intervention arms and hamper participation or introduce bias. **Stepped-wedge cRCTs (swcRCTs)** overcome this issue since all clusters receive the intervention during the trial, and allow estimation of the intervention effect within each cluster. swcRCTs can be logistically challenging to deliver since some clusters may have to wait to introduce the intervention and exposure should be avoided. Furthermore, the analysis of swcRCT is more complex[31]. Randomisation of time of implementation is crucial to ensure independence of the timing of introduction from cluster-level factors. **Cross-over cRCTs** offer the potential to estimate intervention effects in both directions – i.e. introducing and withdrawing, but may not be practicable (e.g. it may not be feasible to withdraw an educational intervention. Alternatively, the washout phase of a cross-over study may be considered an assessment of sustainability for some forms of intervention. Assessment of carried antimicrobial resistance in crossover designs may need to consider the potential for resistance selection to persist.

A particular challenge with evaluation of interventions made at a cluster rather than patient-level is intracluster correlation[32]. This must be incorporated into the sample size calculation otherwise a trial may be underpowered. Intracluster correlation is the extent to which patients are more similar to each other within a cluster than they would be if selected at random. The intracluster correlation coefficient (ICC) of an outcome

is a measure of the relatedness of clustered data by comparing the variance within clusters (e.g. hospitals) with the variance between clusters. A high ICC means that observations within clusters are much more similar to each other than to observations in other clusters, while an ICC of zero means that observations within one cluster are equally similar to each other than to observations in other clusters. In general, if the ICC is large, research designs with cross-over are more efficient, while if the ICC is low, parallel cluster designs are more efficient[32].

Table 3. Design recommendations for experimental evaluations antimicrobial stewardship Interventions

Feature	Recommendations		
	Parallel cRCTs	**Stepped-wedge cRCTs**	**Crossover cRCTs**
Cluster selection	Randomised implementation at the lowest level (e.g. prescriber, ward, hospital, primary care practice) at which contamination can be minimised Define eligibility criteria and document representativeness of included clusters with respect to system from which they are drawn (e.g. size, case mix)		
Cluster allocation and randomisation, timing of intervention	Ensure allocation concealment until the intervention is implemented (as complete blinding to allocation after randomisation is often not feasible).	Conceal timing and order of intervention / cross-over as much as possible Timing of intervention should be determined externally and at random, where possible	
Cluster balance	Pursue good/excellent balance between clusters (e.g. matching, stratified randomisation based on factors likely to be associated with the outcome under study). No lower limit above which randomisation will ensure balance but particularly problematic if there are fewer than 20 clusters per randomised group. Collect data to document balance between clusters.	Good/excellent balance between clusters achieved through design.	
Blinding	Consider the objectivity of the selected outcomes and the extent to which patients and assessors of outcomes can be blinded to the cluster allocation		
Outcomes	Specify a primary or co-primary process outcome Specify a co-primary clinical outcome or at minimum one lead clinical outcome, and specify and report secondary clinical outcomes even if not powered on these Specify and analyse outcomes in each domain – clinical, microbiological, process (quantity or quality of antimicrobial use) Within implementation research, process outcomes should be selected with regard to complex intervention methodology [20] e.g. measures of fidelity, mediators and modifiers of the intended effect and measures of organisational impact Consider all important harms / unintended effects including 'squeezing the balloon' effects in which achieving the intended reduction in antimicrobial overuse results in an unintended increase in harmful overuse elsewhere [14, 15, 37]. Define timing of different cluster-level and individual-level outcomes		
Power calculation	Provide sample size calculations to demonstrate study power – for the primary / co-primary outcome(s), and taking intra-cluster correlation into account		
Analysis	Adjust for secular trends (particularly for stepped-wedge cRCTs)		
Selection of patients for outcome evaluation	Ensure robust consistent inclusion of patients in control and intervention clusters / phases. Report denominators from whom included patients were selected wherever possible.		
Follow-up of patients	Timing of patient follow-up to assess patient-level outcomes should consider relevant timescales for both effectiveness and harms		
Follow-up of clusters	Consider duration of follow up both for immediate effect of the intervention and sustainability		Only possible with short-term interventions with rapid loss of effect post withdrawal
Reporting	Report according to CONSORT criteria for cluster RCTs, stepped-wedge cRCTs, and other CONSORT guidelines as appropriate (e.g. pragmatic trials, non-inferiority trials). Consider using the TiDier checklist to clearly describe any behavioural intervention [38].		

PART I | CHAPTER 3

OPTIMISING DESIGN OF RESEARCH TO EVALUATE ANTIBIOTIC STEWARDSHIP INTERVENTIONS;
CONSENSUS RECOMMENDATIONS OF A MULTINATIONAL WORKING GROUP

Quasi-experimental study designs **(Table 4)**

In situations where randomisation is not feasible or ethically not acceptable (see below), quasi-experimental, before-after-studies have the potential to deliver robust evidence of a causal relationship between an intervention and measured outcomes if they incorporate appropriate controls and analyses which account for time trends. Where control is provided through comparison with centre(s) where the intervention is not introduced, the term **Controlled Before-After (CBA)** study is used. Where control is provided by use of pre-intervention observations within centres, and secular time-trends in the outcomes are specifically accounted for, the term **Interrupted Time Series (ITS)** study is used. In practice, ITS reflects a method of analysis, being used for before and after studies and CBA, rather than a specific study type and can also be applied to CBA studies. CBA studies which do not control for time-trends are unlikely to provide reliable evidence, regardless of external control[19]. The working group agreed that, design of quasi-experimental evaluations of stewardship interventions must always account for changes in time[33, 34]. Such analyses require sufficient pre-intervention time points to incorporate segmented regression analysis, and should consider adjustment for autocorrelation (e.g. using ARIMA models). Such analyses should report immediate effects on outcome and trends before and after the implementation, and assess whether trends are non-linear[29, 35]. Furthermore the timing of intervention implementation must be externally set to avoid the problem of regression to the mean which occurs when sites introduce a stewardship intervention in response to deterioration in the chosen outcome measure. Detailed guidance on conduct of Interrupted Time Series analyses are available through EPOC[30] and described in a recent review[36].

Ethical considerations

Antimicrobial stewardship measures which balance immediate and individual risks against future and societal access to effective antimicrobials raise challenging ethical issues around intergenerational justice, global distributive justice and protection of public health[37]. A key ethical issue in stewardship research is that, by gathering evidence for safety through clinical outcome measures, the possibility of individual harm is acknowledged. Individual patient consent may not be feasible in studies of interventions which act on prescribers or structures such as hospitals or clinics. This may set a higher ethical barrier than for individually randomized studies in which informed consent can be obtained. In this situation the research design process should involve patients to ensure that independent non-research views from the relevant patient population about these trade-offs are heard, actively considered, and incorporated into the final design. Additionally, researchers should be able to justify why the interventions under examination are reasonable choices of practice which could also be made outside the study setting. Studies in which the intervention is made at a cluster level will often still

use individual patient data. Any requirement for individual patient consent to collect data may lead to loss of representativeness and a biased assessment of the intervention effect. Because consent is acquired with knowledge of the intervention, there is an increased risk of selection bias, e.g. if investigators are more motivated to enroll patients during the intervention period. Depending on the national regulations, in some countries study designs can address this issue through use of de-identified or anonymous data (e.g. through electronic patient records) of parameters collected routinely in clinical practice without the need for individual patient consent.

Table 4. Design recommendations for quasi-experimental evaluations antimicrobial stewardship Interventions

Feature	Recommendations
Control	Even in situations where randomisation is not possible (e.g. too few available clusters) allocation to intervention or control group should be made externally if at all possible, i.e. not depending on known factors or clinician preference
	Consider trying to match controls to minimise risk of bias arising from intrinsic differences between control and intervention groups
Timing	Timing of intervention should be externally set OR if this is not possible timing must be explained and described
Data	Data from automated electronic data recording (e.g. antimicrobial use data, routine electronic patient data) can be used retrospectively for pre-intervention data providing that collection/entry is consistent over calendar time, otherwise all data should be collected prospectively
	Measure, report and analyse any concurrent changes in case-mix, changes in methodology of outcome assessment, and care practices
Analysis	Include at least 12 monthly time points before and after the intervention to allow for anticipated secular trends [35, 39]
	Use segmented regression or ARIMA models to account for secular trends.
	Include at least 100 observations per time point [39].
	Check and, if necessary, correct for autocorrelation.
Outcomes	See table 3
Follow-up of patients	Timing of patient follow-up to assess patient-level outcomes should consider relevant timescales for both effectiveness and harms
Follow-up of clusters	Consider duration of follow up both for immediate effect of the intervention and sustainability
Reporting	Report according to relevant recommendations; STROBE-AMS [40] or STROBE [41] and the TiDier checklist [38], SQUIRE to describe in detail quality improvement component of study [42], TREND statement for nonrandomized evaluations of behavioural and public health interventions [43].

KEY DESIGN DECISIONS

The consensus group considered that researchers planning antimicrobial stewardship evaluations must make a set of key decisions (Table 5) which will ultimately determine optimal study design. We have classified these decisions based on whether they apply to the *intervention itself*, the *evaluation setting*, the *outcomes of interest*, the *research objective* and *type of study*. Detailed explanation of the decisions are presented in **supplementary materials**.

Table 5. Key Design Decisions. A detailed explanation of the rationale and how these address different aspects of design is set out in the supplementary materials

Question	Design aspect addressed
Where does knowledge gap the study aims to address lie on a spectrum between 'what' and 'how' questions?	selection and prioritisation of outcomes
What are the risks of contamination?	how clusters will be defined within the study.
Is it possible to remove the intervention after it has been implemented?	what study design will be most appropriate.
Is the intervention impact threatened by sustainability?	selection and timing of study outcomes
What forms of bias threaten the validity of the study?	cluster selection; feasibility of blinding; data collection
What features of the evaluation setting will impact on external validity?	cluster selection; feasibility of blinding; data collection
Is it possible to blindly assess the outcome?	feasibility of blinding

DISCUSSION AND CONCLUSIONS

The theoretical framework and design recommendations we present have been developed by a diverse international working group with broad and substantial expertise in antimicrobial stewardship research and practice. They address aspects of study design which are crucial to translation of research into practice and will, we believe, increase the impact of future research in this field. By drawing on wide-expertise and building our comprehensive systematic review we consider our recommendations relevant across diverse settings of care. Our work has some notable limitations. Although we gave careful consideration to the breadth of expertise required on the group and sought external advice, we did not seek lay input. We cannot discount the possibility that this would have changed our emphasis, around patient reported outcome or experience measures for example. Given the technical nature of our guidance we think it unlikely this would have changed our conclusions. An inherent risk of the consensus-group design is 'group think' in which members trying to reach consensus fail to critically evaluate alternative views. To address this we sought critical evaluation by two highly eminent international experts in this field. Although these were also, of necessity, experts in antimicrobial stewardship research, the impact of their input on our thinking, the breadth and seniority of expertise in our group make it unlikely we have failed to consider major alternative viewpoints. Notwithstanding these caveats, we believe that application of this guidance has the potential to greatly improve the quality and impact of antimicrobial stewardship research.

SUMMARY RECOMMENDATIONS

OUTCOMES

- Researchers should determine whether their study aims to investigate, effectiveness, or implementation ('what or 'how'). This will determine the priority and nature of outcomes.

- All antimicrobial stewardship studies should define process, clinical and microbiological outcomes and specify a primary process outcome(s) to measure effectiveness of the intervention.

- Unless there is pre-existing evidence that a stewardship intervention cannot or will not compromise treatment outcome, an evaluation should attempt to pre-specify a co-primary clinical/microbiological efficacy outcome on which the study is adequately powered, or, at minimum, a single lead clinical outcome.

- Clinical and microbiological data documenting treatment outcome should be collected and reported as pre-specified secondary outcomes even if the study is not powered on them

- Measurement of incidence of infections / colonisation due to multi-drug resistant bacteria and infections due to *C. difficile* infection should be included in the design of stewardship interventions whenever possible. Studies assessing resistance should clarify whether this is related to the infecting pathogen or among colonisers.

OBJECTIVES

- If a relevant clinical benefit can be hypothesised for an intervention, then the research objective should seek superiority for an appropriate primary clinical outcome.

- If not, researchers should seek both superiority for an appropriate process measure and ideally non-inferiority for a co-primary clinical/clinically relevant microbiological outcome.

- Researchers should justify how the non-inferiority margin has been selected and balanced against research costs and feasibility.

- Where this is not possible, as a minimum, researchers should specify, and report point estimates and confidence intervals for, at minimum, a single pre-specified lead clinical outcome.

- In situations where the study size is determined by a co-primary non-inferiority safety outcome, an interim futility analysis of the superiority process outcome should be considered to confirm a relevant change in treatment/management.

STUDY DESIGN

- Cluster randomised controlled trials (including crossover and stepped-wedge designs) are preferable to quasi-experimental before/after studies.

- The threshold for defining clusters should be as low as possible to minimise contamination, allowing the maximum number of clusters to be studied.

- In a parallel cluster RCT, randomisation should not be relied on to control for imbalance between study arms if the number of clusters is <20 per arm and stratified

PART I | CHAPTER 3

OPTIMISING DESIGN OF RESEARCH TO EVALUATE ANTIBIOTIC STEWARDSHIP INTERVENTIONS;
CONSENSUS RECOMMENDATIONS OF A MULTINATIONAL WORKING GROUP

or matched randomisation should be considered

- Designs using within-cluster comparisons (stepped-wedge cRCT, cross-over cRCT or quasi-experimental approaches) are indicated where there are fewer than 10 clusters per arm.

- Quasi-experimental studies should incorporate appropriate controls and analyses to account for time trends

- In quasi-experimental studies, timing of the intervention should be externally set or if this is not possible timing should be explained and described.

- Segmented regression analysis of interrupted time series studies should include 12 time points with at least 100 observations per time point before and after the intervention to allow for anticipated secular trends and test or correct for autocorrelation.

- Single centre studies using a robustly designed and analysed interrupted time series approach including observations before and after the intervention should be considered the lowest quality research design which will impact on clinical practice.

TRANSPARENCY DECLARATION

ACKNOWLEDGEMENTS

REFERENCES

1. The Review on Antimicrobial Resistance. Antimicrobial Resistance: Tackling a crisis for the health and wealth of nations. 2014.

2. Costelloe C, Metcalfe C, Lovering A, Mant D, Hay AD. Effect of antibiotic prescribing in primary care on antimicrobial resistance in individual patients: systematic review and meta-analysis. BMJ. 2010;340:c2096.

3. Goossens H, Ferech M, Vander Stichele R, Elseviers M, ESAC Project Group. Outpatient antibiotic use in Europe and association with resistance: a cross-national database study. Lancet. 2005;365(9459):579-87.

4. Dyar OJ, Huttner B, Schouten J, Pulcini C, Esgap. What is antimicrobial stewardship? Clin Microbiol Infect. 2017;23(11):793-8.

5. McGowan JE, Jr., Gerding DN. Does antibiotic restriction prevent resistance? New Horiz. 1996;4(3):370-6.

6. Molina J, Cisneros JM. Editorial Commentary: A Chance to Change the Paradigm of Outcome Assessment of Antimicrobial Stewardship Programs. Clin Infect Dis. 2015;61(5):807-8.

7. Schweitzer VA, van Heijl I, van Werkhoven CH, Islam J, Hendriks-Spoor KD, Bielicki J, et al. The quality of studies evaluating antimicrobial stewardship interventions: a systematic review. Clin Microbiol Infect. 2018.

8. Hulscher M, Prins JM. Antibiotic stewardship: does it work in hospital practice? A review of the evidence base. Clin Microbiol Infect. 2017;23(11):799-805.

9. Schuts EC, Hulscher M, Mouton JW, Verduin CM, Stuart J, Overdiek H, et al. Current evidence on hospital antimicrobial stewardship objectives: a systematic review and meta-analysis. Lancet Infect Dis. 2016;16(7):847-56.

10. Ramsay C, Brown E, Hartman G, Davey P. Room for improvement: a systematic review of the quality of evaluations of interventions to improve hospital antibiotic prescribing. J Antimicrob Chemother. 2003;52(5):764-71.

11. Hulscher ME, Laurant MG, Grol RP. Process evaluation on quality improvement interventions. Qual Saf Health Care. 2003;12(1):40-6.

12. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care. 2012;50(3):217-26.

13. Burke JP. Antibiotic resistance--squeezing the balloon? JAMA. 1998;280(14):1270-1.

14. Peterson LR. Squeezing the antibiotic balloon: the impact of antimicrobial classes on emerging resistance. Clin Microbiol Infect. 2005;11 Suppl 5:4-16.

15. Toma M, Davey PG, Marwick CA, Guthrie B. A framework for ensuring a balanced accounting of the impact of antimicrobial stewardship interventions. J Antimicrob Chemother. 2017;72(12):3223-31.

16. McGregor JC, Furuno JP. Optimizing research methods used for the evaluation of antimicrobial stewardship programs. Clin Infect Dis. 2014;59 Suppl 3:S185-92.

17. Gillespie D, Francis NA, Carrol ED, Thomas-Jones E, Butler CC, Hood K. Use of co-primary outcomes for trials of antimicrobial stewardship interventions. Lancet Infect Dis. 2018;18(6):595-7.

18. Baur D, Gladstone BP, Burkert F, Carrara E, Foschi F, Dobele S, et al. Effect of antibiotic stewardship on the incidence of infection and colonisation with antibiotic-resistant bacteria and Clostridium difficile infection:

PART I | CHAPTER 3

OPTIMISING DESIGN OF RESEARCH TO EVALUATE ANTIBIOTIC STEWARDSHIP INTERVENTIONS;
CONSENSUS RECOMMENDATIONS OF A MULTINATIONAL WORKING GROUP

a systematic review and meta-analysis. Lancet Infect Dis. 2017;17(9):990-1001.

19. Davey P, Marwick CA, Scott CL, Charani E, McNeil K, Brown E, et al. Interventions to improve antibiotic prescribing practices for hospital inpatients. Cochrane Database Syst Rev. 2017;2:CD003543.

20. Lennox L, Maher L, Reed J. Navigating the sustainability landscape: a systematic review of sustainability approaches in healthcare. Implment. Sci. 2018;13(1):27.

21. Medical Research Council. Developing and evaluating complex interventions. 2019.

22. Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG, Group C. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. JAMA. 2012;308(24):2594-604.

23. Laptook AR, Shankaran S, Tyson JE, Munoz B, Bell EF, Goldberg RN, et al. Effect of Therapeutic Hypothermia Initiated After 6 Hours of Age on Death or Disability Among Newborns With Hypoxic-Ischemic Encephalopathy: A Randomized Clinical Trial. JAMA. 2017;318(16):1550-60.

24. Mauri L, D'Agostino RB, Sr. Challenges in the Design and Interpretation of Noninferiority Trials. N Engl J Med. 2017;377(14):1357-67.

25. Cranendonk DR, Opmeer BC, Prins JM, Wiersinga WJ. Comparing short to standard duration of antibiotic therapy for patients hospitalized with cellulitis (DANCE): study protocol for a randomized controlled trial. BMC Infect Dis. 2014;14:235.

26. Evans SR, Rubin D, Follmann D, Pennello G, Huskins WC, Powers JH, et al. Desirability of Outcome Ranking (DOOR) and Response Adjusted for Duration of Antibiotic Risk (RADAR). Clin Infect Dis. 2015;61(5):800-6.

27. Phillips PP, Morris TP, Walker AS. DOOR/RADAR: A Gateway Into the Unknown? Clin Infect Dis. 2016;62(6):814-5.

28. Schweitzer VA, van Smeden M, Postma DF, Oosterheert JJ, Bonten MJM, van Werkhoven CH. Response Adjusted for Days of Antibiotic Risk (RADAR): evaluation of a novel method to compare strategies to optimize antibiotic use. Clin Microbiol Infect. 2017;23(12):980-5.

29. Lopez Bernal J, Soumerai S, Gasparrini A. A methodological framework for model selection in interrupted time series studies. J Clin Epidemiol. 2018;103:82-91.

30. (EPOC) CEPaOoC. What study designs can be considered for inclusion in an EPOC review and what should they be called? 2017.

31. Hemming K, Taljaard M, McKenzie JE, Hooper R, Copas A, Thompson JA, et al. Reporting of stepped wedge cluster randomised trials: extension of the CONSORT 2010 statement with explanation and elaboration. BMJ. 2018;363:k1614.

32. Hemming K, Taljaard M. Sample size calculations for stepped wedge and cluster randomised trials: a unified approach. J Clin Epidemiol. 2016;69:137-46.

33. Boel J, Andreasen V, Jarlov JO, Ostergaard C, Gjorup I, Boggild N, et al. Impact of antibiotic restriction on resistance levels of Escherichia coli: a controlled interrupted time series study of a hospital-wide antibiotic stewardship programme. J Antimicrob Chemother. 2016;71(7):2047-51.

34. Taggart LR, Leung E, Muller MP, Matukas LM, Daneman N. Differential outcome of an antimicrobial stewardship audit and feedback program in two intensive care units: a controlled interrupted time series study. BMC Infect Dis. 2015;15:480.

35. Lawes T, Lopez-Lozano JM, Nebot CA, Macartney G, Subbarao-Sharma R, Wares KD, et al. Effect of a national

4C antibiotic stewardship intervention on the clinical and molecular epidemiology of Clostridium difficile infections in a region of Scotland: a non-linear time-series analysis. Lancet Infect Dis. 2017;17(2):194-206.

36. de Kraker MEA, Abbas M, Huttner B, Harbarth S. Good epidemiological practice: a narrative review of appropriate scientific methods to evaluate the impact of antimicrobial stewardship interventions. Clin Microbiol Infect. 2017;23(11):819-25.

37. Littmann J, Rid A, Buyx A. Tackling anti-microbial resistance: ethical framework for rational antibiotic use. Eur J Public Health. 2018;28(2):359-63.

38. Davey P, Peden C, Charani E, Marwick C, Michie S. Time for action-Improving the design and reporting of behaviour change interventions for antimicrobial stewardship in hospitals: Early findings from a systematic review. Int J Antimicrob Agents. 2015;45(3):203-12.

39. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ. 2014;348:g1687.

40. Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D. Segmented regression analysis of interrupted time series studies in medication use research. J Clin Pharm Ther. 2002;27(4):299-309.

41. Tacconelli E, Cataldo MA, Paul M, Leibovici L, Kluytmans J, Schroder W, et al. STROBE-AMS: recommendations to optimise reporting of epidemiological studies on antimicrobial resistance and informing improvement in antimicrobial stewardship. BMJ Open. 2016;6(2):e010134.

42. Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. PLoS Med. 2007;4(10):e297.

43. Ogrinc G, Davies L, Goodman D, Batalden P, Davidoff F, Stevens D. SQUIRE 2.0 (Standards for QUality Improvement Reporting Excellence): revised publication guidelines from a detailed consensus process. BMJ Qual Saf. 2016;25(12):986-92.

44. Des Jarlais DC, Lyles C, Crepaz N, Group T. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Public Health. 2004;94(3):361-6.

4

RESPONSE ADJUSTED FOR DAYS OF ANTIBIOTIC RISK (RADAR): EVALUATION OF A NOVEL METHOD TO COMPARE STRATEGIES TO OPTIMISE ANTIBIOTIC USE

Valentijn A. Schweitzer, Maarten van Smeden, Douwe F. Postma, Jan J. Oosterheert, Marc J.M. Bonten, Cornelis H. van Werkhoven

PART I | CHAPTER 4

RESPONSE ADJUSTED FOR DAYS OF ANTIBIOTIC RISK (RADAR): EVALUATION OF A NOVEL METHOD TO COMPARE
STRATEGIES TO OPTIMISE ANTIBIOTIC USE

ABSTRACT

OBJECTIVES

The Response Adjusted for Days of Antibiotic Risk (RADAR)-statistic was proposed to improve efficiency of trials comparing antibiotic stewardship strategies to optimize antibiotic use. We studied the behaviour of RADAR in a non-inferiority trial in which a beta-lactam monotherapy strategy (BL, n=656) was non-inferior to fluoroquinolone monotherapy (FQL, n=888) for moderately-severe community-acquired pneumonia (CAP) patients.

METHODS

Patients were ranked according to clinical outcome, using five or eight categories, and antibiotic use. RADAR was calculated as the probability that the BL group had a more favourable ranking than the FQL group. To investigate the sensitivity of RADAR to detrimental clinical outcome we simulated increasing rates of 90-day mortality in the BL group and performed the RADAR and non-inferiority analysis.

RESULTS

The RADAR of the BL-group compared to the FQL group was 60.3% (95% confidence interval 57.9%-62.7%) using five and 58.4% (95% CI 56.0%-60.9%) using eight clinical outcome categories, all in favour of BL. Sample sizes for RADAR were 38% (250/653) and 89% (580/653) of the non-inferiority sample size calculation, using five or eight clinical outcome categories respectively. With simulated mortality rates, loss of non-inferiority of the BL-group occurred at a relative risk of 1.125 in the conventional analysis, whereas using RADAR the BL-group lost superiority at a relative risk of mortality of 1.25 and 1.5, with eight and five clinical outcome categories, respectively.

CONCLUSIONS

RADAR favoured BL over FQL therapy for CAP. Although RADAR required fewer patients than conventional non-inferiority analysis, the statistic was less sensitive to detrimental outcomes.

INTRODUCTION

Antibiotic resistance is increasing and associated with prolonged hospital stay and increased mortality and healthcare costs[1,2]. Selection of antibiotic resistant bacteria is facilitated through use of antibiotics[3]. Previous studies found high rates of inappropriate antibiotic use in different clinical settings[4–6]. With antibiotic stewardship programs, physicians aim to ensure therapeutic efficacy while limiting adverse events of antibiotic overuse, such as the emergence of resistance, adverse drug events and costs[3,7]. In general, trials to evaluate antibiotic stewardship interventions are designed to show an increase in appropriate antibiotic use without compromising clinical outcome, frequently using a non-inferiority design. However, non-inferiority trials may suffer from analysis of suboptimal or subjective clinical outcomes, such as "clinical cure", and are frequently underpowered to demonstrate non-inferiority for objective clinical outcomes such as mortality[8]. Also, comparing benefit and harm can be difficult, for example when an intervention decreases antibiotic use at the cost of more complications. Finally, non-inferiority trials usually require large numbers of subjects and choosing the optimal non-inferiority margin may be subjective and may lead to discussion after completion of the trial[9,10].

In search for a better method to compare strategies to optimize antibiotic use, which is often pursued by antibiotic stewardship interventions, the Response Adjusted for Days of Antibiotic Risk (RADAR) statistic was proposed[11]. For its computation, patients are first classified based on mutually exclusive, hierarchical levels corresponding to clinical outcome of the patient, e.g. complication free survival, survival with complications and mortality. Within each clinical outcome category, patients are subcategorized according to their level of antibiotic use. All patients are then ranked according to their category, where patients with a better clinical outcome, receiving less antibiotics or antibiotics with a narrower antimicrobial spectrum have a more favourable rank. Theoretically, through this ordinal classification, patients with a worse clinical outcome always have a lower ranking, while within outcome categories, appropriateness of antibiotic prescription determines the ranking. In the analysis, the distributions of the rankings pre and post antibiotic stewardship intervention are compared by combining clinical outcomes and antibiotic use, RADAR allows to analyse antibiotic stewardship interventions as superiority instead of non-inferiority trials, thereby requiring a smaller sample size[11,12]. However, RADAR also may have disadvantages, for example: choosing the components of the hierarchical clinical categories is subjective, RADAR results in a percentage which is difficult to interpret and it is uncertain whether clinical safety can be demonstrated with the reduced sample size[13,14]. Therefore, RADAR needs to be evaluated with real-life clinical trial data. For this purpose, we used data from a non-inferiority trial of empirical

PART I | CHAPTER 4

RESPONSE ADJUSTED FOR DAYS OF ANTIBIOTIC RISK (RADAR): EVALUATION OF A NOVEL METHOD TO COMPARE
STRATEGIES TO OPTIMISE ANTIBIOTIC USE

antibiotic treatment strategies in patients with community-acquired pneumonia (CAP) admitted to non-intensive care unit (non-ICU) wards[15]. We study the application of RADAR in comparison to the conventional non-inferiority analysis to determine the sensitivity to choices in the analysis by (1) calculating RADAR with different clinical outcome categories and different levels of antibiotic use, (2) quantifying the effect of sample size on the certainty of the outcome with both methods, and (3) determining the sensitivity of RADAR to lose superiority and the conventional analysis to lose non-inferiority to increasing worse clinical outcomes by simulations.

METHODS

DATA COLLECTION

The Community-Acquired Pneumonia — Study on the Initial Treatment with Antibiotics of Lower Respiratory Tract Infections (CAP-START) was performed between February 2011 and August 2013 in 7 hospitals in the Netherlands[15,16]. In the CAP-START trial, patients above 18 years of age who were admitted to a non-ICU ward for suspicion of CAP were eligible for study participation. CAP-START was a cluster-randomised cross-over trial. Hospitals participating in the trial were assigned to beta-lactam monotherapy (BL), beta-lactam with a macrolide (BLM), or fluoroquinolone monotherapy (FQL) as preferred empiric treatment strategies for a 4 months period in randomised order without washout periods. Physicians in the participating hospitals were repeatedly reminded of the current antibiotic strategy by local investigators with the use of newsletters and presentations to ensure strategy adherence. Deviation from the assigned treatment or subsequent change was allowed when medically indicated. Depending on the study arm, adherence to the strategy varied between 70 to 80%. The primary outcome of the CAP-START trial was that an empiric treatment strategy of BL was non-inferior to BLM and FQL on 90-day mortality. The study was approved by the ethics review board of the at the University Medical Center Utrecht (reference number 10/148). Written informed consent for data collection was obtained within 72 hours after admission.

CLINICAL OUTCOME RANKINGS

For the current post-hoc analysis we compared the BL group to the FQL group, considering the latter as the control group with a high proportion of patients receiving antibiotics covering atypical pathogens. Implementing BL was considered the antibiotic stewardship intervention aiming to reduce the use of atypical coverage. Clinical outcome categories were constructed, in accordance with RADAR recommendations, with either two, five, or eight mutually exclusive hierarchical levels for clinical outcome (Table 1)[11]. RADAR is not intended to be used with a simple binary clinical outcome ranking[11].

Table 1. The constructed clinical outcome categories based on two, five and eight mutually exclusive hierarchical levels and corresponding ranks.

Ranking	Clinical outcome categories		
	Two	Five	Eight
Rank 1	- Survival (day 90)	- Survival (day 90) - Discharge <7 days - No ICU admission - No septic shock	- Survival (day 90) - Discharge <7 days - No ICU admission - No septic shock - No gastro-intestinal side-effects - No cardiovascular complications
Rank 2	**- Death (90 day)**	- Survival (day 90) **- Discharge ≥7 days** - No ICU admission - No septic shock	- Survival (day 90) - Discharge <7 days - No ICU admission - No septic shock **- Gastro-intestinal side-effects** - No cardiovascular complications
Rank 3	-	- Survival (day 90) - Any discharge day **- ICU admission** - No septic shock	- Survival (day 90) **- Discharge ≥7 days** - No ICU admission - No septic shock - No gastro-intestinal side-effects - No cardiovascular complications
Rank 4	-	- Survival (day 90) - Any discharge day - ICU admission **- Septic shock**	- Survival (day 90) **- Discharge ≥7 days** - No ICU admission - No septic shock - Gastro-intestinal side-effects - No cardiovascular complications
Rank 5	-	**- Death (day 90)** - Any discharge day - Yes/no ICU admission - Yes/no septic shock	- Survival (day 90) - Any discharge day **- No ICU admission** - No septic shock - Yes/no gastro-intestinal side-effects **- Cardiovascular complications**
Rank 6	-	-	- Survival (day 90) - Any discharge day **- ICU admission** - No septic shock - Yes/no gastro-intestinal side-effects - Yes/no cardiovascular complications
Rank 7	-	-	- Survival (day 90) - Any discharge day - ICU admission **- Septic shock** - Yes/no gastro-intestinal side-effects - Yes/no Cardiovascular complications
Rank 8	-	-	**- Death (day 90)** - Any discharge day - Yes/no ICU admission - Yes/no septic shock - Yes/no gastro-intestinal side-effects - Yes/no Cardiovascular complications

We only use the 2-category RADAR to illustrate the behaviour of RADAR in response to the amount of clinical categories used. Categories of antibiotic use were defined either as receipt of any atypical coverage (less favourable ranking) at any time during

PART I | CHAPTER 4

RESPONSE ADJUSTED FOR DAYS OF ANTIBIOTIC RISK (RADAR): EVALUATION OF A NOVEL METHOD TO COMPARE STRATEGIES TO OPTIMISE ANTIBIOTIC USE

hospitalization versus no atypical coverage (more favourable ranking), or as the number of in-hospital days on atypical coverage (more days equals less favourable ranking). Atypical coverage was defined as antibiotic treatment with azithromycin, erythromycin, clarithromycin, moxifloxacin, levofloxacin, doxycyclin or ciprofloxacin. RADAR is a rank-order statistic which reflects the probability that a randomly selected patient assigned to the intervention group (here the BL group) has a more favourable ranking compared to a randomly selected patient from the control group (here the FQL group)[11]. A RADAR-statistic of 50% implies that the BL and FQL group are equally ranked, <50% indicates that either the clinical outcome is worse and/or the antibiotic use is higher (more atypical coverage) in the BL group and >50% indicates that either the clinical outcome is better and/or the antibiotic use is lower (less atypical coverage) in the BL group. Two thousand bootstrap samples were generated to estimate 95% confidence intervals for the RADAR-statistic[11,17]. Statistical significance was declared when the confidence interval did not overlap a predefined clinically relevant value (55% as proposed by Evans et al.). For simplicity of illustration, cluster effects arising from the cluster-randomized design of the study were ignored in the calculation of RADAR; the intra-cluster correlation for 90-day mortality was estimated to be 4.5E-7 and is therefore negligible[15]. All analyses were performed using R software, version 3.0.2 (R Project for Statistical Computing)[18].

EFFECT OF SAMPLE SIZE ON NON-INFERIORITY ANALYSIS FOR MORTALITY

The CAP-START trial was designed to demonstrate non-inferiority on day-90 mortality, with a predicted mortality rate of 5%, a non-inferiority margin of 3%, a one-sided alpha of 0.05, a power of 80%, and negligible intra-cluster correlation, yielding a required sample size of 650 per study arm[15],[16]. The observed day-90 mortality rate was 10%. When re-estimating the sample size assuming 10% mortality, the required sample size is 1126 per study arm. The required sample size for RADAR was calculated to demonstrate a probability of at least 55% that a patient assigned to the BL group had a more favourable ranking than a patient assigned to the FQL group with a two-sided alpha of 0.05 and a power of 80%. The RADAR sample size was calculated using a Wilcoxon-Mann-Whitney test both with the predicted mortality of 5% and the observed mortality of 10%. For the sample size calculation using the observed mortality, the RADAR rank distributions were determined with the clinical outcomes and antibiotic use as observed in the CAP-START trial (Table 2). For the sample size calculation with the predicted mortality, we assumed that the distribution of remaining clinical outcomes and antibiotic use of the patients was the same as observed. Sample size calculations were confirmed with power simulations (results not shown).

To assess the impact of reduced RADAR sample size on non-inferiority of clinical outcome, we analysed the risk difference of 90-day mortality and corresponding 90%

confidence intervals (one sided test with alpha 0.05) if the amount of patients required for the RADAR analysis were included in the trial.

Table 2. Distribution of rankings* in the BL period compared to the FQL period with either two or five clinical outcome categories.

Rank	Any atypical coverage	Two clinical outcome categories		Five clinical outcome categories		Eight clinical outcome categories	
		BL N(%)	FQL N(%)	BL N(%)	FQL N(%)	BL N(%)	FQL N (%)
1	-	366 (56.0)	80 (9.0)	232 (35.5)	48 (5.4)	220 (33.6)	44 (5.0)
	+	229 (35.0)	729 (82.2)	116 (17.7)	445 (50.2)	109 (16.7)	423 (47.7)
2	-	36 (5.5)	12 (1.4)	134 (20.5)	32 (3.6)	8 (1.2)	3 (0.3)
	+	23 (3.5)	66 (7.4)	104 (15.9)	275 (31.0)	6 (0.9)	13 (1.5)
3	-	-	-	0 (0.0)	0 (0.0)	106 (16.2)	26 (2.9)
	+	-	-	6 (0.9)	8 (0.9)	87 (13.3)	234 (26.4)
4	-	-	-	0 (0.0)	0 (0.0)	8 (1.2)	4 (0.5)
	+	-	-	3 (0.5)	1 (0.1)	6 (0.9)	8 (0.9)
5	-	-	-	36 (5.5)	12 (1.4)	24 (3.7)	3 (0.3)
	+	-	-	23 (3.5)	66 (7.4)	12 (1.8)	27 (3.0)
6	-	-	-			0 (0.0)	0 (0.0)
	+	-	-			6 (0.9)	8 (0.9)
7	-	-	-			0 (0.0)	0 (0.0)
	+	-	-			3 (0.5)	1 (0.1)
8	-	-	-			36 (5.5)	12 (1.4)
	+	-	-			23 (3.5)	66 (7.4)

* The definitions of the clinical outcome categories are shown in table 1.

SENSITIVITY OF RADAR-STATISTIC AND THE NON-INFERIORITY ANALYSIS FOR CLINICAL OUTCOME

The sensitivity of RADAR and the non-inferiority analysis for clinical outcome was tested by simulating increased mortality in the BL-group. In each simulation, a random selection of patients from the BL-group were reclassified to the "death within 90 days" clinical outcome category. Simulations were repeated 2000 times. The simulated RADAR for five and eight clinical categories, and the mortality risk differences between BL and FQL were plotted against the simulated relative risk of mortality. To determine the sensitivity of the statistics for clinical outcome, we used the simulated relative risk at which superiority (for RADAR) or non-inferiority (for the conventional analysis) was lost.

PART I | CHAPTER 4

RESPONSE ADJUSTED FOR DAYS OF ANTIBIOTIC RISK (RADAR): EVALUATION OF A NOVEL METHOD TO COMPARE
STRATEGIES TO OPTIMISE ANTIBIOTIC USE

RESULTS

RADAR OUTCOME OF THE CAP-START TRIAL

From the 2,283 patients included in the CAP-START study, 656 were assigned to BL and 888 to FQL. When defining antibiotic use as 'any atypical coverage', the RADAR of the BL-group compared to the FQL group was 71.5% (95% confidence interval 69.2%-73.7%) using two clinical outcome categories, 60.3% (95% confidence interval 57.9%-62.7%) using five clinical outcome categories, and 58.4% (95% confidence interval 56.0%-60.9%) using eight clinical outcome categories, all in favour of BL (Figure 1). The RADAR-statistics were comparable when using 'days of atypical coverage' (Figure 1).
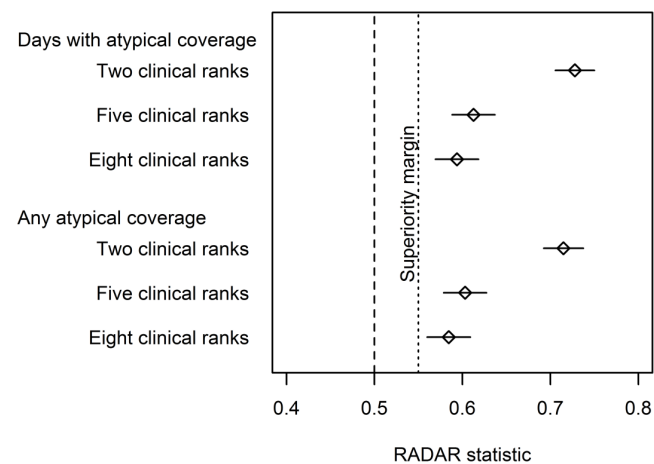
Figure 1. RADAR outcome of the CAP-START trial with two, five or eight clinical outcome categories and with antibiotic use defined as either any atypical coverage during hospitalization or days of atypical coverage received.

EFFECT OF SAMPLE SIZE ON NON-INFERIORITY ANALYSIS OF MORTALITY

The calculated required sample size for RADAR using five clinical categories to demonstrate superiority with the lower bound of the confidence interval above 55% was 250 patients per study arm, i.e. 38% of the non-inferiority sample size of the original CAP-START study using the predicted 5% mortality or 360 patients per study arm, i.e. 32% of the non-inferiority sample size using the observed 10% mortality and a non-inferiority margin of 3% (Table 3). The required sample size for the RADAR-analysis using eight clinical categories was 580 per study arm, i.e. 89% of the non-inferiority sample size using the predicted 5% mortality or 875 per study arm, i.e. 78% of the non-inferiority sample size using the observed 10% mortality. Based on the RADAR sample size calculation with five clinical outcome categories and the predicted 5% mortality, 250 patients per study arm, or 500 consecutive patients enrolled which would yield a risk difference in 90-day mortality of -0.21% (-4.1% to 3.6%, 90% confidence interval). The required sample size with eight clinical outcome categories would correspond to 580 patients per study arm, which would result in a risk difference of -0.06% (-2.6% to 2.4%, 90% confidence interval). The corresponding power to demonstrate non-inferiority would be 45.7% when using five clinical outcome categories and 75.8% using eight clinical outcome categories. This demonstrates that, when using five clinical outcome categories, RADAR would allow a marked reductions in sample size but the method would have insufficient power to demonstrate non-inferiority on 90-day mortality. In contrast, when using eight clinical outcome categories, RADAR requires a similar sample size and consequently would have similar power to demonstrate non-inferiority on 90-day mortality.

Table 3. Estimated required sample size per study arm for RADAR and the non-inferiority design

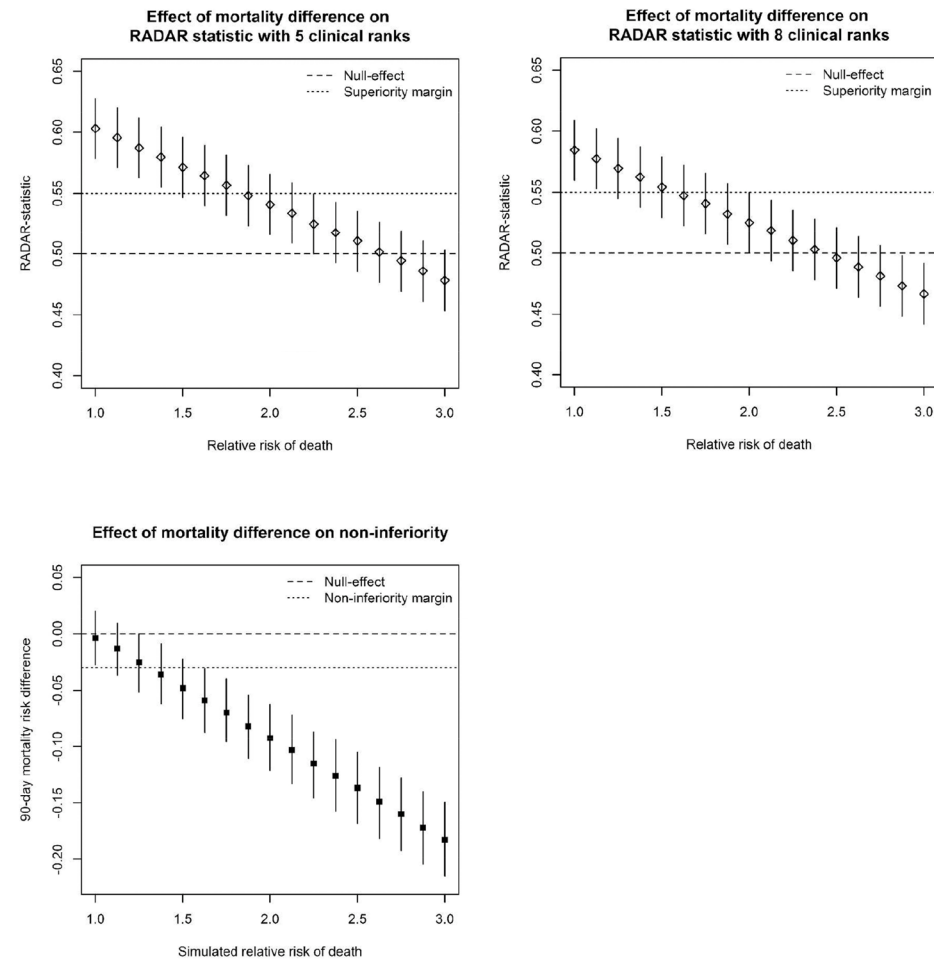|	RADAR		Non-inferiority
	5 clinical outcomes	8 clinical outcomes	
Predicted clinical outcome/observed antibiotic use	250	580	653
Observed clinical outcome/ observed antibiotic use	360	875	1126

SENSITIVITY OF RADAR-STATISTIC AND THE NON-INFERIORITY ANALYSIS FOR CLINICAL OUTCOME

By simulating increased mortality in the BL-group, RADAR remained statistically significant above the 55% threshold up to an increased relative risk of death of 1.5 in the analysis with five clinical outcome categories and up to an increased relative risk of death of 1.25 in the analysis with eight clinical outcome categories (Figure 2). Thus, in spite of this marked increase in mortality in the BL group, RADAR continued to show a favourable outcome for BL. In contrast, the conventional non-inferiority approach already failed to demonstrate non-inferiority after the first step increase of the relative risk of death at a level of 1.125.

PART I | CHAPTER 4

RESPONSE ADJUSTED FOR DAYS OF ANTIBIOTIC RISK (RADAR): EVALUATION OF A NOVEL METHOD TO COMPARE
STRATEGIES TO OPTIMISE ANTIBIOTIC USE

Figure 2. Relation between a simulated increase in mortality and the RADAR-statistic or non-inferiority. The definition of the five and eight clinical outcome categories are defined in table 1.

DISCUSSION

Application of the recently proposed RADAR on data from a multicentre cluster randomized trial, designed to demonstrate non-inferiority of a strategy of BL for 90-day all-cause mortality in patients with moderate-severe CAP, revealed a significantly more favourable ranking in the BL group compared to the FQL group. In our study, the value of RADAR was shown to be strongly influenced by the number of clinical outcome categories used but not to the number of categories for antibiotic use. Despite the favourable RADAR outcome, we showed that the obtained value of RADAR varies with the number of clinical outcome categories chosen in the design phase. The performed

sample size calculations show that RADAR may require fewer than the conventional non-inferiority analysis. However, RADAR and the reduced sample size showed to be insensitive to detecting worse clinical outcomes.

The positive RADAR confirm the absence of differences in clinical outcomes with substantially reduced atypical coverage when using the BL strategy[15]. However, we consider interpretation of RADAR as difficult. While we observed no differences in clinical outcome in the CAP-START trial, the statistic was substantially higher using two clinical categories as compared to five or eight clinical categories. Furthermore, a RADAR-statistic >50% should indicate that the clinical outcome is better and/or that antibiotic use has improved (i.e. shorter duration of antibiotic use, less broad-spectrum antibiotic use, depending on the definition of antibiotic exposure chosen for RADAR). However, the simulations with increasing mortality rates in the BL strategy revealed that positive RADAR-statistics can still be found in the setting of a worse clinical outcome. In addition, RADAR represents the probability of more favourable ranking, without quantification of intervention effectiveness.

Our findings demonstrate that the construction of the clinical outcome categories with the hierarchical levels needs to be considered carefully, as it influences both the sample size calculation and the number of predefined clinical outcome categories directly determined the value of RADAR. The latter can be explained by a reduced contribution of antibiotic use to RADAR when more hierarchical levels are used. In the extreme case, if the resolution of clinical outcome is infinite, every patient would get its own clinical outcome category, negating the effect of antibiotic use completely, as this only affects patients that are at the same clinical outcome category. As a result, the value of RADAR can be influenced by choosing the amount of clinical outcome categories. In addition, there is large variability in the definition of antibiotic use. This makes comparison of RADAR-statistics between studies only possible if the same ranking categories and definition of antibiotic use have been used. Also, it may be difficult to identify the most appropriate outcome categories beforehand. Incorrectly adding or missing clinical outcomes has direct consequences for the primary endpoint of the study through the number of clinical outcome categories used.

An attractive argument to adopt RADAR would be a required lower sample size[11]. When comparing the calculated required sample size of RADAR to the original sample size required for the classical non-inferiority design of CAP-START, there was a 61% reduction in required sample size when using five clinical outcome categories but only a 11% reduction when using eight clinical outcome categories. However, as RADAR is a composite endpoint, separate analysis of clinical outcomes is still required to

PART I | CHAPTER 4

RESPONSE ADJUSTED FOR DAYS OF ANTIBIOTIC RISK (RADAR): EVALUATION OF A NOVEL METHOD TO COMPARE
STRATEGIES TO OPTIMISE ANTIBIOTIC USE

ensure safety of the reduction in antibiotic use, and a lower sample size results in less confidence of safety to the intervention for clinical outcome[13]. This was illustrated by the simulated trial analysis with the lower RADAR sample size, that did not allow the conclusion of non-inferiority for 90-day mortality when the RADAR method required a marked lower sample size, as was the case when using five clinical outcome categories. We, therefore, concur with others that RADAR should not be used as the main criterion to determine the required sample size[13,14].

Strengths of the current analysis include the application of the CAP-START trial, constituting real-life trial data, to evaluate RADAR. The design of the CAP-START trial is very suitable to evaluate the RADAR method, as the randomized antibiotic strategies allowed for deviation. This resembles daily practice of antibiotic stewardship interventions where deviations from a proposed strategy are frequently justified. Also, the sample size calculations were performed using both the estimated and the actual clinical outcomes and antibiotic use as found in the CAP-START study. However, the RADAR sample size calculation was performed using only observed non-mortality outcome data. The RADAR sample size using predicted outcomes might have been different due to uncertainty in expected outcomes. Other strengths include the completeness of data collection and the sample size allowing comparison between a conventional non-inferiority and RADAR.

A limitation of the current analysis is that we did not adjust for clustering in the sample size calculations and in the calculation of RADAR. Due to the cross-over design, the intra-cluster correlation of 90-day mortality was negligible in this study, therefore, the impact of not adjusting for clustering was expected to be minimal[15]. However, the intra-cluster correlation of antibiotic use could be higher and taking this into account might have resulted in a higher required sample size for the RADAR method. Another limitation is that clinical outcome categories were defined post-hoc, as this was a post-hoc analysis. However, different outcome categories would not have influenced the conclusion regarding sensitivity to the number of clinical outcome categories and detrimental outcome.

To conclude, we present the application of the novel RADAR methodology to compare strategies to optimize antibiotic use on a previously conducted non-inferiority trial. RADAR inventively combines clinical outcome with antibiotic use as a composite endpoint on a patient level. However, we believe that it should not replace non-inferiority of patient-relevant clinical endpoints as the primary analysis. Therefore, we recommend that mortality is used for sample size calculations (as this endpoint requires the largest number of patients) and that mortality and an appropriate antibiotic use outcome are reported as the two primary endpoints in antibiotic stewardship trials.

PART I | CHAPTER 4

RESPONSE ADJUSTED FOR DAYS OF ANTIBIOTIC RISK (RADAR): EVALUATION OF A NOVEL METHOD TO COMPARE
STRATEGIES TO OPTIMISE ANTIBIOTIC USE

REFERENCES

1. European Centre for Disease Prevention and Control. Annual epidemiological report 2014. Antimicrobial resistance and healthcare-associated infections 2014:28.

2. Rottier WC, Ammerlaan HSM, Bonten MJM. Effects of confounders and intermediates on the association of bacteraemia caused by extended-spectrum -lactamase-producing enterobacteriaceed and patient outcome: A meta-analysis. J Antimicrob Chemother 2012;67:1311–20. doi:10.1093/jac/dks065.

3. Bell BG, Schellevis F, Stobberingh E, Goossens H, Pringle M. A systematic review and meta-analysis of the effects of antibiotic consumption on antibiotic resistance. BMC Infect Dis 2014;14:13. doi:10.1186/1471-2334-14-13.

4. Fitzpatrick JM, Biswas JS, Edgeworth JD, et al. Gram-negative bacteraemia; A multi-centre prospective evaluation of empiric antibiotic therapy and outcome in English acute hospitals. Clin Microbiol Infect 2016;22:244–51. doi:10.1016/j.cmi.2015.10.034.

5. Jones BE, Jones MM, Huttner B, et al. Trends in antibiotic use and nosocomial pathogens in hospitalized veterans with pneumonia at 128 medical centers, 2006-2010. Clin Infect Dis 2015;61:1403–10. doi:10.1093/cid/civ629.

6. Dekker ARJ, Verheij TJM, van der Velden AW. Inappropriate antibiotic prescription for respiratory tract indications: Most prominent in adult patients. Fam Pract 2015;32:401–7. doi:10.1093/fampra/cmv019.

7. Dellit TH, Owens RC, McGowan JE, et al. Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America Guidelines for Developing an Institutional Program to Enhance Antimicrobial Stewardship. Clin Infect Dis 2007;44:159–77. doi:10.1097/IPC.0b013e318068b1c0.

8. Harbarth S, von Dach E, Pagani L, et al. Randomized non-inferiority trial to compare trimethoprim/ sulfamethoxazole plus rifampicin versus linezolid for the treatment of MRSA infection. J Antimicrob Chemother 2015;70:264–72. doi:10.1093/jac/dku352.

9. Chetchotisakd P, Chierakul W, Chaowagul W, et al. Trimethoprim-sulfamethoxazole versus trimethoprimsulfamethoxazole plus doxycycline as oral eradicative treatment for melioidosis (MERTH): A multicentre, double-blind, non-inferiority, randomised controlled trial. Lancet 2014;383:807–14. doi:10.1016/S0140-6736(13)61951-0.

10. Nie W, Li B. Antibiotics for Community-Acquired Pneumonia in Adults. N Engl J Med 2015;373:683–6. doi:10.1056/NEJMc1506892#SA4.

11. Evans SR, Rubin D, Follmann D, et al. Desirability of outcome ranking (DOOR) and response adjusted for duration of antibiotic risk (RADAR). Clin Infect Dis 2015;61:800–6. doi:10.1093/cid/civ495.

12. Wang SJ, James Hung HM, Tsong Y, Cui L. Group sequential test strategies for superiority and non-inferiority hypotheses in active conrolled clinical trials. Stat Med 2001;20:1903–12. doi:10.1002/sim.820.

13. Molina J, Cisneros JM. A chance to change the paradigm of outcome assessment of antimicrobial stewardship programs. Clin Infect Dis 2015;61:807–8. doi:10.1093/cid/civ496.

14. Phillips PPJ, Morris TP, Walker AS. DOOR/RADAR: A Gateway into the Unknown? Clin Infect Dis 2015;62:814–5. doi:10.1093/cid/civ1002.

15. Postma DF, van Wekhoven CH, van Elden LJ, et al. Antibiotic Treatment Strategies for Community-Acquired Pneumonia in Adults. NEJM 2015;372;14:1312–23. doi:10.1056/NEJMoa1406330.

16. Van Werkhoven CH, Postma DF, Oosterheert JJ, Bonten MJM. Antibiotic treatment of moderate-severe community-acquired pneumonia: Design and rationale of a multicentre cluster-randomised cross-over trial. Neth J Med 2014;72:170–8.

17. Efron B. Bootstrap Methods: Another Look at the Jackknife. Ann Stat 1979;7:1–26. doi:10.1214/aos/1176344552.

18. R foundation. The R project for statistical computing (http://www.r-project.org).

5

CONFOUNDING BY INDICATION OF THE SAFETY OF DE-ESCALATION IN COMMUNITY-ACQUIRED PNEUMONIA: A SIMULATION STUDY EMBEDDED IN A PROSPECTIVE COHORT.

Inger van Heijl, Valentijn A. Schweitzer, C.H. Edwin Boel, Jan J. Oosterheert, Susanne M. Huijts, Wendelien Dorigo-Zetsma, Paul D. van der Linden, Marc J.M. Bonten, Cornelis H. van Werkhoven

PART II | CHAPTER 5

CONFOUNDING BY INDICATION OF THE SAFETY OF DE-ESCALATION IN COMMUNITY-ACQUIRED PNEUMONIA:
A SIMULATION STUDY EMBEDDED IN A PROSPECTIVE COHORT.

ABSTRACT

Observational studies have demonstrated that de-escalation of antimicrobial therapy is independently associated with lower mortality. This most probably results from confounding by indication. Reaching clinical stability is associated with the decision to de-escalate and with survival. However, studies rarely adjust for this confounder. We quantified the potential confounding effect of clinical stability on the estimated impact of de-escalation on mortality in patients with community-acquired pneumonia. Data were used from the Community-Acquired Pneumonia immunization Trial in Adults (CAPiTA). The primary outcome was 30-day mortality. We performed Cox proportional-hazards regression with de-escalation as time-dependent variable and adjusted for baseline characteristics using propensity scores. The potential impact of unmeasured confounding was quantified through simulating a variable representing clinical stability on day three, using data on prevalence and associations with mortality from the literature. Of 1,536 included patients, 257 (16.7%) were de-escalated, 123 (8.0%) were escalated and in 1156 (75.3%) the antibiotic spectrum remained unchanged. Crude 30-day mortality was 3.5% (9/257) and 10.9% (107/986) in the de-escalation and continuation groups, respectively. The adjusted hazard ratio of de-escalation for 30-day mortality (compared to patients with unchanged coverage), without adjustment for clinical stability, was 0.39 (95%CI: 0.19-0.79). If 90% to 100% of de-escalated patients were clinically stable on day three, the fully adjusted hazard ratio would be 0.56 (95%CI: 0.27-1.12) to 1.04 (95%CI: 0.49-2.23), respectively. The simulated confounder was substantially stronger than any of the baseline confounders in our dataset. Quantification of effects of de-escalation on patient outcomes without proper adjustment for clinical stability results in strong negative bias. This study suggests the effect of de-escalation on mortality needs further well-designed prospective research to determine effect size more accurately.

INTRODUCTION

The aim of antimicrobial stewardship is improving antibiotic use, without compromising clinical outcomes on the individual level[1]. De-escalation of empirical antimicrobial therapy is highly recommended in antimicrobial stewardship programs. In a recent systematic review de-escalation of empirical antimicrobial therapy was associated with a 56% (95% CI 34%-70%) relative risk reduction in mortality[2]. Although it seems a safe strategy, most studies evaluating de-escalation and reporting mortality were observational with a high risk of bias, high clinical heterogeneity and not sufficiently powered to demonstrate safety for mortality. To the best of our knowledge, there are two randomized trials evaluating de-escalation, and these trials did not show a survival benefit for de-escalation[3, 4]. A possible physiological mechanism for decreased mortality due to de-escalation could be a result of a more effective strategy by narrow-spectrum antibiotics or in case of continuation of unnecessary broad-spectrum antibiotics due to more (severe) side-effects. However, it seems highly unlikely that this would lead to increased mortality in the population. Therefore, the association between de-escalation and improved survival in observational studies is most likely biased by unmeasured confounding by indication. Confounding by indication is present if the indication for the intervention (here: de-escalation of empirical antimicrobial therapy) is also a prognostic factor for the outcome (mortality). De-escalation is usually only performed when clinical stability is reached in the first days after starting antimicrobial therapy and this also is a strong prognostic factor for patient outcome. However, hardly any of the observational studies adjusts for clinical stability during admission. In the aforementioned systematic review[2] only one of nineteen observational studies corrected for this confounder[5]. Potentially they did not consider this to be an important confounder, or they lacked data on clinical stability during admission. Not taking this into account causes a negative bias (towards a protective effect). However, the magnitude of this bias has never been established. The aim of the current study was to quantify the potential effect of unmeasured confounding by indication due to clinical stability in the association between de-escalation and patient outcome in patients with community-acquired pneumonia.

MATERIALS AND METHODS

DATA COLLECTION

Data were used from the Community-Acquired Pneumonia immunization Trial in Adults (CAPiTA)[6]. This study was a parallel-group, randomized, placebo-controlled, double blind trial to assess the efficacy of a 13-valent pneumococcal conjugate vaccine. The study included 84,496 immunocompetent community-dwelling adults, 65 years of age

PART II | CHAPTER 5

CONFOUNDING BY INDICATION OF THE SAFETY OF DE-ESCALATION IN COMMUNITY-ACQUIRED PNEUMONIA:
A SIMULATION STUDY EMBEDDED IN A PROSPECTIVE COHORT.

and above. Surveillance for suspected pneumonia was performed in 58 hospitals in the Netherlands, in the period September 2008 - August 2013. The study was approved by the Central Committee on Research Involving Human Subjects and by the Ministry of Health, Welfare and Sport in the Netherlands and all the participants provided written informed consent. For the current analysis, patients receiving antibiotics on the day of admission and with a working diagnosis of CAP admitted to a non-intensive care unit (ICU) were included. We think the effect of de-escalation on mortality in the ICU population is different than in non-ICU population and including these patients will result in a more heterogeneous population. Moreover, factors such as culture results and clinical stability may play a very different role in that population. Patients were excluded from the current analysis if they participated in a simultaneously running interventional trial evaluating different antibiotic regimens for CAP[7], since this trial interfered with the choice of empirical antibiotic treatment, or if they died within 24 hours of admission because these are not eligible for de-escalation.

DEFINITIONS

To define de-escalation, antibiotics were ranked based on their spectrum of activity against CAP pathogens, from rank 1 ('narrow-spectrum') to rank 3 ('extended / restricted spectrum') antibiotics (Table 1). The ranking was performed by a team of experts: two clinical microbiologists (CHEB, MJMB), one infectious diseases specialist (JJO), two clinical pharmacists (IvH, PDvdL) and one epidemiologist (CHvW). In the Dutch setting, penicillin and amoxicillin are in general classified as narrow-spectrum antibiotics. For mild CAP in primary care and moderate-severe CAP (non-ICU ward) these antibiotics are first choice treatment with tetracyclines as an alternative in case of allergies[8]. Sweden and Denmark have similar policies[9, 10]. These antibiotics were classified as rank 1. Antibiotics with a 'restricted' label, advised by the national guide for antibiotic stewardship teams were classified as rank 3[11]. All other regimens were classified as rank 2. In patients with combination therapy, the highest rank of any individual antibiotic was counted, except for combination therapy of β-lactam therapy and a macrolide, which was considered as rank 3, as for respiratory pathogens this combination results in a much broader spectrum than any of the individual antibiotics. Therapy adjustment was defined as the first switch from empirical therapy to another antimicrobial class during hospitalization, independent of the reason for switching. De-escalation and escalation were defined as a change to a lower rank or a higher rank, respectively. Continued regimens or adjustments to an equivalent rank were defined as continuation.

Table 1. Antibiotic ranking.

| Rank 1 (Narrow spectrum)	Rank 2 (Broad spectrum)	Rank 3 (Extended / restricted spectrum)
Penicillin	1st generation cephalosporins	3d generation cephalosporins
Amoxicillin	2nd generation cephalosporins	4th generation cephalosporins
Tetracyclines	Co-amoxi-clav	Fluoroquinolones
	Co-trimoxazole	Aminoglycosides
	Clindamycine	Piperacillin/tazobactam
	Macrolides	Carbapenems
		Vancomycin

STATISTICAL ANALYSIS

Descriptive statistics were used to describe clinical practice of de-escalation. Differences in patient characteristics between patients with a de-escalation versus no de-escalation were compared using Student's t test or χ2 tests. Frequencies of de-escalation, escalation and continuation were described visually and numerically. We tested the proportional hazard assumptions for a follow-up period of 90 days, which revealed that the hazards were proportional up to 30 days and not thereafter (see Fig 1). Therefore we used 30-day mortality as the outcome. To determine the effect of de-escalation on clinical outcome we excluded patients starting in rank 1, since they are not able to de-escalate. We performed Cox proportional hazards regression with de-escalation as time-dependent variable and adjusted for baseline characteristics using propensity score analyses. Propensity scores were calculated from a logistic regression model to estimate a patients propensity for de-escalation and included the variables: age, gender, smoking status, history of diabetes mellitus, history of chronic pulmonary disease, antibiotic use two weeks before admission, rank on day 1, season of admission, weekday vs. weekend day (the latter defined as Saturday or Sunday), culture results and all variables from the Pneumonia Severity Index (PSI) score (nursing home resident, comorbidities (neoplastic disease, liver disease history, congestive heart failure history, cerebrovascular disease history, renal disease history), altered mental status, respiratory rate, systolic blood pressure, temperature, heart rate, pH, blood urea nitrogen, sodium, glucose, hematocrit, partial pressure of oxygen and pleural effusion on x-ray). Propensity scores were then included as a continuous variable in the Cox proportional hazard regression model. Patients with escalation of therapy were censored at the time of escalation so that only the days before escalation contributed to the analysis. Other patients were censored at day 30.

EFFECT OF CONFOUNDING BY INDICATION

To quantify the effect of unmeasured confounding by indication we simulated clinical stability during hospital admission as a new confounder. We defined clinical stability during admission as a binary variable evaluated at 72 hours, because clinical stability in patients with CAP is often reached within 48 hours and therapy is often evaluated

PART II | CHAPTER 5

CONFOUNDING BY INDICATION OF THE SAFETY OF DE-ESCALATION IN COMMUNITY-ACQUIRED PNEUMONIA:
A SIMULATION STUDY EMBEDDED IN A PROSPECTIVE COHORT.

after three days (with culture results also available)[8, 12, 13]. The strength of any given confounder is determined by the following three parameters: (1) the prevalence in the group with the determinant (de-escalation), (2) the prevalence in group without the determinant (continuation) and (3) the association with patient outcome (mortality). For the simulation of clinical stability at 72 hours we reviewed the literature for reasonable assumptions for the three parameters.
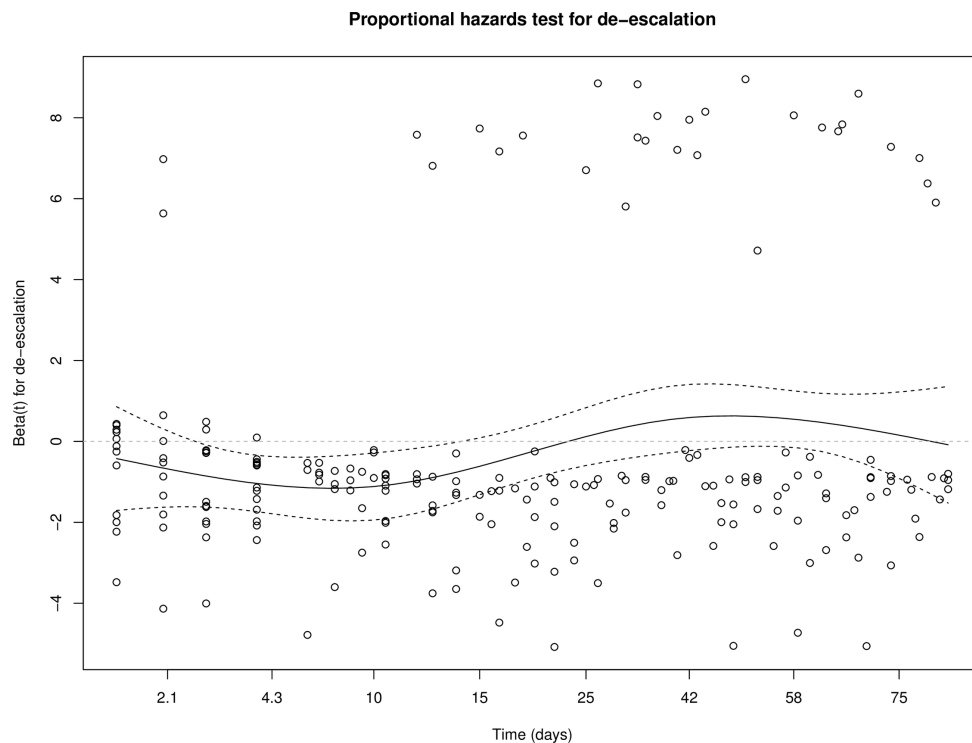
Figure 1. Proportional hazards test for de-escalation. The figure shows that the hazard ratio for de-escalation is not constant over time (i.e. proportion hazards assumption violated): the association between de-escalation and mortality is present within approx. 20-days after admission but is around zero afterwards.

We assumed that 80% of CAP patients admitted to a non-ICU ward will be clinically stable at day three, based on three randomized controlled trials evaluating intravenous to oral switches in patients[14-16]. As the prevalence of clinical stability in the total study population is a weighted average of the prevalence of clinical stability in the de-escalation and the continuation group, the prevalence in one group can be calculated

from the prevalence in the other group. We assumed a high prevalence for clinical stability in the de-escalation group, so we varied the prevalence from 80% to 100%, with corresponding calculated prevalence's in the continued group between 80% and 75% to arrive at the overall prevalence of 80%. The assumed crude odds ratio (OR) between clinical stability at 72 hours and 30-day mortality was 0.14, based on unpublished data of a randomized controlled trial evaluating the effect of adjunct prednisone therapy versus placebo on time to clinical stability for patients with CAP (Courtesy of dr. Blum)[17]. In this trial, clinical stability was measured every 12 hours during hospital stay and was defined as time (days) until stable normalized vital signs for $\geq$ 24 hours: temperature $\leq$ 37.8°C without antipyretic agents, heart rate $\leq$ 100 beats per minute, spontaneous respiratory rate $\leq$ 24 per minute, systolic blood pressure $\geq$ 90 mmHg ($\geq$100 mmHg for patients diagnosed with hypertension) without vasopressor support, mental status back to level before CAP, oxygenation on room air or oxygen therapy (PaO2 $\geq$60 mmHg or pulse oximetry $\geq$ 90%, or PaO2 or pulse oximetry measurement back to baseline for patients with chronic hypoxemia or chronic oxygen therapy)[17]. To simulate the confounder of clinical stability at 72 hours in our dataset, we randomly assigned the presence and the absence of clinical stability such that the aforementioned assumptions about the three parameters were met. Subsequently, the HR of de-escalation on mortality adjusted for clinical stability was determined by including clinical stability as an extra covariate in the propensity score adjusted model. The robustness of the resulting adjusted HRs was tested by repeating the random assignment three times with a different random seed, which verified that the same adjusted HRs was achieved. In the end we plotted the crude and adjusted HR without clinical stability and the resulting HRs for different prevalence's of clinical stability.

We also quantified the strength of each confounder as the change in HR of the model with or without each confounder. For the simulated confounder (clinical stability) we used the corresponding adjusted HR when added to the model with prevalence's of resp. 90% and 100% in the de-escalation group. Data analysis was performed using SPSS for Windows, v.25.0 (SPSS, Chicago, IL, USA) and R v.3.4.3 http://www.R-projects.org/.

RESULTS

ASSOCIATION BETWEEN DE-ESCALATION AND MORTALITY

The study cohort consisted of 3,243 patients admitted with a clinical suspicion of pneumonia. After applying the in- and exclusion criteria 1,536 patients were included for analysis (Fig 2). Empirical treatment was rank-1 in 211 (13.7%), rank-2 in 624 (40.6%), and rank-3 in 701 (45.6%) patients. De-escalation occurred in 257 patients (16.7%)

PART II | CHAPTER 5

CONFOUNDING BY INDICATION OF THE SAFETY OF DE-ESCALATION IN COMMUNITY-ACQUIRED PNEUMONIA:
A SIMULATION STUDY EMBEDDED IN A PROSPECTIVE COHORT.

and escalation occurred in 123 (8.0%) patients. Most patients (1156, 75.3%) continued treatment without a change in rank of antimicrobial therapy during admission (Fig 3). Median time to de-escalation was 3.0 days (IQR 2.0 – 4.0 days). Compared to patients with continued (no de-escalation) regimens, patients with de-escalation less often were current smokers (21/257 (8.2%) vs. 148/1068 (13.9%)), more often had a pathogen identified (107/257 (41.6%) vs. 303/1068 (28.4%)) and had a higher median PSI-score (103 vs. 99) (Table 2). Patients in rank 2 de-escalated less often than patients in rank 3 (6.7% vs. 30.1%; p<0.001). Of the 257 patients with de-escalated therapy, therapy was later escalated in 14 patients (5.5%; 0.9% of all included patients) during admission.

Crude 30-day mortality was 3.5% (9/257) and 10.9% (107/986) in the de-escalation and continuation groups, respectively. The crude and adjusted hazard ratios for de-escalation, compared to continuation, were 0.40 (95% CI: 0.20 – 0.80) and 0.39 (95% CI: 0.19 – 0.79) for day-30 mortality. The AUC of the propensity score was 0.76 (95% CI: 0.73 – 0.79) and was considered acceptable.
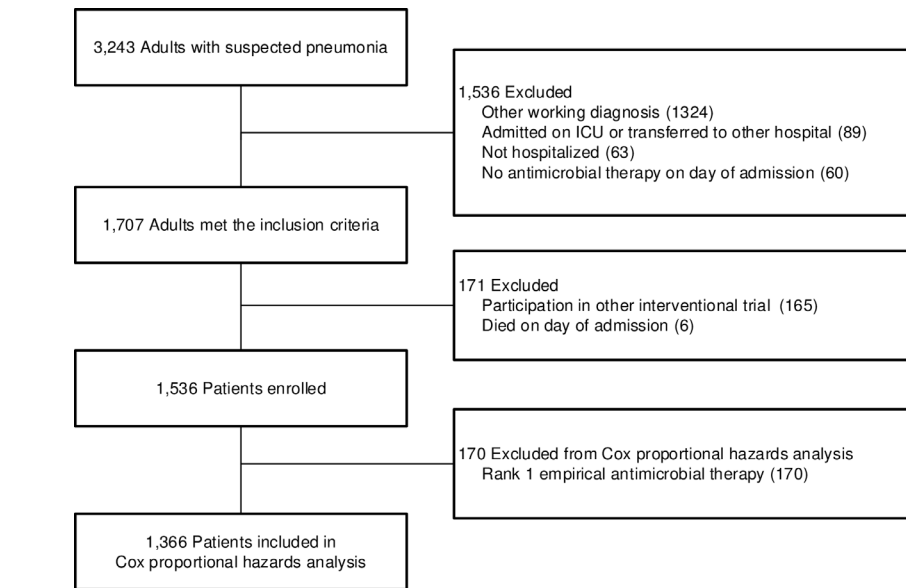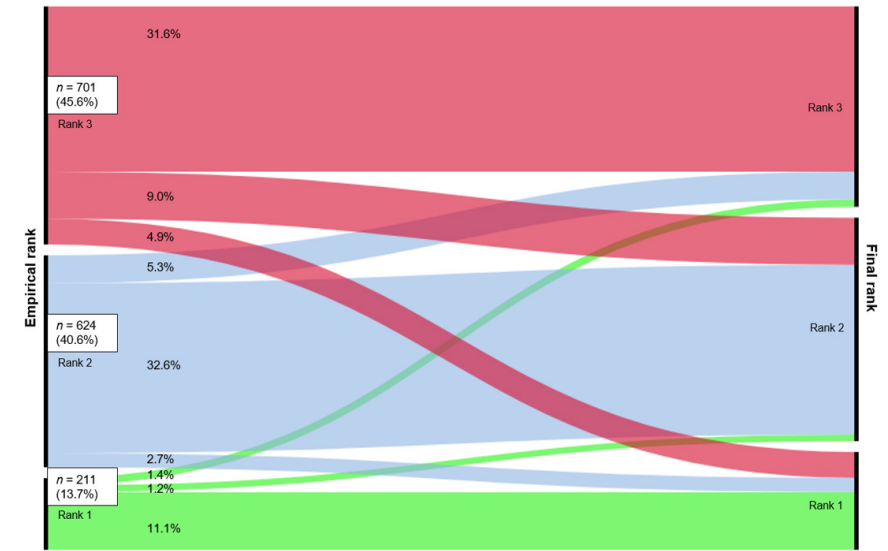

Figure 3. Alluvial-diagram of adjustment of empirical therapy. Rank-1 = narrow-spectrum-antibiotics, rank-2 = broad-spectrum-antibiotics, rank-3 = extended-spectrum-antibiotics. Continued regimen is a straight line, de-escalation is a falling line, escalation is a rising line.

Fig 2. Patient selection.

Table 2. Baseline characteristics. [a]

Patients with a continued regimen and patients with an escalation.

	Total cohort	De-escalation	No de-escalation (rank2-3)[a]	No de-escalation (rank 1)[a]
Patients (N, %)	1536 (100)	257 (16.7)	1068 (69.5)	211 (13.7)
Age (y, median, range)	77 (65 – 100)	77 (66 – 95)	77 (65 – 99)	78 (66 – 100)
Male gender (n, %)	1093 (71.2)	189 (73.5)	768 (71.9)	136 (64.5)
Smoker (n, %)	194 (12.6)	21 (8.2)	148 (13.9)	25 (11.8)
Co-morbidities (n, %)				
Chronic pulmonary disease	849 (55.3)	131 (51.0)	608 (56.9)	110 (52.1)
Chronic cardiovascular disease	650 (42.3)	123 (47.9)	446 (41.8)	81 (38.4)
Chronic renal disease	11 (0.7)	3 (1.2)	5 (0.5)	3 (1.4)
Chronic liver disease	17 (1.1)	5 (1.9)	8 (0.7)	5 (1.9)
Diabetes mellitus	322 (21.0)	53 (20.6)	215 (20.1)	54 (20.6)
PSI score (median, IQR)	99 (82 – 117)	103 (84-121)	99 (82 – 118)	92 (80 – 111)
Antibiotic use before admission (n, %)	493 (32.1)	84 (32.7)	365 (34.2)	44 (20.9)
Pathogen identified (n, %)	469 (30.5)	107 (41.6)	303 (28.4)	59 (28.0)
Day of admission (n, %)				
Weekend	613 (39.9)	71 (27.6)	263 (24.6)	59 (28.0)
Empirical rank on day 1 (n, %)				
Rank 1	211 (13.7)	NA	NA	211 (100)
Rank 2	624 (40.6)	42 (16.3)	582 (54.5)	NA
Rank 3	701 (45.6)	215 (83.7)	486 (45.5)	NA

PART II | CHAPTER 5

CONFOUNDING BY INDICATION OF THE SAFETY OF DE-ESCALATION IN COMMUNITY-ACQUIRED PNEUMONIA:
A SIMULATION STUDY EMBEDDED IN A PROSPECTIVE COHORT.

Effect of confounding by indication due to clinical stability

The results of the simulation analysis are depicted in Fig 4. Not using clinical stability for adjustment yields the afore-mentioned HR of 0.39. When using the assumed odds ratio between clinical stability at 72 hours and 30-day mortality of 0.14, the adjusted HR for de-escalation gradually increased to 1.04 with an increasing prevalence of clinical stability in patients with de-escalation up to 100%. The upper boundary of 95% confidence interval crosses 1 if the prevalence of clinical stability in the de-escalated patients was >= 87%. Determination of the strength of the simulated confounder, clinical stability, revealed that it was substantially stronger than any of the observed confounders in our dataset (Table 3).

Table 3. Strength of known and simulated confounders to the crude HR for 30-day mortality.

Confounder	% change of crude HR
Smoking	+ 1.5
Renal disease	+ 1.7
Respiratory rate	- 1.9
Nursing home resident	+ 2.0
Congestive heart failure	- 2.2
Liver disease	- 2.3
Heart rate	- 2.4
pH	- 2.8
Propensity score	- 3.9
Partial pressure of oxygen	- 4.2
Blood urea nitrogen	- 6.5
Neoplastic disease	+ 7.4
Rank on day 1	- 11.4
Clinical stability (simulated)	
With prevalence in de-escalated group of 90%	+ 37.4
With prevalence in de-escalated group of 100%	+ 157.8

Variables with a change less than 1.5%: diabetes mellitus, sodium, systolic blood pressure, hematocrit, cerebrovascular disease, antibiotic use before admission, day of admission, glucose, chronic pulmonary disease, pleural effusion, altered mental status, age, culture results, season of admission, temperature and gender.

The line reflects the Hazard Ratios for 30-day mortality (based on Cox proportional hazard regression analysis adjusted with propensity scores) with 95% Confidence Interval (shaded area) for different prevalence's of clinical stability in patients with and without de-escalation (horizontal axis). At the left side the weighted average of the two proportions is fixed at 80%, which reflects the adjusted Hazard Ratio without adjustment for clinical stability. The dashed line represents a HR of 1. The HR rises from 0.39 to 1.04 when the prevalence of clinical stability increases to 100% in the de-escalated group. From a prevalence of clinical stability of 87% and above in the de-escalated group the upper limit of the 95% confidence interval included 1. For example a prevalence of 90% in de-escalated results in an adjusted HR of 0.56 (95% CI: 0.27 – 1.12) and a prevalence of 100% results in a HR of 1.04 (95% CI: 0.49 – 2.23).
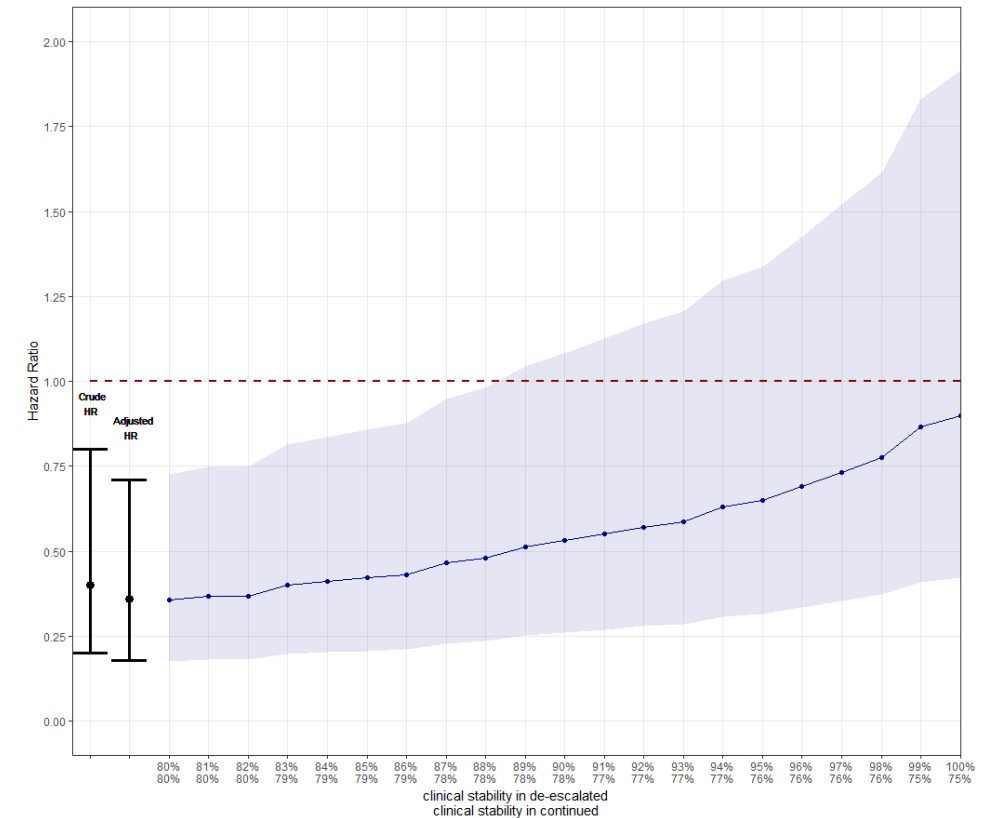
Figure 4. Effect of simulated confounder (clinical stability at 72 hours) on 30-day mortality.

DISCUSSION

In this observational study of patients hospitalized with CAP, after adjustment for observed baseline confounders de-escalation of antimicrobial therapy was associated with a 61% lower hazard of day-30 mortality. However, our simulations have demonstrated that clinical stability at 72 hours, which was not measured in our study, could fully explain this effect under reasonable, literature based assumptions. Based on these findings we conclude that the effects of de-escalation on patient outcome cannot be reliably quantified without adjustment for clinical stability and that the true effect of de-escalation on mortality needs to be quantified by a well-designed prospective study.

De-escalation occurred in 16.7% of the patients. During the enrolment period of our study antibiotic stewardship was not yet well established. Therefore, we expect the proportion of de-escalation in current practice to be larger. In our population, most

PART II | CHAPTER 5

CONFOUNDING BY INDICATION OF THE SAFETY OF DE-ESCALATION IN COMMUNITY-ACQUIRED PNEUMONIA:
A SIMULATION STUDY EMBEDDED IN A PROSPECTIVE COHORT.

patients continued the antibiotic regimen, even though the majority should be clinically stable based on data from the literature. In the absence of antibiotic stewardship, physicians might be more inclined to continue the regimen when it appears to be effective.

In a systematic review including different infectious diseases, de-escalation of empirical antimicrobial therapy was associated with a large reduction in mortality[2]. Although our study only included CAP patients, we expect that the mechanism of bias applies to all infectious diseases for which empirical broad-spectrum antibiotic treatment is common practice. This bias, introduced by not including clinical stability during admission, applies to all previous studies evaluating de-escalation in patients with CAP hospitalized at a non-ICU ward[18-22]. To the best of our knowledge, there are four observational studies on the association between de-escalation and mortality that adjusted for clinical stability or a similar time-varying confounder. In the first study by Joung et al. patients with intensive care unit-acquired pneumonia were included and clinical stability during admission was measured as two scores; APACHE-II (Acute Physiology and Chronic Health Evaluation II) and modified CPIS (clinical pulmonary infection score) both measured on day 5 after development of pneumonia. Both high APACHE II score ($\geq$24) on day 5 and a high CPIS ($\geq$10) on day 5 were associated with an increased 30-day pneumonia-related mortality. By including these confounders, next to other baseline covariates into the multivariable analysis the association between no de-escalation of antibiotics and 30-day mortality resulted in an aHR of 3.988 (95% CI 0.047 – 6.985)[23]. The study objective was to determine independent risk factors for mortality, hence the focus of model building was not on selecting appropriate confounders and one should be careful to interpret the results as a causal effect. In the second study by Garnacho-Montero et al. patients admitted to the ICU with severe sepsis or septic shock were included and clinical stability during admission was measured as Sequential Organ Failure Assessment (SOFA) score on the day when culture results were available. A high SOFA score at culture result day was associated with a higher in-hospital mortality. When including this covariate next to other covariates the association between de-escalation and in-hospital mortality resulted in an aOR of 0.55 (95% CI 0.32 – 0.98, p = 0.022)[5]. In the third study by Montravers et al. patients admitted with health care-associated intra-abdominal infection admitted to ICU were included and clinical stability during admission was measured by SOFA score. Here a decreased SOFA score at day three after initiation of empirical antimicrobial therapy was associated with a lower 28-day mortality. By including this covariate next to other covariates in the analysis this resulted in an aHR of 0.566 (95% CI 0.2503 – 1.278, p = 0.171) for association between de-escalation and 28-day mortality. However, this multivariate analysis also had the purpose to identify risk factors for 28-day mortality, not on selecting appropriate confounders[24]. The fourth study by Lee et al. included

patients with community-onset monomicrobial *Escherichia coli*, *Klebsiella* species and *Proteus mirabilis* bacteremia treated empirically with broad-spectrum beta-lactams and clinical stability during admission was measured by the Pitt bacteremia score. A high Pitt bacteremia score ($\geq$4) at day three was associated with 4-week mortality. After propensity score matching there was no statistically significant difference in mortality rates between de-escalation and no-switch regarding 2-week, 4-week and 8-week mortality[25]. Comparison of the studies is difficult because different criteria for de-escalation and different definitions of disease severity during admission were used, and different populations were studied. The first three studies included ICU patients, and in this setting registering scores representing clinical stability is part of routine care, which makes it more feasible to include such parameters in observational studies. Although the definition for clinical stability for CAP as provided by Halm et al.[13] is widely accepted, in clinical practice patients can be declared stable based on other criteria (e.g. feeling well, eating and drinking) even if they do not meet the formal criteria. A critique of the aforementioned studies is that all used de-escalation as a fixed variable. However, de-escalation is performed on a different day for each individual and should be analyzed as a time-dependent variable, otherwise it introduces immortal time bias[26].

It is recommended to include sensitivity analyses to estimate the potential impact of unmeasured confounding in every non-randomized study on causal associations[27]. However, for observational studies evaluating de-escalation of antimicrobial therapy this has never been done before. To strengthen our sensitivity analysis we based our assumptions about the prevalence of clinical stability and association with mortality on existing high-quality data. We further assumed that physicians will only de-escalate when a patient is clinically stable or to initiate targeted treatment for an identified pathogen. In the latter case, we still expect that most patients in whom the physician decides to de-escalate will be clinically stable. We, therefore, expect that at least 90% and probably close to 100% of de-escalated patients will be clinically stable on day three.

Strengths of our study include the pragmatic approach of using prospectively collected data of a large patient population treated with empiric antibiotics and a working diagnosis of CAP. This included patients without an identified pathogen, which increases the generalizability of our study results. The effect of de-escalation on mortality may be different from one country to another, or even between hospitals within one country, depending on local antibiotic practices. However, we think that the confounding effect of clinical stability is generalizable to other countries and also applies to other severe bacterial infections, because clinical stability will always be a major determinant of de-escalation. A limitation of our study is that we had to exclude 165 patients due to participation in a concurrent trial which could result in selection bias. However this was

PART II | CHAPTER 5

CONFOUNDING BY INDICATION OF THE SAFETY OF DE-ESCALATION IN COMMUNITY-ACQUIRED PNEUMONIA:
A SIMULATION STUDY EMBEDDED IN A PROSPECTIVE COHORT.

a small number of patients and participation was hospital dependent, so the influence of selection bias will be small. Another limitation of our study was that we had to make assumptions for the prevalence of clinical stability in the de-escalated and continued group and for the association between clinical stability and day-30 mortality. These were derived from different study populations, all representing CAP patients hospitalized to a non-ICU ward. Our findings suggest that adjustment for clinical stability will result in a non-significant effect of de-escalation on mortality, which would be biologically plausible. Our findings also demonstrate that the individual baseline confounders, as measured in our study, are poorly predictive for de-escalation, indicating that their correlation with clinical stability is probably also weak.

Another simplification in our analysis was that we modelled clinical stability as a binary variable on day three, which does not well represent reality. For future studies we recommend to measure clinical stability repeatedly over time, as a time-varying confounder and on a continuous scale. Finally, we did not have information on quality of our sputum samples on which the pathogen was identified. Quality of sputum samples is also a prognostic factor for de-escalation of empirical antimicrobial therapy, however we could not correct for this in our model.

The results of our analysis may also suggest that possibility of clinically relevant harm due to de-escalation cannot be excluded, as the upper boundary of the 95% confidence interval for the HR was over to 2 in the most extreme scenario. The scientific evidence for safety of de-escalation is de facto based on two RCTs. However, both RCTs are not powered for mortality. The first prospective, open-label, randomized clinical trial included patients with hospital-acquired pneumonia in an ICU without inclusion criteria regarding baseline clinical stability. After randomization de-escalation was performed three to five days after initiation if empirical treatment when culture results were available. For the association between de-escalation and 14-day mortality the RR was 0.67 (95% CI 0.31 – 1.43), for 28-day mortality the RR was 0.75 (95%CI 0.46 – 1.23) and for in-hospital mortality the RR was 0.64 (95%CI 0.37 – 1.13), (calculated by the authors based on the data reported in[3]. The other multicenter non-blinded randomized non-inferiority trial evaluated the safety of de-escalation with 90-day mortality as secondary outcome in patients with severe sepsis admitted to an ICU without inclusion criteria regarding baseline clinical stability. After randomization de-escalation was performed after culture results were available (IQR 2-4 days after initiation of empirical therapy). In the de-escalation group 18 of 59 patients (31%) died within 90-days, compared to 13 of 57 patients (23%) in the continuation group, yielding an adjusted HR of 1.7 (95% CI 0.79 – 3.49, p = 0.18). Although not statistically significant, this trend may indicate potential harm rather than improved outcome due to de-escalation[4]. As we have demonstrated,

observational studies performed so far do not contribute to determining the safety of de-escalation because the amount of confounding by indication due to clinical stability is insurmountable. As appropriate adjustment of confounding by indication was not performed in the majority of the published observational studies on de-escalation, the ones that adjusted for clinical stability had other important limitations, and only two small RCTs have been performed, we conclude that the safety of this widely propagated antibiotic stewardship intervention should be studied more appropriately. We recommend that future observational studies addressing this research question include clinical stability in the analysis, preferably as a time-varying variable because clinical stability may change over time. It has been suggested that in the case of time-varying confounders a marginal structural model is appropriate[28]. Ultimately, although more expensive, de-escalation would be optimally studied in a pragmatic randomized controlled trial.

To conclude, the previously observed protective effect of de-escalation on mortality is likely due to confounding by unobserved factors such as clinical stability during admission. This study suggests the effect of de-escalation on mortality needs further prospective research to determine effect size more accurately.

REFERENCES

1. Davey P, Marwick CA, Scott CL, Charani E, McNeil K, Brown E, et al. Interventions to improve antibiotic prescribing practices for hospital inpatients. Cochrane Database Syst Rev. 2017;2:CD003543. Epub 2017/02/09. doi: 10.1002/14651858.CD003543.pub4. PubMed PMID: 28178770; PubMed Central PMCID: PMCPMC6464541.

2. Schuts EC, Hulscher MEJL, Mouton JW, Verduin CM, Stuart JWTC, Overdiek HWPM, et al. Current evidence on hospital antimicrobial stewardship objectives: a systematic review and meta-analysis. The Lancet Infectious Diseases. 2016;16(7):847-56. doi: 10.1016/s1473-3099(16)00065-7.

3. Kim JW, Chung J, Choi SH, Jang HJ, Hong SB, Lim CM, et al. Early use of imipenem/cilastatin and vancomycin followed by de-escalation versus conventional antimicrobials without de-escalation for patients with hospital-acquired pneumonia in a medical ICU: a randomized clinical trial. Crit Care. 2012;16(1):R28. Epub 2012/02/18. doi: 10.1186/cc11197. PubMed PMID: 22336530; PubMed Central PMCID: PMCPMC3396273.

4. Leone M, Bechis C, Baumstarck K, Lefrant JY, Albanese J, Jaber S, et al. De-escalation versus continuation of empirical antimicrobial treatment in severe sepsis: a multicenter non-blinded randomized noninferiority trial. Intensive Care Med. 2014;40(10):1399-408. Epub 2014/08/06. doi: 10.1007/s00134-014-3411-8. PubMed PMID: 25091790.

5. Garnacho Montero J, Gutiérrez Pizarraya A, Escoresca Ortega A, Corcia Palomo Y, Fernández Delgado E, Herrera Melero I, et al. De-escalation of empirical therapy is associated with lower mortality in patients with severe sepsis and septic shock. Intensive care medicine. 2014;40(1):32-40. doi: 10.1007/s00134-013-3077-7.

6. Bonten MJM, Huijts S, Bolkenbaas M, Webber C, Patterson S, Gault S, et al. Polysaccharide conjugate vaccine against pneumococcal pneumonia in adults. The New England journal of medicine. 2015;372(12):1114-25. doi: 10.1056/NEJMoa1408544.

7. Postma D, van Werkhoven C, van Elden LJR, Thijsen SFT, Hoepelman AIM, Kluytmans JAJW, et al. Antibiotic treatment strategies for community-acquired pneumonia in adults. The New England journal of medicine. 2015;372(14):1312-23. doi: 10.1056/NEJMoa1406330.

8. Wiersinga WJ, Bonten MJ, Boersma WG, Jonkers RE, Aleva RM, Kullberg BJ, et al. Management of community-acquired pneumonia in adults: 2016 guideline update from the Dutch Working Party on Antibiotic Policy (SWAB) and Dutch Association of Chest Physicians (NVALT). Netherlands journal of medicine. 2018;76(1):4-13.

9. Danish Society of Infectious Diseases. Behandling af samfundserhvervet pneumoni 2011 [August 13, 2019]. Available from: http://www.infmed.dk/guidelines.

10. Spindler C, Stralin K, Eriksson L, Hjerdt-Goscinski G, Holmberg H, Lidman C, et al. Swedish guidelines on the management of community-acquired pneumonia in immunocompetent adults--Swedish Society of Infectious Diseases 2012. Scandinavian journal of infectious diseases. 2012;44(12):885-902. Epub 2012/07/27. doi: 10.3109/00365548.2012.700120. PubMed PMID: 22830356.

11. SWAB (The Dutch Working Party on Antibiotic Policy). A-teams Praktijkgids Antimicrobial Stewardship in Nederland. 2 ed. Leusden: Buro Hans; 2018. 21-2 p.

12. Niederman MS, Mandell LA, Anzueto A, Bass JB, Broughton WA, Campbell GD, et al. Guidelines for the management of adults with community-acquired pneumonia. Diagnosis, assessment of severity, antimicrobial therapy, and prevention. Am J Respir Crit Care Med. 2001;163(7):1730-54. Epub 2001/06/13. doi: 10.1164/ajrccm.163.7.at1010. PubMed PMID: 11401897.

13. Halm EA, Fine MJ, Marrie TJ, Coley CM, Kapoor WN, Obrosky DS, et al. Time to clinical stability in patients hospitalized with community-acquired pneumonia: implications for practice guidelines. Jama. 1998;279(18):1452-7. Epub 1998/05/26. doi: 10.1001/jama.279.18.1452. PubMed PMID: 9600479.

14. Castro-Guardiola A, Viejo-Rodriguez AL, Soler-Simon S, Armengou-Arxe A, Bisbe-Company V, Penarroja-Matutano G, et al. Efficacy and safety of oral and early-switch therapy for community-acquired pneumonia: a randomized controlled trial. Am J Med. 2001;111(5):367-74. Epub 2001/10/05. doi: 10.1016/s0002-9343(01)00868-3. PubMed PMID: 11583639.

15. Kohno S, Yanagihara K, Yamamoto Y, Tokimatsu I, Hiramatsu K, Higa F, et al. Early switch therapy from intravenous sulbactam/ampicillin to oral garenoxacin in patients with community-acquired pneumonia: a multicenter, randomized study in Japan. Journal of infection and chemotherapy : official journal of the Japan Society of Chemotherapy. 2013;19(6):1035-41. Epub 2013/05/23. doi: 10.1007/s10156-013-0618-5. PubMed PMID: 23695232.

16. Oosterheert JJ, Bonten MJ, Schneider MM, Buskens E, Lammers JW, Hustinx WM, et al. Effectiveness of early switch from intravenous to oral antibiotics in severe community acquired pneumonia: multicentre randomised trial. BMJ (Clinical research ed). 2006;333(7580):1193. Epub 2006/11/09. doi: 10.1136/bmj.38993.560984.BE. PubMed PMID: 17090560; PubMed Central PMCID: PMCPMC1693658.

17. Blum CA, Nigro N, Briel M, Schuetz P, Ullmer E, Suter-Widmer I, et al. Adjunct prednisone therapy for patients with community-acquired pneumonia: a multicentre, double-blind, randomised, placebo-controlled trial. Lancet (London, England). 2015;385(9977):1511-8. Epub 2015/01/23. doi: 10.1016/s0140-6736(14)62447-8. PubMed PMID: 25608756.

18. Carugati M, Franzetti F, Wiemken T, Kelley RR, Peyrani P, Blasi F, et al. De-escalation therapy among bacteraemic patients with community-acquired pneumonia. Clin Microbiol Infect. 2015;21(10):936 e11-8. Epub 2015/06/28. doi: 10.1016/j.cmi.2015.06.015. PubMed PMID: 26115864.

19. Cremers AJ, Sprong T, Schouten JA, Walraven G, Hermans PW, Meis JF, et al. Effect of antibiotic streamlining on patient outcome in pneumococcal bacteraemia. J Antimicrob Chemother. 2014;69(8):2258-64. Epub 2014/04/15. doi: 10.1093/jac/dku109. PubMed PMID: 24729585.

20. Viasus D, Simonetti AF, Garcia-Vidal C, Niubo J, Dorca J, Carratala J. Impact of antibiotic de-escalation on clinical outcomes in community-acquired pneumococcal pneumonia. J Antimicrob Chemother. 2017;72(2):547-53. Epub 2016/11/01. doi: 10.1093/jac/dkw441. PubMed PMID: 27798219.

21. West DM, McCauley LM, Sorensen JS, Jephson AR, Dean NC. Pneumococcal urinary antigen test use in diagnosis and treatment of pneumonia in seven Utah hospitals. ERJ Open Res. 2016;2(4). Epub 2017/01/06. doi: 10.1183/23120541.00011-2016. PubMed PMID: 28053969; PubMed Central PMCID: PMCPMC5152836.

22. Yamana H, Matsui H, Tagami T, Hirashima J, Fushimi K, Yasunaga H. De-escalation versus continuation of empirical antimicrobial therapy in community-acquired pneumonia. J Infect. 2016;73(4):314-25. Epub 2016/07/11. doi: 10.1016/j.jinf.2016.07.001. PubMed PMID: 27394401.

PART II | CHAPTER 5

CONFOUNDING BY INDICATION OF THE SAFETY OF DE-ESCALATION IN COMMUNITY-ACQUIRED PNEUMONIA:
A SIMULATION STUDY EMBEDDED IN A PROSPECTIVE COHORT.

23. Joung MK, Lee JA, Moon SY, Cheong HS, Joo EJ, Ha YE, et al. Impact of de-escalation therapy on clinical outcomes for intensive care unit-acquired pneumonia. Crit Care. 2011;15(2):R79. Epub 2011/03/04. doi: 10.1186/cc10072. PubMed PMID: 21366903; PubMed Central PMCID: PMCPMC3219332.

24. Montravers P, Augustin P, Grall N, Desmard M, Allou N, Marmuse JP, et al. Characteristics and outcomes of anti-infective de-escalation during health care-associated intra-abdominal infections. Crit Care. 2016;20:83. Epub 2016/04/08. doi: 10.1186/s13054-016-1267-8. PubMed PMID: 27052675; PubMed Central PMCID: PMCPMC4823898.

25. Lee CC, Wang JL, Lee CH, Hung YP, Hong MY, Tang HJ, et al. Clinical benefits of antimicrobial de-escalation in adults with community-onset monomicrobial Escherichia coli, Klebsiella species and Proteus mirabilis bacteremia. Int J Antimicrob Agents. 2017;50(3):371-6. Epub 2017/07/12. doi: 10.1016/j.ijantimicag.2017.03.024. PubMed PMID: 28694235.

26. Jones M, Fowler R. Immortal time bias in observational studies of time-to-event outcomes. J Crit Care. 2016;36:195-9. Epub 2016/11/05. doi: 10.1016/j.jcrc.2016.07.017. PubMed PMID: 27546771.

27. Groenwold RH, Nelson DB, Nichol KL, Hoes AW, Hak E. Sensitivity analyses to estimate the potential impact of unmeasured confounding in causal research. International journal of epidemiology. 2010;39(1):107-17. Epub 2009/12/02. doi: 10.1093/ije/dyp332. PubMed PMID: 19948779.

28. Hernan MA, Hernandez-Diaz S, Werler MM, Mitchell AA. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. Am J Epidemiol. 2002;155(2):176-84. Epub 2002/01/16. doi: 10.1093/aje/155.2.176. PubMed PMID: 11790682.

6

RELEVANCE OF HEALTHCARE-ASSOCIATED PNEUMONIA FOR EMPIRICAL ANTIBIOTIC THERAPY IN THE NETHERLANDS

Valentijn A. Schweitzer, Cornelis H. van Werkhoven, Inger van Heijl, Rinck F. Smits, C.H. Edwin Boel, Marc J.M. Bonten, Douwe F. Postma, Jan J. Oosterheert

ABSTRACT

BACKGROUND

There is no consensus whether patients with healthcare-associated pneumonia (HCAP) should be treated as hospital-acquired pneumonia (HAP), with broad-spectrum antibiotics, or community-acquired pneumonia (CAP), with narrow-spectrum antibiotics. HCAP research has focused mostly on the predictive value for non-susceptibility to broad-spectrum antibiotics and multi-drug resistant pathogens, in settings with moderate to high levels of antibiotic resistance. We investigated whether HCAP criteria predicts non-susceptibility to different empirical strategies, including narrow-spectrum antibiotics in the Dutch setting.

METHODS

In a post-hoc analysis of patients with moderate-severe CAP in seven Dutch hospitals we compared in vitro antibiotic susceptibilities of definite and probable causative pathogens of CAP and HCAP to amoxicillin and broader antibiotic regimens. In a sensitivity analysis, pathogens with missing susceptibilities were assumed susceptible (best-case scenario) or non-susceptible (worst-case scenario).

RESULTS

Among 2,283 patients with moderate-severe CAP, 23.1% (n=527) were classified as HCAP. Non-susceptibility to amoxicillin ranged from 11.3% (95% CI 9.9-12.8%) (best-case) to 14.4% (95% CI 12.8-16.1%) (worst-case) in CAP patients and from 16.7% (95% CI 13.8-20.1%) (best-case) to 19.7% (95% CI 16.6-23.3%) (worst-case) in HCAP patients. The largest reduction in non-susceptibility was achieved with adding ciprofloxacin to amoxicillin in both CAP patients (10% absolute risk reduction) and HCAP patients (11-16% reduction).

CONCLUSIONS

In the Netherlands, HCAP criteria predict higher amoxicillin non-susceptibility in patients hospitalized with moderate-severe CAP. Although broadening the antibiotic spectrum of empiric treatment reduced the likelihood of non-susceptibility, absolute reductions of non-susceptibility in HCAP patients were too low to justify the universal use of broad-spectrum empirical therapy.

KEYWORDS

Antibiotic resistance, community-acquired pneumonia, empirical antibiotic treatment, healthcare-associated pneumonia

INTRODUCTION

Traditionally, pneumonia is categorized as Community-acquired (CAP), Hospital-acquired (HAP) or Ventilator-associated (VAP), as the aetiology and empirical antibiotic treatment differs depending on where and how the infection is acquired. In 2005, Healthcare-associated pneumonia (HCAP) was introduced as a novel category in the American Thoracic Society (ATS) and the Infectious Diseases Society of America (IDSA) HAP and VAP guideline[1]. Patients with HCAP often present at the emergency department, but are distinguished from CAP patients by their recent contact with healthcare institutions. As a consequence, HCAP patients may have a different bacterial aetiology of infection and increased risk for colonization and infection with antibiotic resistant or healthcare-associated pathogens, such as *Staphylococcus aureus*, Gram-negative Enterobacteriaceae, and *Pseudomonas aeruginosa*[1]. Therefore, the guideline recommended to empirically treat HCAP with broad-spectrum antibiotics, similar to HAP and VAP[1]. This has led to a large increase of broad-spectrum antibiotic use without apparent clinical benefit for these patients[2–5]. Recent evidence suggests that the predictive value of HCAP criteria for the need of broad-spectrum antibiotic treatment might be lower than anticipated[2,6–11]. In response to these findings, HCAP was removed from the 2016 ATS HAP/VAP guidelines and it was suggested to consider implementing HCAP recommendations in CAP guidelines, as both CAP and HCAP patients are initially cared for in the emergency department[8]. Several studies have already evaluated the predictive value of HCAP criteria for bacterial aetiology in CAP patients[6,7]. However, the appropriateness of introducing HCAP in CAP guidelines depends on the prevalence of pathogens requiring broader antibiotic treatment and the preferred empirical treatment for CAP patients, which differs per geographical region. In the Netherlands, the first choice empirical treatment for moderate-severe CAP is narrow-spectrum beta-lactam monotherapy. Current HCAP research focuses on predicting multi-drug resistant pathogens (including Methicillin-resistant Staphylococcus aureus (MRSA) and extended spectrum beta-lactamase (ESBL)–producing Enterobacteriaceae) and non-susceptibility to broad-spectrum beta-lactams (ceftriaxone or ampicillin-sulbactam), macrolides and fluoroquinolones[17–21]. In the Netherlands, infections caused by these resistant pathogens are rare and pneumonia acquired in nursing homes is usually considered as HAP. This is why the relevance and predictive value of HCAP criteria for the Northern European or Dutch setting, i.e. for the non-susceptibility to narrow-spectrum beta-lactams, remains unknown. Our main study objective was to evaluate the predictive value of HCAP criteria for narrow-spectrum beta-lactam (i.e. amoxicillin) non-susceptibility (thus needing broad-spectrum treatment) in patients hospitalized for moderate-severe CAP. In addition, we assessed the predictive value of HCAP criteria for non-susceptibility to broader antibiotic regimens, including amoxicillin-clavulanic

acid, ceftriaxone, moxifloxacin, amoxicillin plus azithromycin, and amoxicillin plus ciprofloxacin.

MATERIALS AND METHODS

STUDY SUBJECTS AND DESIGN

We performed a post-hoc analysis of an observational cohort study, nested within the Community-Acquired Pneumonia — Study on the Initial Treatment with Antibiotics of Lower Respiratory Tract Infections (CAP-START trial), which was a cluster randomised trial performed between February 2011 and August 2013 in seven hospitals in the Netherlands[22]. Patients above 18 years of age who were admitted to a non-ICU ward for suspicion of pneumonia were eligible for study participation. The study was approved by the ethics review board of the University Medical Center Utrecht (reference number 10/148). Written informed consent for data collection was obtained within 72 hours after hospital admission.

DATA COLLECTION

Data on HCAP criteria, co-morbidities, clinical presentation, antibiotic use, complications, and clinical outcome were retrieved prospectively from medical records by trained research nurses after patient inclusion. As pneumonia acquired in nursing homes are considered as HAP in the Netherlands, these patients were not included in the original trial. Therefore, the following HCAP definition was used: hospitalization within the last 90 days, residence in long-term care facilities other than nursing homes, receiving wound care or intravenous therapy in the previous 30 days, or attending haemodialysis clinics[1].

MICROBIOLOGY

Sputum and blood cultures, urinary antigen tests, and antibiotic susceptibility testing was performed as part of routine care. Susceptibility was determined by routinely performed microbiological tests. To account for the possibility of false-positive tests due to colonization, the causative pathogen per patient was determined accounting for the specificity of the different microbiological tests, where positive urine antigen tests and blood cultures were assumed to have a higher specificity for causative pathogen than sputum cultures. For example: in a patient with a positive pneumococcal urinary antigen test and *S. aureus* cultured from sputum, *Streptococcus pneumoniae* was considered the causative pathogen due to the higher specificity of the urinary antigen test and *S. aureus* was considered as colonization. Susceptibility testing was reported as S (sensitive), I (intermediate), or R (resistant) by participating microbiology laboratories.

I and R results were considered as non-susceptible for all the analyses. In patients with multiple possible causative pathogens (i.e. multiple pathogens in sputum culture and no pathogens from blood culture or urinary antigen tests), susceptibility to antibiotics was determined by the most resistant pathogen. In case of missing susceptibility data, susceptibility per antibiotic was imputed and assumed to be susceptible (S) if the prevalence of antibiotic resistance to the antibiotic was ≤10% in national surveillance data, non-susceptible (R) if the prevalence was ≥90% or unknown (U) if the prevalence was between 10 and 90% (table S1). Pathogens were considered susceptible (S) to combination antibiotic therapy if susceptible to any of the two antibiotics, unknown (U) if susceptible to one antibiotic and unknown to the other antibiotic or if unknown to both antibiotics and non-susceptible (R) if non-susceptible to both antibiotics.

STATISTICAL ANALYSIS

Descriptive statistics were used to compare baseline characteristics between CAP and HCAP patients. Sensitivity analyses were performed for cases with unknown (U) antibiotic susceptibility which were either assumed to be all susceptible (best-case scenario) or assumed to be all non-susceptible (worst-case scenario). Predictive values, sensitivity and specificity for non-susceptibility per empirical antibiotic strategy were calculated using 2x2 contingency tables. 95% confidence intervals were calculated using the Wilson score interval method[23]. Analyses were performed using the Statistical Package for the Social Sciences for Windows (Version SPSS 21.0.0.0). Graphs were created using GraphPad PRISM (Version 7.02).

RESULTS

A total of 2,283 patients with moderate-severe CAP were included in the CAP-START study, of which 527 (23.1%) were classified as HCAP. Of the patients with HCAP, 318 (60%) were hospitalized within the last 90 days; 111 (21%) resided in an elderly home; 166 (32%) received intravenous therapy in the previous 30 days; 94 (18%) received wound care in the previous 30 days; and 17 (3%) were on chronic haemodialysis. In comparison to patients with CAP, patients with HCAP were older, had more co-morbidities, had higher disease severity scores (PSI on admission), had higher influenza vaccination rates, were more often dependent on daily living activities (ADL), and more often had treatment restrictions (Table 1). Clinical outcomes of patients with HCAP were worse, with a higher in-hospital, 30-day, and 90-day mortality. There were no differences between patients with CAP and HCAP regarding the frequency with which microbiological testing was performed, except for a slightly higher rate of Legionella urinary antigen testing in patients with CAP (Table 1).

Table 1. Descriptive statistics of CAP and HCAP patients

	CAP (n= 1,756)	HCAP (n=527)
Male sex (n, %)	994 (56.6)	139 (61.1)
Age (y, median, IQR)	70 (58-79)	72 (62-81)
PSI-score (mean, SD)	132 (20.5)	137 (27.1)
Received antibiotics before admission (%)	32.2	34.9
Received pneumococcal vaccination (%)	1.9	2.5
Received influenza vaccination (%)	63.2	72.3
ADL dependent (%)	22.8	27.5
Any treatment restriction (%)	23.8	46.1
Co-morbidities		
Immunocompromised* (%)	18.4	39.8
Cardiovascular disease (%)	20.0	24.3
COPD or asthma (%)	38.7	45.4
Cerebrovascular disease (%)	9.2	14.2
Diabetes mellitus (%)	16.0	18.8
Malignancy (%)	10.6	22.6
Chronic renal failure (%)	0.5	4.2
Microbiologic testing performed		
Sputum culture (%)	46.1	44.4
Blood culture (%)	76.1	76.1
Pneumococcal urinary antigen test (%)	79.2	77.0
Legionella urinary antigen test (%)	77.1	72.3
Clinical outcome		
ICU admission during hospital stay (%)	1.9	2.1
All-cause mortality		
In-hospital (%)	2.8	4.7
Day 30 (%)	4.3	8.3
Day 90 (%)	7.1	17.8
Length of hospital stay (d, median, IQR)	6 (4-9)	6 (4-10)

Data are presented as n (%) unless otherwise indicated. Abbreviations: ADL: activities of daily living, COPD: chronic obstructive pulmonary disease, PSI: Pneumonia severity index.

*Immunocompromised is defined by a history of HIV, AIDS, leukaemia, lymphoma, Hodgkin lymphoma, multiple myeloma, generalized malignancy, chronic renal failure, nephrotic syndrome, immunosuppressive therapy or transplantation

MICROBIOLOGY

A bacterial pathogen was identified in 566 (32%) of the CAP patients and 178 (34%) of the HCAP patients, most frequently based on sputum culture (n=368, 50%), urinary antigen testing (n=224, 30%), blood culture (n=98. 13%), BAL (n=22, 3%), or serology (n=13, 2%). The most frequent causative pathogen was *S. pneumoniae* in both CAP and HCAP patients (14.0% and 11.2%, respectively, Table 2). In comparison to CAP, HCAP was less frequently caused by *S. pneumoniae* and *H. influenzae* and more frequently caused by *S. aureus, P. aeruginosa,* and *E. coli*, and more often had multiple pathogens identified. Of all the 4,464 bacterial pathogen / antibiotic strategy combinations, 20% (n=909) were confirmed by susceptibility testing, 65% (n=2,921) were assumed to be sensitive or resistant based on intrinsic resistance or national surveillance and 15% (n=634) were unknown.

Table 2. Pathogens in CAP and HCAP patients

	CAP n= 1,756	HCAP n= 527
Streptococcus pneumoniae	246 (14.0)	59 (11.2)
Haemophilus influenzae	92 (5.2)	17 (3.2)
Multiple pathogens*	52 (3.0)	27 (5.1)
Staphylococcus aureus	30 (1.7)	15 (2.8)
Pseudomonas aeruginosa	27 (1.5)	15 (2.8)
Escherichia coli	23 (1.3)	18 (3.4)
Mycoplasma pneumoniae	23 (1.3)	1 (0.2)
Legionella pneumophila	17 (1.0)	1 (0.2)
Moraxella catarrhalis	12 (0.7)	4 (0.8)
Klebsiella pneumoniae	7 (0.4)	2 (0.4)
Other gram-negative bacteria	31 (1.8)	13 (2.5)
Other gram-positive bacteria	6 (0.3)	6 (1.1)
Total	566 (32.2)	178 (33.8)

Data are given as n(%).

* Most frequent multiple pathogen combinations were *S. pneumoniae* with *H. influenzae* (20%) and *H. influenzae* with *S. aureus* (15%) in CAP and *H. influenzae* with *M. catarrhalis* (16%) and *H. influenzae* with *E. coli* (12%) in HCAP patients.

## PREDICTIVE VALUE FOR AMOXICILLIN NON-SUSCEPTIBILITY

The prevalence of non-susceptibility to amoxicillin for the best-case and worst-case scenario was 11.3% (95% confidence interval (CI) 9.9%-12.8%) and 14.4% (95% CI 12.9%-16.1%) in CAP patients and 16.7% (95% CI 13.6%-20.1%) and 19.7% (95% CI 16.6%-23.3%) in HCAP patients, respectively (Figure 1 A and B). The corresponding negative predictive values, which is the prevalence of amoxicillin susceptibility in CAP patients without HCAP criteria, were 88.7% (95% CI 87.2%-90.1%) and 85.6% (95% CI 83.9%-87.2%) (1 minus non-susceptibility rate in CAP patients) for the best- and worst-case scenarios, respectively, with sensitivities of 30.8% (95% CI 25.7%-36.3%) and 29.2% (95% CI 24.7%-34.1%), and specificities of 78.0% (95% CI 76.1%-79.8%) and 78.0% (95% CI 76.1%-79.8%).

PREDICTIVE VALUE FOR BROAD-SPECTRUM NON-SUSCEPTIBILITY

When comparing antibiotic non-susceptibility rates we used the non-susceptibility rate for amoxicillin as a reference, which was 11.3-14.4% in CAP patients and 16.7-19.7% in HCAP patients. In comparison to this reference, other antibiotic combinations reduced the proportion of CAP patients with non-susceptibility by 5-10% and by 7-16% in HCAP patients. The largest reduction in non-susceptibility compared to amoxicillin was achieved by adding ciprofloxacin to amoxicillin in both CAP and HCAP patients. In CAP patients, the 11.3-14.4% (best-case, worst-case) non-susceptibility to amoxicillin was reduced by 10% to a non-susceptibility of 0.8% (95% CI 0.4%-1.3%) (best-case) and 4.1% (95% CI 3.3%-5.1%) (worst-case). In HCAP patients, the 16.7-19.7% (best-case, worst-case) was reduced by 11-16% to a non-susceptibility of 1.3% (95% CI 0.6%-2.7%) (best-case) and 8.3% (95% CI 6.3%-11.0%) (worst-case) with amoxicillin plus ciprofloxacin. (figure 1C and D).

Table 3. Diagnostic values of HCAP to predict for antibiotic resistance

		Resistance rate			
	Scenario	CAP (n=1,756) (n (%, 95% CI))	HCAP (n=527) (n (%, 95% CI))	Sensitivity (%, 95% CI)	Specificity (%, 95% CI)
Amoxicillin	Best-case	198 (11.3 (9.8-12.8))	88 (16.7 (13.8-20.1))	30.8 (25.7-36.3)	78.0 (76.1-79.8)
	Worst-case	252 (14.4 (12.8-6.1))	104 (19.7 (16.6-23.3))	29.2 (24.7-34.1)	78.0 (76.1-79.8)
Amoxicillin-clavulanic acid	Best-case	117 (6.7 (5.6-7.9))	52 (9.9 (7.6-12.7))	30.8 24.3-38.1)	77.5 (75.7-79.3)
	Worst-case	131 (7.5 (6.3-8.8))	62 (11.8 (9.3-14.8))	32.1 (25.9-39.0)	77.8 (75.9-79.5)
Ceftriaxone	Best-case	90 (5.1 (4.2-6.3))	35 (6.6 (4.8-9.1))	28.0 (20.9-36.4)	77.2 (75.4-78.9)
	Worst-case	108 (6.2 (5.2-7.4))	45 (8.5 (6.4-11.2))	29.4 (22.8-37.1)	77.4 (75.5-79.1)
Moxifloxacin	Best-case	33 (1.9 (1.3-2.6))	27 (5.1 (3.5-7.4))	45.0 (33.1-57.5)	77.5 (75.7-79.2)
	Worst-case	119 (6.8 (5.7-8.0))	71 (13.5 10.8-16.7)	37.4 (30.8-44.4)	78.2 (76.4-79.9)
Amoxicillin + azithromycin	Best-case	59 (3.4 (2.6-4.3))	28 (5.3 (3.7-7.6))	32.2 (23.3-42.6)	77.3 (75.5-79.0)
	Worst-case	132 (7.5 (6.4-8.8))	51 (9.7 (7.4-12.5))	27.9 (21.9-34.8)	77.3 (75.5-79.1)
Amoxicillin + ciprofloxacin	Best-case	14 (0.8 (0.5-1.3))	7 (1.3 (0.6-2.7))	33.3 (17.2-54.6)	77.0 (75.2-78.7)
	Worst-case	72 (4.1 (3.3-5.1))	44 (8.3 (6.3-11.0))	37.9 (29.6-47.0)	77.7 (75.9-79.4)
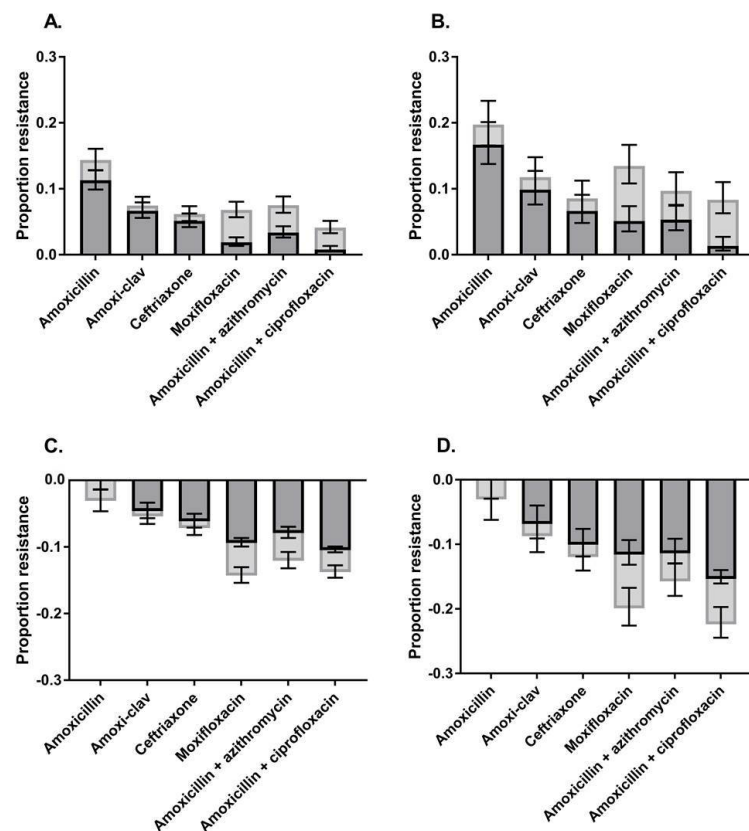
Figure 1. Non-susceptibility for antibiotics in CAP (A) and HCAP patients (B) and the difference in non-susceptibility compared to amoxicillin (CAP: C and HCAP: D). Dark grey indicates the best-case scenario, white grey indicates the worst-case scenario. Confidence intervals are given for both scenarios.

DISCUSSION

In this study of patients with a clinical diagnosis of CAP admitted to non-ICU wards, non-susceptibility of CAP pathogens to amoxicillin was 5-6% higher in patients who met the HCAP criteria compared to patients without HCAP criteria. The most commonly identified pathogens were *S. pneumoniae* and *H. influenzae* in CAP patients and *S. pneumoniae*, multiple pathogens, *S. aureus, P. aeruginosa,* and *E. coli* in HCAP patients. These findings are comparable to previous reports[18–20,24]. Naturally, this difference in non-susceptibility could be reduced by broadening the empiric antibiotic spectrum for HCAP patients. To our knowledge, this is the first study to assess the predictive value of HCAP criteria for non-susceptibility to narrow-spectrum beta-lactams. As such, the presented data may be useful in the discussion of whether HCAP should be implemented in CAP guidelines in settings where narrow-spectrum beta–lactam monotherapy is the first choice treatment.

Despite the differences in aetiology between CAP and HCAP patients, several observational studies from the US have failed to show benefit of broad-spectrum empirical antibiotics on the clinical outcome of HCAP patients, with some even showing worse clinical outcome[3–5]. However, these observational studies most likely suffered from confounding by indication, where underlying conditions, frailty, severity of disease, and treatment restrictions may have influenced the association between treatment and outcome of HCAP patients. In addition, being able to predict non-susceptibility to empirical antibiotics does not necessarily mean that such patients would benefit from broader empirical therapy. It may also be safe to start with narrow-spectrum antibiotics and escalate treatment based on culture or urine antigen testing results. Therefore, proper randomised trials are required to validly assess treatment effects on clinical outcome in HCAP patients.

As the criteria for HCAP have often been questioned, multiple studies have evaluated other risk factors or scores to predict antibiotic resistance in CAP patients. The risk factors evaluated to date include: family members with resistant bacteria, severe pneumonia, prior antibiotic use, functional status, ICU admission, immunosuppression, co-morbidities (cerebrovascular disease, diabetes, COPD), gastric acid suppression medication, tube feeding, prior infection with a drug resistant pathogen, and MRSA colonization[17,18,20,21,24–26]. However, many of these risk factors were evaluated in settings with a high prevalence of antibiotic resistance. Whether their predictive value can be generalised to settings with a low prevalence of antibiotic resistance remains to be elucidated. Moreover, many of the aforementioned risk factors, such as previous colonization with MRSA, are not appropriate for settings with low antibiotic resistance

prevalence. In an additional analysis, we explored the predictive value of severe pneumonia (CURB-65 score>2), prior antibiotic use, functional status (ADL dependence), immunosuppression, and co-morbidities (cardiovascular disease, diabetes mellitus, COPD) in a multivariable model. From these variables, only HCAP, immunosuppression, cerebrovascular disease and diabetes mellitus were predictive for amoxicillin non-susceptibility (Table S2). However, the discriminative capacity of the multivariable model remained limited. The predictive value of these variables, in combination with other promising predictors (such as previous colonisation with resistant bacteria) should be evaluated in a prospective cohort study.

This study has several strengths. We used high quality data from a prospective multicentre trial, including consecutive patients with moderate-severe community-acquired pneumonia, irrespective of whether a bacterial pathogen was isolated. In contrast to including patients with positive cultures only, the predictive values presented here are directly relevant for clinical practise[17,21,24]. In addition, we used extensive antibiotic susceptibility data and assessed non-susceptibility over a range of different empirical antibiotic treatment regimens. There were also several limitations. First, patients that resided in nursing homes were excluded because their disease was not considered to be CAP. Therefore, one could argue that we did not include the entire spectrum/domain of HCAP and the presented results might not be generalizable to the international HCAP definition. However, these patients would generally be considered as hospital- or nursing-home acquired pneumonia and treated as such. Second, diagnostic testing was performed as part of routine care, which is why blood cultures, sputum cultures, and urinary antigen testing were not uniformly performed. Yet, although we cannot exclude the possibility of bias in outcome assessment, there were no major differences between rates of microbiological testing in CAP and HCAP patients. Third, although there were missing susceptibility data for individual antibiotics in certain pathogens, imputed susceptibility data were based on local surveillance data and, therefore, generalizable to settings with low antibiotic resistance. In addition, we performed sensitivity analyses on susceptibility patterns that remained unknown with a best-case and worst-case scenario where the unknown susceptibilities were either all susceptible or non-susceptible. These sensitivity analyses yielded only small variations in non-susceptibility for the different antibiotic regimens. Fourth, the probable causative pathogen was in many cases based on sputum cultures, which might represent colonisation rather than infection. Yet, therefore we only considered plausible pneumonia pathogens in our analyses. Lastly, we assumed cases of pneumonia without a causative pathogen to be susceptible to all antibiotics, which might not be true in case of false-negative culture results for resistant pathogens in a subset of patients.

To conclude, HCAP criteria predict for higher non-susceptibility rates to amoxicillin in patients hospitalized with CAP to non-ICU wards in the Netherlands. However, we consider the absolute risk difference of non-susceptibility to amoxicillin between CAP and HCAP patients as being too low to justify treating all HCAP patients with broad-spectrum antibiotics. Future research should focus on identifying and validating risk factors to predict for narrow-spectrum beta-lactam antibiotic non-susceptibility that are appropriate for settings with low antibiotic resistance. Furthermore, prediction rules need to be evaluated in randomised clinical trials to show benefit on clinical outcome.

REFERENCES

1. American Thoracic Society; Infectious Diseases Society of America. Guidelines for the Management of Adults with Hospital-acquired, Ventilator associated, and Healthcare-associated Pneumonia. Am J Respir Crit Care Med, 171 (2005), pp. 388-416

2. Chalmers JD, Rother C, Salih W, Ewig S. Healthcare-associated pneumonia does not accurately identify potentially resistant pathogens: A systematic review and meta-analysis. Clin Infect Dis, (2014);58:330−9. doi:10.1093/cid/cit734.

3. Madaras-Kelly KJ, Remington RE, Sloan KL, Fan VS. Guideline-based antibiotics and mortality in healthcare-associated pneumonia. J Gen Intern Med 2012;27:845−52. doi:10.1007/s11606-012-2011-y.

4. Rothberg MB, Zilberberg MD, Pekow PS, Priya A, Haessler S, Belforti R, et al. Association of guideline-based antimicrobial therapy and outcomes in healthcare-associated pneumonia. J Antimicrob Chemother 2014;70:1573−9. doi:10.1093/jac/dku533.

5. Attridge RT, Frei CR. Health care-associated pneumonia: An evidence-based review. Am J Med 2011;124:689−97. doi:10.1016/j.amjmed.2011.01.023.

6. Garcia-Vidal C, Viasus D, Roset A, Adamuz J, Verdaguer R, Dorca J, et al. Low incidence of multidrug-resistant organisms in patients with healthcare-associated pneumonia requiring hospitalization. Clin Microbiol Infect 2011;17:1659−65. doi:10.1111/j.1469-0691.2011.03484.x.

7. Gross AE, Van Schooneveld TC, Olsen KM, Rupp ME, Bui TH, Forsung E, et al. Epidemiology and predictors of multidrug-resistant community-acquired and health care-associated pneumonia. Antimicrob Agents Chemother 2014;58:5262−8. doi:10.1128/AAC.02582-14.

8. Kalil AC, Metersky ML, Klompas M, Muscedere J, Sweeney DA, Palmer LB, et al. Management of Adults With Hospital-acquired and Ventilator-associated Pneumonia: 2016 Clinical Practice Guidelines by the Infectious Diseases Society of America and the American Thoracic Society. Clin Infect Dis 2016;63:e61−111. doi:10.1093/cid/ciw353.

9. Yap V, Datta D, Metersky ML. Is the Present Definition of Health Care-Associated Pneumonia the Best Way to Define Risk of Infection with Antibiotic-Resistant Pathogens? Infect Dis Clin North Am 2013;27:1−18. doi:10.1016/j.idc.2012.11.002.

10. Jones BE, Jones MM, Huttner B, Stoddard G, Brown KA, Stevens VW, et al. Trends in antibiotic use and nosocomial pathogens in hospitalized veterans with pneumonia at 128 medical centers, 2006-2010. Clin Infect Dis 2015;61:1403−10. doi:10.1093/cid/civ629.

11. Valles J, Martin-Loeches I, Torres A, Diaz E, Seijas I, Lopez MJ, et al. Epidemiology, antibiotic therapy and clinical outcomes of healthcare-associated pneumonia in critically ill patients: A Spanish cohort study. Intensive Care Med 2014;40:572−81. doi:http://dx.doi.org/10.1007/s00134-014-3239-2.

12. Mandell L, Wunderink RG, Anzueto A, Bartlett JG, Campbell GD, Dean NC, et al. Infectious Diseases Society of America/American Thoracic Society consensus guidelines on the management of community-acquired pneumonia in adults. Clin Infect Dis 2007;44 Suppl 2:S27-72. doi:10.1086/511159.

13. Lim WS, Baudouin S V, George RC, Hill AT, Jamieson C, Le Jeune I, et al. BTS guidelines for the management of community acquired pneumonia in adults: update 2009. Thorax 2009;64:iii1-iii55. doi:10.1136/thx.2009.121434.

14. Spindler C, Strålin K, Eriksson L, Hjerdt-Goscinski G, Holmberg H, Lidman C, et al. Swedish guidelines on the management of community-acquired pneumonia in immunocompetent adults - Swedish Society of Infectious Diseases 2012. Scand J Infect Dis 2012;44:885−902. doi:10.3109/00365548.2012.700120.

15. Wiersinga WJ, Bonten MJ, Boersma WG, Jonkers RE, Aleva RM, Kullberg BJ, et al. SWAB/NVALT (Dutch working party on antibiotic policy and Dutch association of chest physicians) guidelines on the management of community-acquired pneumonia in adults. Neth J Med 2012;70:90−101.

16. Behandling af samfundserhvervet pneumoni (2011), http://www.infmed.dk/guidelines, accessed at December 2017

17. Webb BJ, Dascomb K, Stenehjem E, Vikram HR, Agrwal N, Sakata K, et al. Derivation and multicenter validation of the drug resistance in pneumonia clinical prediction score. Antimicrob Agents Chemother 2016;60:2652−63. doi:10.1128/AAC.03071-15.

18. Shindo Y, Ito R, Kobayashi D, Ando M, Ichikawa M, Shiraki A, et al. Risk factors for drug-resistant pathogens in community-acquired and healthcare-associated pneumonia. Am J Respir Crit Care Med 2013;188:985−95. doi:10.1164/rccm.201301-0079OC.

19. Self WH, Wunderink RG, Williams DJ, Barrett TW, Baughman AH, Grijalva CG. Comparison of clinical prediction models for resistant bacteria in community-onset pneumonia. Acad Emerg Med 2015;22:730−40. doi:10.1111/acem.12672.

20. Aliberti S, Di Pasquale M, Zanaboni AM, Cosentini R, Brambilla AM, Seghezzi S, et al. Stratifying risk factors for multidrug-resistant pathogens in hospitalized patients coming from the community with pneumonia. Clin Infect Dis 2012;54:470−8. doi:10.1093/cid/cir840.

21. Shorr AF, Zilberberg MD, Reichley R, Kan J, Hoban A, Hoffman J, et al. Validation of a clinical score for assessing the risk of resistant pathogens in patients with pneumonia presenting to the emergency department. Clin Infect Dis 2012;54:193−8. doi:10.1093/cid/cir813.

22. Postma DF, van Werkhoven CH, van Elden LJ, Thijsen SF, Hoepelman AI, Kluytmans JA, et al. Antibiotic Treatment Strategies for Community-Acquired Pneumonia in Adults. NEJM 2015;372;14:1312−23. doi:10.1056/NEJMoa1406330.

23. DasGupta A, Cai TT, Brown LD. Interval Estimation for a Binomial Proportion. Stat Sci 2001;16:101−33. doi:10.1214/ss/1009213286.

24. Park SC, Kim EY, Kang YA, Park MS, Kim YS, Kim SK, et al. Validation of a scoring tool to predict drug-resistant pathogens in hospitalised pneumonia patients. Int J Tuberc Lung Dis 2013;17:704−9. doi:10.5588/ijtld.12.0723.

25. Schreiber MP, Chan CM, Shorr AF. Resistant pathogens in nonnosocomial pneumonia and respiratory failure: Is it time to refine the definition of health-care-associated pneumonia? Chest 2010;137:1283−8. doi:10.1378/chest.09-2434.

26. Brito V, Niederman MS. Healthcare-associated pneumonia is a heterogeneous disease, and all patients do not need the same broad-spectrum antibiotic therapy as complex nosocomial pneumonia. Curr Opin Infect Dis 2009;22:316−25. doi:10.1097/QCO.0b013e328329fa4e.

SUPPLEMENT

Table S1. Assumed antibiotic susceptibility patterns of pathogens in case of missing resistance data.

	AMO	ACL	CTR	AZI	MOX	CIP
Streptococcus pneumoniae	S	S	S	U	S	R
Haemophilus influenzae	U	S	S	U	S	S
Escherichia coli	U	U	S	R	U	U
Staphylococcus aureus	R	S	S	U	U	U
Pseudomonas aeruginosa	R	R	R	R	R	U
Mycoplasma pneumoniae	R	R	R	S	S	S
Legionella pneumophila	R	R	R	S	S	S
Moraxella catarrhalis	R	S	S	S	S	S
Klebsiella pneumoniae	R	U	U	R	U	U
Stenotrophomonas species	R	R	R	R	S	R
beta-hemolytic streptococcus	S	S	S	S	S	U
Serratia marcescens	R	R	R	R	U	U
Enterobacter species	R	R	R	R	U	U
Klebsiella oxytoca	R	U	U	R	U	U
Enterobacter cloacae	R	R	U	R	U	U
Morganella morganii	R	R	U	R	U	U
Pneumocystis jirovecii	R	R	R	R	R	R
Acinetobacter species	R	R	U	R	U	U
Citrobacter freundii	R	R	S	U	U	U
Citrobacter species	R	R	S	U	U	U
Coxiella burnetti	R	R	R	U	U	U
Enterobacter aerogenes	R	R	U	R	S	S
Klebsiella ozaenea	R	S	S	R	S	S
Proteus mirabilis	U	S	S	U	S	S
Serratia liquefaciens	R	R	U	R	S	S

AMO: amoxicillin, ACL: amoxicillin-clavulanic acid, CTR: ceftriaxone, AZI: azithromycin, MOX: moxifloxacin, CIP: ciprofloxacin

(S) susceptible; prevalence of antibiotic resistance to the antibiotic $\leq$10%

(R) non-susceptible; prevalence of antibiotic resistance to the antibiotic $\geq$90%

(U) unknown; prevalence of antibiotic resistance to the antibiotic >10% and <90%

Table S2. Multivariable and univariate prediction model results for amoxicillin non-susceptibility

		Bootstrapped OR (95% CI)	AUC* of ROC curve (95% CI)
Best-case scenario	HCAP (univariate)	1.6 (1.2-2.1)	0.54 (0.5-0.58)
	HCAP	1.5 (1.1-2.0)	0.58 (0.54-0.61)
	History of cerebrovascular disease	1.8 (1.2-2.6)	
	Diabetes mellitus	0.7 (0.4-0.9)	
Worst-case scenario	HCAP (univariate)	1.5 (1.1-1.9)	0.54 (0.5-0.57)
	HCAP	1.4 (1.1-1.8)	0.56 (0.53-0.6)
	History of cerebrovascular disease	1.5 (1.0-2.1)	
	Diabetes mellitus	1.5 (1.1-2.2)	
	Immunosuppression	0.7 (0.5-1.0)	

*Area under the curve of the receiver operating characteristic (ROC) curve

7

PREDICTORS FOR INDIVIDUAL PATIENT ANTIBIOTIC TREATMENT EFFECT IN HOSPITALISED COMMUNITY-ACQUIRED PNEUMONIA PATIENTS

Antonella F. Simonetti, Cornelis H. van Werkhoven, Valentijn A. Schweitzer, Diego Viasus, Jordi Carratalà, Douwe F. Postma, Jan J. Oosterheert, Marc J.M. Bonten

ABSTRACT

OBJECTIVE

Our objective was to identify clinical predictors of antibiotic treatment effects in non-ICU hospitalised CAP patients.

METHODS

Post-hoc analysis of three prospective cohorts (from the Netherlands and Spain) of CAP adult patients admitted to a non-ICU having received either beta-lactam monotherapy (BL), beta-lactam + macrolide (BLM), or fluoroquinolone-based therapy (FQL) as empiric antibiotic treatment. We evaluated candidate clinical predictors of treatment effects in multiple mixed-effects models by including interactions of the predictors with empiric antibiotic choice and using 30-day mortality, ICU admission, and length of hospital stay (LOS) as outcomes.

RESULTS

Among 8,562 patients, empiric treatment was BL in 4,399 (51.4%), FQL in 3,373 (39.4%), and BLM in 790 (9.2%). Older age (interaction OR 1.67, 95% CI 1.23 − 2.29, p-value 0.034) and current smoking (interaction OR 2.36, 95% C.I. 1.34 − 4.17, p-value 0.046) were associated with lower effectiveness of FQL on 30-day mortality. Older age was also associated with lower effectiveness of BLM on LOS (interaction effect ratio 1.13, 95% CI 1.08 − 1.18, p-value < 0.0001).

CONCLUSIONS

Older age and smoking could influence the response to specific antibiotic regimens. The effect modification of age and smoking should be considered hypothesis generating to be evaluated in future trials.

INTRODUCTION

Community-acquired pneumonia (CAP) is a leading cause of hospitalization and death worldwide[1-3]. Although recent studies described a downward trend in 30-day mortality in hospitalized patients with CAP over the last 20 years[4-5], the reported hospital mortality in these patients remains high, ranging from 4% to 15%[4-7].

For CAP patients admitted to a non−intensive-care-unit (non-ICU), international guidelines recommend either beta-lactam monotherapy (BL), beta-lactam macrolide combination therapy (BLM) or respiratory fluoroquinolone monotherapy (FQL) as empiric treatment[8-10]. However, the necessity for atypical coverage in non-severe CAP patients is uncertain as beneficial effects on mortality were only found in observational studies, but not in randomized controlled trials[11-12]. Moreover, the use of macrolides and fluoroquinolones has been related to increased risks of antimicrobial resistance and adverse drug effects[13-17]. A limitation of the studies performed so far is that they compared interventions within the whole domain of hospitalized CAP (e.g. at the population level), lacking power for proper subgroup analyses.

Despite important advancements in diagnostic testing, a causative pathogen is not detected in the majority of CAP patients; and if detected there is often a delay of up to 48 hours[2]. Initial antibiotic treatment is therefore almost always empiric. However, CAP is a heterogeneous disease due to heterogeneity in both host and pathogen factors. Therefore, an individualized antibiotic treatment approach might prove beneficial.

The concept of individualized medicine, initially referred to the use of genomics in clinical care, has extended to recognizing the heterogeneity of each individual patient, particularly their risk factors for developing disease or having poor outcomes, and using this to inform treatment decisions. Biomarkers and clinical predictors have been widely studied in CAP in an attempt to predict the microbial etiology[18, 19] or clinical outcomes, such as early treatment failure or all-cause mortality[20-25]. Yet, predictors of pathogens are weak at best, and predictors of all-cause mortality do not inform the treating physician about the necessity to adjust empiric therapy. To pave the way for individualized medicine for CAP, it is necessary to take a step further and assess differences in treatment response based on multiple patient factors.

The objective of this study was to find candidate predictors at individual patient level for effect modification of empiric antibiotic regimens (BL, BLM and FQL) in CAP patients hospitalized to non-ICU wards.

PATIENTS AND METHODS

SETTING, STUDY POPULATION AND RESEARCH DESIGN

This is a post-hoc analysis of three cohorts of hospitalized patients with CAP, two from the Netherlands and one from Spain[4, 12, 26]. The Dutch cohorts were from two large randomized clinical trials conducted in the Netherlands. All patients hospitalized for CAP from The Community-Acquired Pneumonia immunization Trial in Adults (CAPiTA), and all patients included in the Community-Acquired Pneumonia — Study on the Initial Treatment with Antibiotics of Lower Respiratory Tract Infections (CAP-START) were included.

The Spanish (Bellvitge) cohort includes all patients with X-ray confirmed CAP admitted via the emergency department of Bellvitge University Hospital. **Supplementary table 1** shows the main characteristics of the three cohorts. For the purpose of this study, we only analysed patients who received BL, BLM or FQL as empiric antibiotic treatment.

DATA COLLECTION

Empiric antibiotic treatment was defined as the antibiotic treatment administered in the first calendar day of hospitalization (Dutch cohorts) or prospectively collected as a specific item in the data collection form (Bellvitge cohort), as the first antibiotic regimen administered to the patient after admission.

Data on clinical presentation, laboratory, microbiologic test results, antibiotic use, and clinical outcome were retrieved from medical records. In the absence of notes in clinical records, the following variables were assumed to be absent/negative: pneumococcal or influenza vaccination, clinical symptoms (cough, purulent sputum, pleuritic chest pain, headache, gastro-intestinal symptoms, chills), confusion, hypotension, tachycardia, positive urinary antigen for *S. pneumoniae*. Definitions of predictors and empiric antibiotic treatment are explained in the **Supplement**.

All studies were approved by the Institutional Review Board in the participating hospitals and the informed consent covered the current analysis. To protect personal privacy, data were anonymized.

OUTCOMES

The primary outcome was all-cause mortality within 30 days after admission. The 30-day mortality was either assessed at a long-term follow-up visit (Bellvitge), from General Practitioner (GP) medical records (Bellvitge, CAPiTA), or from the municipal records database (CAP-START). The secondary outcomes were ICU admission after the first day of hospitalization and length of hospital stay. All outcomes were measured and analyzed at the individual patient level.

PREDICTORS

Through an extensive search in PubMed we selected a list of candidate clinical predictors of treatment effects on CAP. These clinical predictors should be present and known at admission and associated either to specific CAP etiology or to clinical outcome. A complete list of the predictors chosen for the analysis and the correspondent bibliography are shown in the **Supplement**.

In addition, the year of admission was included as a confounding variable, categorized in 4 periods of 5 years each, as following: 1995-1999, 2000-2004, 2005-2009, 2010-2014.

STATISTICAL ANALYSIS

Data are presented as percentages and numbers, means with SDs, medians with interquartile ranges (IQRs), or proportions with 95% CIs, as appropriate.

For binary outcomes we used mixed-effects logistic regression models (see **Supplement** for details). To identify candidate predictors of treatment effects we applied a two-step approach. First, we estimated for each candidate predictor the interaction effect with antibiotic treatment in separate models, including the fixed effects, random effects, and the single interaction effect. Interaction variables with a two-sided p-value of <0.10 using the Wald test were included in the second step of our analysis. There we constructed a mixed-effects model including all selected interactions from the first step and all afore mentioned fixed and random effects. P-values of the second-step model were corrected for multiple testing using the Benjamini-Hochberg (BH) method[28]. Two-sided BH adjusted p-values <0.05 were considered statistically significant. Associations are given as ORs with 95% CIs. Effect modifiers for the length of hospital stay (LOS) were tested similarly with mixed-effects linear regression models, after log-transforming length of stay. The exponent of the regression coefficients was interpreted as the effect ratio, e.g. an effect ratio of 2 for factor X implies that a patient with X has a two time longer length of stay compared to a patient without X.

We performed sensitivity analyses including only patients with radiologically confirmed CAP and we performed analyses stratified per cohort. Assumptions of the models were tested visually by plotting residuals. Missing data on smoking habits (6.6% of missing data), pre-hospital antibiotics use (2.5%), elderly home living (12.4%), serum sodium concentration (12.4%), leukocyte count (0,2%), and PSI (0.1%) were imputed by multiple imputations (ten imputation datasets), assuming data missing at random. Descriptive statistics and multiple imputations were performed using the Statistical Package for the Social Sciences for

Windows (Version SPSS 21.0.0.0). Mixed-effects models were performed with R (R Core Team, 2015), and the R-package lme4 (Bates, Maechler, Bolker, Walker 2015).

RESULTS

A total of 8,562 patients were included: 2,184 (25.5%) from the CAPiTA cohort, 2,154 (25.2%) from the CAP-START cohort and 4,224 (49.3%) from the Bellvitge cohort (**supplementary figure 1**). Patient characteristics are described in **table 1**. A probable or definite microbiological diagnosis was made in 46.3% of patients. The diagnostic work-up by cohorts is described in **supplementary table 2**. The causative pathogens identified per age group are summarized in **supplementary table 3**. The majority of patients received BL as empiric treatment (4,399; 51.4%), followed by FQL (3,373; 39.4%) and BLM (790; 9.2%). The different empirical antibiotics administered in each cohort, either in monotherapy or in combination, are listed in **supplementary table 4**.

Table 1. Principal clinical characteristics and outcomes in each cohort.

	CAPiTA [26] n=2184(25.5%)	CAP-START [12] n=2154(25.2%)	BELLVIGE [4] n=4224(49.3%)	ALL n=8562
Age, years(IQR)	76.0 (72-82)	70.0 (59-79)	70.5 (58-79)	73.0 (63-80)
Male sex, n(%)	1545 (70.7)	1250 (58.0)	2860 (67.7)	5655 (66.0)
Elderly home, n(%)	81 (4.1)	102 (4.8)	234 (6.9)	417 (5.6)
Current smoker, n(%)	323 (19.0)	441 (21.1)	1037 (24.7)	1801 (22.5)
Influenza season, n(%)	1565 (71.7)	1553 (72.1)	3230 (76.5)	6348 (74.1)
S. pneumoniae vaccination, n(%)	1066 (48.8)	44 (2.0)	710 (16.8)	1820 (21.3)
Influenza virus vaccination, n(%)	1916 (87.7)	1396 (64.8)	2001 (47.4)	5313 (62.1)
Outpatient antibiotic, n(%)	656 (31.0)	639 (30.4)	882 (21.4)	2177 (26.1)
Beta-lactams, n(%)	373 (17.8)	366 (17.7)	538 (13.2)	1277 (15.5)
Atypical coverage, n(%)	296 (14.1)	251 (12.1)	327 (8.0)	874 (10.6)
Comorbidities				
Cerebrovascular disease, n(%)	278 (12.7)	221 (10.3)	343 (8.1)	842 (9.8)
COPD, n(%)	1351 (61.9)	973 (45.2)	1230 (29.1)	3554 (41.5)
Malignancy, n(%)	301 (13.8)	364 (16.9)	414 (9.8)	1079 (12.6)
Cardiovascular, n(%)	909 (41.6)	454 (21.1)	1042 (24.7)	2405 (28.1)
Immunosuppression, n(%)	235 (10.8)	210 (9.7)	337 (8.0)	782 (9.1)
Symptoms days, days (IQR)	3 (1-6)	3(1-7)	3 (2-6)	3 (1-7)
Cough, n(%)	1509 (69.1)	1776 (82.5)	3585 (84.9)	6870 (80.2)
Purulent sputum, n(%)	924 (42.3)	1247 (57.9)	2022 (47.9)	4193 (49.0)
Gastro-intestinal symptoms, n(%)	167 (7.6)	291 (13.5)	635 (15.0)	1093 (12.8)
Pleuritic chest pain, n(%)	225 (10.3)	294 (13.6)	1767 (41.8)	2286 (26.7)
Headache, n(%)	78 (3.6)	99 (4.6)	618 (14.6)	795 (9.3)
Chills, n(%)	320 (14.7)	426 (19.8)	1927 (45.6)	2673 (31.2)
Confusion, n(%)	291 (13.3)	193 (9.0)	586 (13.9)	1070 (12.5)
Fever, n(%)	786 (36.7)	1206 (57.1)	2013 (48.1)	4005 (47.5)
Hypotension, n(%)	343 (15.7)	293 (13.6)	635 (15.0)	1271 (14.8)
Heart rate > 125 bpm, n(%)	202 (9.2)	269 (12.5)	352 (8.3)	823 (9.6)
Respiratory failure, n(%)	528 (24.2)	837 (38.9)	2435 (57.6)	3800 (44.4)
Bilateral infiltrate on chest X ray, n(%)	185 (8.5)	190 (8.8)	627 (14.8)	1002 (11.7)
Pleural fluid on chest X ray, n(%)	206 (9.4)	146 (6.8)	708 (16.8)	1060 (12.4)
Positive urinary antigen for S. pneumoniae, n(%)	166 (7.6)	197 (9.1)	939 (22.2)	1302 (15.2)

Table 1 continued.

	CAPiTA [26] n=2184(25.5%)	CAP-START [12] n=2154(25.2%)	BELLVIGE [4] n=4224(49.3%)	ALL n=8562
Comorbidities				
PSI score, points (IQR)	107 (91-125)	86 (66-107)	99 (77-124)	98 (79-120)
PSI class I, n(%)	0	0	184 (4.4)	184 (2.2)
PSI class II, n(%)	34 (1.6)	644 (29,9)	672 (16.0)	1350 (15.8)
PSI class III, n(%)	506 (23.2)	556 (25.8)	859 (20.4)	1921 (22.5)
PSI class IV, n(%)	1228 (56.2)	770 (35.7)	1641 (39,0)	3639 (42.6)
PSI class V, n(%)	416 (19.0)	184 (8.5)	857 (20.3)	1457 (17.0)
Antibiotic empiric treatment				
Beta-lactam monotherapy, n(%)	1493 (68.4)	730 (33.9)	2176 (51.5)	4399 (51.4)
Beta-lactam + Macrolide, n(%)	64 (2.9)	536 (24.9)	190 (4.5)	790 (9.2)
Fluoroquinolone-based, n(%)	627 (28.7)	888 (41.2)	1858 (44.0)	3373 (39.4)
Outcomes				
30-day mortality, n(%)	195 (9.2)	114 (5.3)	261 (6.2)	570 (6.7)
Early mortality, n(%)	55 (2.5)	12 (0.6)	89 (2.1)	156 (1.8)
ICU admission, n(%)	112 (5.1)	41 (1.9)	207 (4.9)	360 (4.2)
Length Of Hospital Stay, days (IQR)	7 (5-11)	6 (4-9)	8 (5-11)	7 (5-10)

IQR: interquartile range. COPD: chronic obstructive pulmonary disease.
PSI: Pneumonia Severity Index.
FQL based treatment was defined as any regimen including a FQL (FQL in monotherapy or in combination therapy).
Early mortality: mortality for any cause in the first 48 hours from admission.
ICU: Intensive Care Unit.

CLINICAL PREDICTORS FOR TREATMENT EFFECT: 30-DAY MORTALITY

In the first step models, five interactions between a clinical predictor and antibiotic empiric treatment were significant at a p-value of <0.10 for 30-day mortality: age, current smoking, tachycardia at admission (heart rate >125 bpm), confusion at admission, and pleuritic chest pain. In the second step we tested the combination of these five interactions (**table 2**). After correction for multiple testing, the following predictors of treatment effect for 30-day mortality were statistically significant: increasing age with the use of FQL vs. BL (interaction OR 1.67, per unit increase of standardized age, 95% CI 1.23 − 2.29, BH adjusted p-value 0.034) and active smoking with the use of FQL vs. BL (interaction OR 2.36, 95% C.I. 1.34 − 4.17, BH adjusted p-value 0.046).

Table 2. 30-day mortality: difference in response to antibiotic empiric strategy by clinical predictors in the second step mixed-effects logistic regression model.

	Adjusted interaction OR (95% IC)	BH p-value for interaction
Age*BLM	1.67 (1.03-2.72)	0.282
Age*FQL	1.67 (1.23-2.29)	0.034
Smoker*BLM	1.10 (0.40-2.99)	>0.999
Smoker*FQL	2.36 (1.34-4.17)	0.046
Heart rate>125 bpm*BLM	0.36(0.11-1.20)	0.487
Heart rate>125 bpm*FQL	1.32 (0.73-2.41)	>0.999
Confusion*BLM	0.73 (0.33-1.60)	>0.999
Confusion*FQL	0.53 (0.32-0.87)	0.123
Pleuritic chest pain*BLM	2.47 (1.01-6.02)	0.282
Pleuritic chest pain*FQL	0.99 (0.53-1.83)	>0.999

FQL: fluoroquinolone-based. BLM: beta-lactam plus macrolide. BH: Benjamini − Hochberg method.

CLINICAL PREDICTORS FOR TREATMENT EFFECT: ICU ADMISSION

In the first step models, three interactions between clinical predictors and antibiotic empiric treatment were statistically significant at a p-value of <0.10 for ICU admission: admission during influenza season, having a positive urinary antigen test for *S. pneumoniae*, and leukopenia (leukocyte count less than 4000 cells/µL) or extreme leukocytosis (leukocyte count more than 20000 cells/µL) at admission. In the second step we tested the combination of these three interactions (**table 3**). After correction for multiple testing, the only statistically significant predictor of treatment effect for ICU admission was extreme leukocytosis for the use of BLM vs. BL (interaction OR 4.42, 95% CI 1.83 – 10.66, BH adjusted p-value 0.029).

Table 3. Intensive Care Unit admission: difference in response to antibiotic empiric strategy by clinical predictors in the second step mixed-effects logistic regression model.

	Adjusted interaction OR (95% CI)	BH p-value for interaction
Influenza season*BLM	0.76 (0.29-1.90)	>0.999
Influenza season*FQL	0.66 (0.37-1.16)	>0.999
S.pneumoniae+Ag*BLM	0.45 (0.09-2.19)	>0.999
S.pneumoniae+Ag*FQL	0.46 (0.25-0.84)	0.117
Leukocyte count <4000 cells/µL*BLM	3.27(0.60-17.83)	>0.999
Leukocyte count <20000 cells/µL *BLM	4.42 (1.83-10.66)	0.029
Leukocyte count <4000 cells/µL *FQL	3.71 (1.34-10.28)	0.117
Leukocyte count <20000 cells/µL *FQL	1.30 (0.69-2.46)	>0.999

FQL: fluoroquinolone-based. BLM: beta-lactam plus macrolide. BH: Benjamini – Hochberg method.

CLINICAL PREDICTORS FOR TREATMENT EFFECT: LENGTH OF HOSPITAL STAY

In the first step models, six interactions between clinical predictors and antibiotic empiric treatment were statistically significant at a p-value of <0.10 for LOS: increasing age, previous outpatient antibiotic treatment with atypical coverage, history of cardiovascular disease, presentation with gastro-intestinal symptoms, having a positive urinary antigen test for *S. pneumoniae*, and presentation with bilateral infiltrates or pleural fluid on chest X-ray. In the second step we tested the combination of these six interactions (**table 4**). After correction for multiple testing, the only statistically significant predictor of treatment effect for LOS was increasing age with the use of BLM vs. BL (interaction effect ratio 1.13 per unit increase of standardized age, 95% CI 1.08 – 1.18, BH adjusted p-value < 0.0001).

SENSITIVITY ANALYSES

Sensitivity analyses of the three final models in patients with radiologically confirmed CAP did not reveal substantial changes in the estimates of interactions (**supplementary table 4**). Subsequently, we performed the analyses in each of the three cohorts separately

(**supplementary table 4**). In the 30-day mortality model, the ORs for the interaction between increasing age and FQL use were consistent in the three cohorts, ranging from 1.62 to 1.75, while the OR for the interaction between being an active smoker and FQL use showed larger variation (1.45 to 3.97) albeit all in the same direction. In the LOS model, the effect size for the interaction between increasing age and BLM treatment ranged from 1.01 to 1.62. In the ICU admission model, the ORs for the interaction of leukocytosis with BLM use showed substantial inter-cohort differences (from 1.58 to 48.91).

Finally, since the analyses yielded similar interaction effect estimates in models without inclusion of confounders, confounding by indication appeared to be limited for the interaction effect (**supplementary table 4**).

Table 4. Length of Hospital Stay: difference in response to antibiotic empiric strategy by clinical predictors in the second step mixed-effects linear regression model.

	Adjusted interaction effect ratio (95% CI)	BH p-value for interaction
Age*BLM	1.13 (1.08-1.18)	0.00003
Age*FQL	1.03 (1.006-1.07)	0.087
Outpatient atypical coverage*BLM	0.81 (0.70-0.94)	0.074
Outpatient atypical coverage*FQL	0.92 (0.84-1.001)	0.204
History of cardiovascular disease*BLM	1.04 (0.93-1.17)	>0.999
History of cardiovascular disease*FQL	1.006 (0.94-1.07)	>0.999
Gastro-intestinal symptoms*BLM	0.87 (0.76-1.01)	0.204
Gastro-intestinal symptoms*FQL	0.93 (0.86-1.005)	0.204
S.pneumoniae+ Urinary Antigen*BLM	1.20 (1.02-1.41)	0.117
S.pneumoniae+ Urinary Antigen*FQL	1.11 (1.03-1.19)	0.076
Pleural fluid on chest X-ray*BLM	1.05 (0.90-1.22)	>0.999
Pleural fluid on chest X-ray * FQL	1.11 (1.02-1.21)	0.076
Bilateral infiltrate on chest X-ray *BLM	0.95 (0.82-1.09)	>0.999
Bilateral infiltrate on chest X-ray *FQL	1.12 (1.02-1.21)	0.076

FQL: fluoroquinolone-based. BLM: beta-lactam plus macrolide. BH: Benjamini – Hochberg method.

INDIVIDUAL PREDICTED TREATMENT EFFECT ON 30-DAY MORTALITY

Focusing on our primary outcome, we refitted the step 2 model, restricted to the significant interaction variables (increasing age and to be a current smoker), to construct a predictive model of 30-day mortality based on the provided antibiotic treatment (**figure 1**). According to this model, in older currently smoking patients empiric treatment with FQL is associated with higher 30-day mortality than empiric treatment with BL. Yet, in young non-smoking patients, FQL empiric treatment was predicted to be associated with lower 30-day mortality. There were no clear effects for BLM vs. BL.
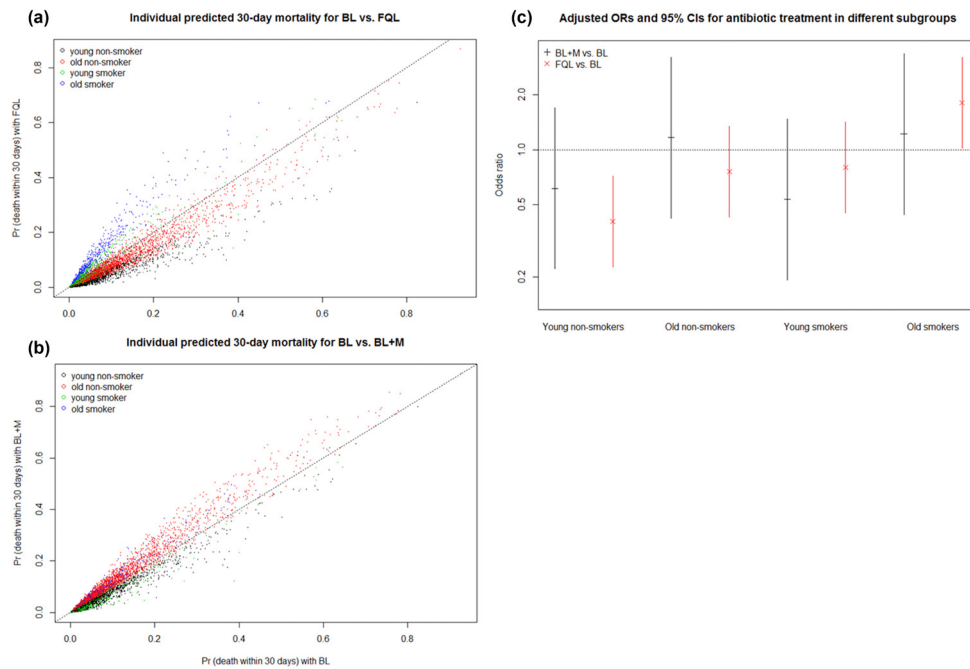
Figure 1. Predicted 30-day mortality at individual patient level. A. Individual predicted 30-day mortality in a logistic regression model restricted to the significant interaction variables (age and smoke habit), comparing patients who receive BL vs. patients who receive FQL as empiric treatment. **B.** Individual predicted 30-day mortality in a logistic regression model restricted to the significant interaction variables (age and smoke habit), comparing patients who receive BL vs. patients who receive BLM as empiric treatment. **C.** Adjusted (BH method) Odds Ratio with 95% Confidence Interval for 30-day mortality in different subgroups of patients, divided for their group age and smoke habit. FQL: fluoroquinolone-based. BLM: beta-lactam plus macrolide. BH: Benjamini – Hochberg method.

DISCUSSION

In this post-hoc analysis of three prospective cohorts from the Netherlands and Spain we identified age and smoking as candidate clinical predictors for the response to empiric antibiotic treatment, from an individualized patient perspective. In a previous clinical trial comparing BL with BLM[11] authors indicate an interaction effect of PSI high classes classification and monotherapy, with a reduced HR for clinical stability. Conversely, in a recent register-based cohort study comparing narrow vs. broad spectrum beta-lactams therapy in CAP patients, the authors did not find significant interaction effects of clinical variable with antibiotic effectiveness[29].

Our findings suggest that older age and smoking are associated with increased 30-day mortality in patients receiving FQL as empiric treatment, either alone or combined with beta-lactams. In older patients the beneficial effects of atypical coverage could be less than in younger patients partly due to a lower incidence of CAP caused by atypical pathogens, as reported in different series[19, 30, 31] and also observed in our data (**supplementary table 3**). Moreover, adverse effects and toxicity of FQL (among them the QT interval prolongation[32]) could be more pronounced in older patients, possibly due to a decline in renal function and changes in pharmacokinetics[33]. Older age was also related with decreased effectiveness of BLM, with an interaction OR of 1.67. However, presumably due to the lower number of patients with this regimen, the association was not statistically significant.

Yet, the direction of the effect of smoking was unexpected, especially in the light of studies reporting a higher proportion of smokers in *Legionella pneumophila* patients, which should, in contrast to our findings, favour fluoroquinolone-based treatment in smokers[34, 35]. This finding raises new questions about a possible interaction between smoking and antibiotic effectiveness. To the best of our knowledge, currently there is no mechanism that could explain such an interaction. We can only hypothesize that smoking patients might have malignancies, COPD, or other unexplored characteristics, which were not yet recognized and/or reported in the medical chart, which could interact with fluoroquinolone use in a detrimental way. Still, due to the large variability of the ORs between cohorts, this finding should be interpreted with caution.

Older age was related to an increase in LOS in patients who received BLM as empiric treatment, with an addition of one day on the median LOS of 7 days. As mentioned above, the lower incidence of atypical pathogens in older patients could lead to less beneficial effects of BLM in these patients. Furthermore, this finding could refer to the well described association between macrolide use and cardiac events [15, 16], which more frequently occur in older patients. Unfortunately, our data did not allow testing of this hypothesis. Moreover, we observed that the effect size of the interaction between age and BLM use was highly variable between the three cohorts, raising uncertainty on the generalizability of this finding.

Similarly, the large confidence interval of the OR and the wide range of ORs between the three cohorts for the association between ICU admission and leukocyte count over 20,000 in patients who received BLM prohibit firm conclusions.

Of note, the interaction between PSI score and empiric antibiotic treatment showed no effect on clinical outcome. In current clinical practice, the choice of empiric antibiotic

treatment is mainly based on clinical severity criteria, supported by disease severity scores such as the PSI score[8, 10]. Our findings suggest that the PSI score does not predict whether a patient will respond better to one empiric antibiotic treatment over another, suggesting that we need to re-evaluate how we select empiric antibiotics to treat CAP patients.

The key strengths of this study are the large number of patients from different cohorts allowing us to assess treatment effects in subgroup analyses, the high quality prospective data collection, and the inclusion of all possible relevant clinical predictors in the analysis. This study could serve as a prototype for future research in CAP, being the first study in using the novel approach of identifying predictors for the effect of empiric treatment strategies, instead of looking at predictors for clinical outcome or causative pathogen. One source of weakness in this study is the presence of some important differences between cohorts. In Bellvitge cohort all patient included have a confirmed CAP on chest X-ray, unlike the Dutch cohorts. Whereas radiologically confirmed CAP patients represent a more well-defined disease entity, the Dutch cohorts included all patients that are treated for a clinical diagnosis of CAP, improving generalizability of the results to daily clinical practice. However, a sensitivity analysis which included only X-ray confirmed CAP showed similar results. Furthermore, there is a large variability in the presence of some clinical signs and symptoms between the three cohorts (**table 1**), which is probably due to a lack of uniformity in the collection of clinical data. The possibility of misreporting clinical characteristics could underestimate their modifying effect on treatment and hence influence results. To correct for clustering within the cohorts, we used mixed-effects regression models. In addition, we performed a sensitivity analysis stratified by cohorts to assess the robustness of our findings in each of the cohorts.

Importantly, these are all observational data, and we could not rule out confounding by indication of the different empiric antibiotic treatments used, although we adjusted for multiple confounders in the multivariate models. Yet, as we focus on the interaction effect of clinical factors with empiric antibiotic treatment, we can postulate that the same bias is present in all the different strata, thus not largely biasing the direction and size of the interaction effect.

Moreover, as we cannot rule out bias on the direct effects of antibiotics, the same interaction effect could either mean benefit for one group, or harm for the other group. For example, we cannot claim that fluoroquinolone-based treatment is harmful in older smoking patients, as our results could be also interpreted the other way round, meaning that they are beneficial in younger and non-smoking patients. Considering this limitation,

our results should be considered hypothesis generating and need to be confirmed in a randomized controlled trial designed to estimate these interaction effects.

In conclusion, it is plausible that older age influences the response to specific antibiotic treatment, as we found a relationship with both the use of FQL and increased 30-day mortality and BLM use and LOS in older patients. Current smoking was also associated with a decreased response to FQL. Future trials evaluating antibiotic strategies for CAP could assess the treatment effects in patients of different age categories and smoking status. In addition, further research illuminating the causal mechanism underlying the identified associations needs to be performed.

SUPPORT STATEMENT

CONFLICT OF INTEREST
The authors declare no conflict of interest.

REFERENCES

1. Thomas CP, Ryan M, Chapman JD, et al. Incidence and cost of pneumonia in medicare beneficiaries. Chest 2012; 142: 973-81.

2. Jain S, Self WH,Wunderink RG, et al; CDC EPIC Study Team. Community-acquired pneumonia requiring hospitalization among US adults. N Engl J Med. 2015;373:415-427.

3. Yu H, Rubin J, Dunning S, Li S, Sato R. Clinical and economic burden of community- acquired pneumonia in the Medicare fee-for-service population. J Am Geriatr Soc 2012; 60: 2137-43.

4. Simonetti AF, Garcia-Vidal C, Viasus D, et al. Declining mortality among hospitalized patients with community-acquired pneumonia. Clin Microbiol Infect 2016;22:567.e1–7.

5. Daniel P, Woodhead M, Welham S, et al. Mortality reduction in adult community-acquired pneumonia in the UK (2009-2014): results from the British Thoracic Society audit programme. Thorax. 2016;71(11):1061-1063.

6. Woodhead M, Welch CA, Harrison DA, Bellingan G, Ayres JG. Community-acquired pneumonia on the intensive care unit: secondary analysis of 17,869 cases in the ICNARC Case Mix Programme Database. Crit Care 2006;10(Suppl. 2):S1.

7. Lee JS, NsaW, Hausmann LR, et al. Quality of care for elderly patients hospitalized for pneumonia in the United States, 2006 to 2010. JAMA Intern Med. 2014;174:1806-1814.

8. Lim WS, Baudouin SV, George RC, et al. BTS guidelines for the management of community acquired pneumonia in adults: update 2009. Thorax 2009; 64: Suppl 3: iii1- iii55.

9. Mandell LA, Wunderink RG, Anzueto A, et al. Infectious Diseases Society of America/American Thoracic Society consensus guidelines on the management of community-acquired pneumonia in adults. Clin Infect Dis 2007; 44: Suppl 2: S27-S72.

10. Wiersinga WJ, Bonten MJ, Boersma WG, et al. SWAB/NVALT (Dutch Working Party on Antibiotic Policy and Dutch Association of Chest Physicians) guidelines on the management of community acquired pneumonia in adults. Neth J Med 2012; 70: 90-101.

11. Garin N, Genné D, Carballo S, et al. -Lactam monotherapy vs -lactam-macrolide combination treatment in moderately severe community-acquired pneumonia: a randomized noninferiority trial. JAMA Intern Med. 2014;174: 1894-1901.

12. Postma DF, vanWerkhoven CH, van Elden LJ, et al; CAP-START Study Group. Antibiotic treatment strategies for community-acquired pneumonia in adults. N Engl J Med. 2015;372:1312-1323.

13. Fuller JD, Low DE. A review of Streptococcus pneumoniae infection treatment failures associated with fluoroquinolone resistance. Clin Infect Dis 2005; 41: 118-21.

14. Malhotra-Kumar S, Lammens C, Coenen S, Van Herck K, Goossens H. Effect of azithromycin and clarithromycin therapy on pharyngeal carriage of macrolide- resistant streptococci in healthy volunteers: a randomised, double-blind, placebo- controlled study. Lancet 2007; 369:482-90.

15. Mortensen EM, Halm EA, Pugh MJ, et al. Association of azithromycin with mortality and cardiovascular events among older patients hospitalized with pneumonia. JAMA. 2014;311:2199-2208.

16. Ray WA, Murray KT, Hall K, Arbogast PG, Stein CM. Azithromycin and the risk of cardiovascular death. N Engl J Med. 2012;366: 1881-1890.

17. Vanderkooi OG, Low DE, Green K, Powis JE, McGeer A; Toronto Invasive Bacterial Disease Network. Predicting antimicrobial resistance in invasive pneumococcal infections. Clin Infect Dis. 2005;40:1288-1297

18. Masiá M, Gutiérrez F, Padilla S, et al. Clinical characterisation of pneumonia caused by atypical pathogens combining classic and novel predictors. Clin Microbiol Infect. 2007 ;13:153-61.

19. Raeven VM, Spoorenberg SM, Boersma WG, et al. Atypical aetiology in patients hospitalised with community-acquired pneumonia is associated with age, gender and season; a data-analysis on four Dutch cohorts. BMC Infect Dis. 2016;16:299.

20. Lim WS, van der Eerden MM, Laing R, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. Thorax. 2003;58:377-382.

21. Fine MJ, Auble TE, Yealy DM, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. N Engl J Med. 1997;336:243-250.

22. Rosón B, Carratalà J, Fernández-Sabé N, Tubau F, Manresa F, Gudiol F. Causes and factors associated with early failure in hospitalized patients with community-acquired pneumonia. Arch Intern Med. 2004;164:502-8.

23. Hoogewerf M, Oosterheert JJ, Hak E, Hoepelman IM, Bonten MJ. Prognostic factors for early clinical failure in patients with severe community-acquired pneumonia. Clin Microbiol Infect. 2006;12:1097-104.

24. Garcia-Vidal C, Fernández-Sabé N, Carratalà J, Díaz V, Verdaguer R, Dorca J, Manresa F, Gudiol F. Early mortality in patients with community-acquired pneumonia: causes and risk factors. Eur Respir J. 2008;32:733-9.

25. Kolditz M, Ewig S, Klapdor B, Schütte H, Winning J, Rupp J, Suttorp N, Welte T, Rohde G; CAPNETZ study group. Community-acquired pneumonia as medical emergency: predictors of early deterioration. Thorax. 2015;70:551-8.

26. Bonten MJM, Huijts SM, Bolkenbaas M et al. Polysaccharide conjugate vaccine against pneumococcal pneumonia in adults. N. Engl. J. Med.2015; 372:1114–1125.

27. Schouten JA, Prins JM, Bonten MJ, Degener J, Janknegt RE, Hollander JM, Jonkers RE, Wijnands WJ, Verheij TJ, Sachs AP, Kullberg BJ; Dutch Working Party on Antibiotic Policy. Revised SWAB guidelines for antimicrobial therapy of community-acquired pneumonia. Neth J Med. 2005;63:323-35.

28. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 1995;57: 289–300.

29. Rhedin S, Galanis I, Granath F, Ternhag A, Hedlund J, Spindler C, et al. Narrow-spectrum ß-lactam monotherapy in hospital treatment of community-acquired pneumonia: a register-based cohort study. Clin Microbiol Infect. 2016 Dec 23. doi: 10.1016/j.cmi.2016.12.015. [Epub ahead of print].

30. Klapdor B, Ewig S, Pletz MW, Rohde G, Schütte H, Schaberg T, Welte T; CAPNETZ Study Group. Community-acquired pneumonia in younger patients is an entity on its own. Eur Respir J. 2012;39:1156-61.

31. Torres A, Blasi F, Peetermans WE, Viegi G, Welte T. The aetiology and antibiotic management of community-acquired pneumonia in adults in Europe: a literature review. Eur J Clin Microbiol Infect Dis. 2014;33:1065-79.

32.   Briasoulis A, Agarwal V, Pierce WJ. QT prolongation and torsade de pointes induced by fluoroquinolones: infrequent side effects from commonly used medications. Cardiology 2011;120:103−110.

33.   Stahlmann R., Lode H. Safety considerations of fluoroquinolones in the elderly. An update. Drugs Aging 2010; 27: 193-209

34.   Almirall J, Blanquer J, Bello S. Community-acquired pneumonia among smokers. Arch Bronconeumol. 2014 Jun;50(6):250-4.

35.   Fernández-Sabé N1, Rosón B, Carratalà J, Dorca J, Manresa F, Gudiol F. Clinical diagnosis of Legionella pneumonia revisited: evaluation of the Community-Based Pneumonia Incidence Study Group scoring system. Clin Infect Dis. 2003 Aug 15;37:483-9.

8

INAPPROPRIATE USE OF ANTIMICROBIALS FOR LOWER RESPIRATORY TRACT INFECTIONS IN ELDERLY PATIENTS: PATIENT- AND COMMUNITY-RELATED IMPLICATIONS AND POSSIBLE INTERVENTIONS

Inger van Heijl, Valentijn A. Schweitzer, Lufang Zhang, Paul D. van der Linden, Cornelis H. van Werkhoven, Douwe F. Postma

PART II | CHAPTER 8

INAPPROPRIATE USE OF ANTIMICROBIALS FOR LOWER RESPIRATORY TRACT INFECTIONS IN ELDERLY PATIENTS: PATIENT- AND COMMUNITY-RELATED IMPLICATIONS AND POSSIBLE INTERVENTIONS

ABSTRACT

Elderly are more susceptible to infections which is reflected in the incidence and mortality of lower respiratory tract infections (LRTI) increasing with age. Several aspects of antimicrobial use for LRTIs in elderly patients should be considered to determine appropriateness. We discuss possible differences in microbial etiology between elderly and younger adults, definitions of inappropriate antimicrobial use for LRTIs currently found in the literature and their results, possible negative impact of antimicrobial therapy on an individual and community level, and finally, we propose that both antimicrobial stewardship interventions and novel rapid diagnostic techniques may optimize antimicrobial use in elderly patients with LRTI.

KEY WORDS

Inappropriate antimicrobial use, lower respiratory tract infections, elderly

KEY POINTS

- Reports on (in)appropriate antimicrobial use lack a reference standard for defining and measuring appropriateness of treatment.
- Quinolones or macrolides should be restricted to selected cases empirically given the low incidence of atypical pathogens in elderly patients and higher risks of adverse drug events and drug-drug interactions.
- Use of low-dose CT scanning, point-of-care ultrasonography, or PoC-PCR testing for viral pathogens are promising research areas to decrease the inappropriate use of antimicrobials.

INTRODUCTION

Elderly people (adults over 65 years of age) comprised a fifth of the total population in Europe in 2016; this proportion may further increase to 25% in 2030[1, 2]. Elderly are more susceptible to infections and their sequelae than younger adults[3, 4]. which is reflected in the incidence and mortality of lower respiratory tract infections (LRTIs) increasing with age[5-7]. Next to increased incidences of comorbidities, it is thought that age-related altered immune regulation, often referred to as 'immunosenescence', also contributes to this[8], Several aspects of LRTIs in elderly make it increasingly difficult to determine the most appropriate antimicrobial therapy for individual patients. First, the etiology of LRTI in elderly patients could be slightly different compared to younger adults, which would require adjusted empirical antimicrobial therapy. Also, diagnosis of LRTI in elderly could be more challenging which might lower the threshold for prescribing antimicrobials. Furthermore, with advancing age, the human body changes in composition and organ function, resulting in alteration of pharmacokinetics and pharmacodynamics of antimicrobials[9]. When combined with the increasing frequency of co-morbidities and/ or polypharmacy, this facilitates the occurrence of adverse drug events (ADE) and drug-drug interactions.

We will discuss the etiology, the currently used definitions for appropriate use of antimicrobials, and the different negative consequences of antimicrobial therapy in individual patients and the community for LRTIs in elderly patients. Finally, we propose targeted interventions to improve antimicrobial prescribing in these patients.

MICROBIOLOGICAL ETIOLOGY

Seven studies, all from Europe, have made head-to-head comparisons of etiology in elderly and younger adult patients with LRTI[10-16]. A cut-off of 65 years of age was used to define categories. The ranges of the most commonly identified pathogens are summarized in Table 1. *Streptococcus pneumoniae* was the most frequently identified pathogen in both age groups. The most discernible differences between the two groups were that gram-negatives, especially Enterobacteriaceae, were found more frequent in elderly, whereas certain atypical pathogens (*Legionella pneumophila, Mycoplasma pneumoniae,* and *Coxiella burnetti*) were more frequent in younger patients. Increasing age and nursing-home residency are associated with colonization of gram-negative bacteria in the oropharynx. Micro-aspiration is considered an etiologic pathway in the development of LRTIs in elderly patients, which could explain the increase in gram-negatives[10, 17, 18]. The lower prevalence of atypical pathogens in elderly may be caused

PART II | CHAPTER 8

INAPPROPRIATE USE OF ANTIMICROBIALS FOR LOWER RESPIRATORY TRACT INFECTIONS IN ELDERLY PATIENTS:
PATIENT- AND COMMUNITY-RELATED IMPLICATIONS AND POSSIBLE INTERVENTIONS

by less frequent exposure to risk factors[10].

Table 1. Identified causal pathogen in studies with head to head comparison between younger and elderly patients

Pathogens	Young patients ᵃ*	Elderly patients ᵇ*	References
Streptococcus pneumoniae	9 – 35%	8.6 – 36%	[10-16]
Staphylococcus aureus	0.3 – 4%	0.0 – 5%	[11-13, 15, 16]
Haemophilus influenzae	1 – 2%	0.7 – 10%	[10-13, 15, 16]
Gram-negatives	0 – 7%	1.4 – 15%	[10-13, 15]
Enterobacteraceae	0.4% – 1.3%	0.9 – 2.6%	[13, 16]
Atypical pathogens	11-37%	1-15%	[10, 11, 13-15]
Legionella pneumophila	3.4% – 5.2%	1 – 5%	[10-16]
Mycoplasma pneumoniae	2.8 – 15%	0 – 3.2%	[10-16]
Coxiella burnetti	0.7 – 15.8%	0 – 3.5%	[10, 13-15]
Chlamydia pneumoniae	0.1 – 8.2%	0 – 6.7%	[10, 12-16]
Total viral pathogens	3.6 – 4%	4.5 – 13.4%	[10, 11, 14, 15]
Influenza	1.2 – 3.0%	0.3 – 4.8%	[10, 12, 14-16]
Parainfluenza	1.3%	1 – 8.6%	[10, 14]
RSV	0.0 – 0.4%	0.7 – 1.8%	[10, 12, 15]
Unknown	24 – 79%	40 – 80%	[10-16]

* Range of rapported prevalances of pathogens found in the literature

ᵃ < 65 years, exception van Vught et al. [11] < 50 years. Fernández Sabé et al. [10] has been excluded for this specific younger age group as their cut-off was 80 years. Otherwise this younger age group would also contain patients aged 65-80 years.

ᵇ > 65 years, exceptions are Fernández Sabé et al. [10] > 80 years, Van Vught et al [11] > 80 years

Higher incidences of viral pathogens in elderly patients have been reported,[12, 14, 16] although this was not demonstrated consistently[10, 11, 15]. Studies showing higher incidences of viral pathogens were mostly population based, including outpatients from general practitioners, while contradicting studies were hospital based, which may explain conflicting results. Also, bias could have been introduced because viral testing was not uniformly performed and age may have influenced the decision to test[19]. Two recent cohort studies of hospitalized community-acquired pneumonia (CAP) patients, one from the US (N=2,259) and one from the Netherlands (N=408) routinely performed testing and found a virus as the only pathogen in 22% and 13% of cases, respectively. In both studies the most common detected viral pathogens were human rhinovirus and influenza virus[20, 21].

The occurrence of a viral etiology was similar for young and elderly patients in the US study[20] The study from the Netherlands even found more viral pathogens with increasing age[21]. The contribution of viral pathogens as a cause for CAP in all age groups might be larger than previously thought.

Elderly patients more often had unknown etiology in LRTI compared to younger adults[10, 11, 13, 15, 16]. Possible reasons for less detection of the causal pathogen could be a more conservative diagnostic approach in this group of patients or difficulty in obtaining good quality material for culturing[13].

In recent years, Health-Care Associated Pneumonia (HCAP) has been proposed as a separate entity from CAP. HCAP may be more frequently caused by gram-negatives and MDRO, and is associated with higher mortality[22-25]. Yet, a meta-analysis of 24 studies concluded that HCAP criteria do not accurately predict MDRO, although low study quality and heterogeneous designs preclude a firm conclusion[22]. European studies tend to report community-like etiology, while studies from Asia and the US show increased MDRO rates in HCAP patients[25-28]. To date, the clinical relevance of HCAP remains unclear[23, 24]. In fact both IDSA and ESCMID guidelines do not address HCAP which leaves a gap in the recommendations regarding treatment for patients from the long-term care facilities (LTCFs).

The microbial etiology does not justify routine empirical coverage of *Legionella pneumophila* as the low incidence is further decreased in elderly patients to around 1-5% (Table 1). As members of our group have suggested before, beta-lactam monotherapy, preferably aminopenicillins, should generally be the first choice of empirical therapy. Naturally, the severity of disease, local epidemiologic data, prior cultures or known colonization of individual patients, comorbidities, or allergies could lead to another antibiotic choice. Doxycyclin, the addition of a macrolide to beta-lactam therapy, or the newer 4th generation fluoroquinolones are potent therapies, but higher risks of adverse drug events and drug-drug interactions should be considered.

INAPPROPRIATE USE OF ANTIMICROBIALS

Whether to start antimicrobial therapy for LRTI and choosing the specific class depends on multiple factors in daily clinical practice. Most factors in this decision pathway are dependent on clinical judgement, which interferes with evaluating appropriateness in a standardized way. Deviation from protocol for empirical therapy, definitive drug selection, and duration of therapy might be justifiable for individual patients for reasons that are not captured in guidelines.

These difficulties are also reflected in the different definitions of (in)appropriate antimicrobial therapy found in the literature. Some studies have assessed (in) appropriate antimicrobial therapy by evaluating empirical therapy and/or definitive drug selection through expert opinion[32-37]. Others have aimed at appropriateness of diagnosis, dosage, route of administration, or duration of antimicrobial therapy[38-42].

PART II | CHAPTER 8

INAPPROPRIATE USE OF ANTIMICROBIALS FOR LOWER RESPIRATORY TRACT INFECTIONS IN ELDERLY PATIENTS:
PATIENT- AND COMMUNITY-RELATED IMPLICATIONS AND POSSIBLE INTERVENTIONS

A less subjective method to assess appropriateness is to evaluate therapy according to in-vitro susceptibilities, yet this requires positive microbiological testing results and cannot be solely relied upon[32-35, 37]. Another method focuses on costs, defining inappropriateness as unnecessary use of combination therapy with same spectrum[43]. Lastly, the indication for starting antimicrobial therapy, i.e. unnecessary antimicrobial therapy, can be evaluated, where inappropriateness did seem to increase with age[44]. The (in)appropriateness criteria for several studies specifically addressing antimicrobials for LRTIs in elderly patients are summarized in Table 2. Generally, the proportion of appropriate antimicrobial treatment ranged from 60 to 80%. Although none of these studies found an association with increasing age, it has been suggested that both nursing-home or LTCF residency and polypharmacy, all occurring more frequently with advancing age, increase the risk for inappropriate prescriptions[45]. For example, the point prevalence of antimicrobial use in nursing homes is about 10% and the proportion of prescribed antimicrobial courses deemed unnecessary or inappropriate after post-hoc review ranged from 20 to 75%[46-51].

It is clear that a reference standard for measuring inappropriate antimicrobial use is currently lacking, which was also concluded by a specific review of this subject [52]. More work on definitions and standardization is heavily needed to ensure evaluation of appropriateness will become more reliable and less dependent of the inter-observer variation inherent to expert evaluation[53].

Although it has been suggested that certain potential inappropriate prescriptions could be automatically recognized by an electronic health records system today, we think this is unfeasible as human judgement is almost always needed[54]. For example, a patient with documented allergy to guideline recommended treatment could be automatically flagged by such a system as appropriate deviation, yet this is dependent on accurate interpretation and registration of allergy data, the latter often being incorrect[55].

Despite the heterogeneity in definitions, and resulting heterogeneity in inappropriateness rates, there is a clear consensus on the adverse consequences of inappropriate antimicrobial use on the patient and community level[56, 57]. Both will be addressed in the next section.

Table 2. Examples from literature of different definitions of (in)appropriate antimicrobial use for LRTI in elderly .

Setting	Definition of inappropriate antimicrobial use	Appropriateness of RTI treatment	References
Tobia et al. 2008 Outpatients at emergency department N = 153	Medication Appropriateness Index (MAI): 1. indication (e.g. presence of symptoms) 2. effectiveness 3. dosage 4. directions (e.g. route) 5. practicality (e.g. adherence) 6. drug-drug interactions 7. drug-disease interactions 8. unnecessary duplication 9. duration 10. expensiveness (least expensive alternative) Rating: A: appropriate, B: marginal and C: inappropriate	n (%): 99 (65%)	[97]
Van Buul et al. 2015 Long-term care facility N = 208	• Algorithm for RTI based on guidelines and national expert panel • Distinction between: (1) acute cough, (2) or no acute cough but fever, (3) or no cough and fever. • Then presence/ absence abnormalities on lung auscultation, COPD, CRP results, other airway and not airway symptoms and certain risk factors Rating: A: appropriate, B: probably appropriate, C: probably inappropriate, D: inappropriate, E: insufficient information	n (%; range) 180 (86.5%; 60.0- 96.2)	[42]
Vergidis et al. 2011 Long-term care facility N = 752	Appropriate (with/without antimicrobial prescription): • With: when effective drug was used • Without: when use of an antimicrobial was not indicated Inappropriate: • With: when a more-effective drug was indicated • Without: Unjustified: • With: use of any antimicrobial was not indicated • Without: when use of an antimicrobial was indicated Insufficient information for categorization.	n (%): 592 (79%)	[51]
Loeb et al. 2001 Long-term care facility N = 646	• Assessment of prescriptions to see if they fulfilled diagnostic criteria • At least 3 of the following: (1) new/increased cough, (2) new/ increased sputum production, (3) fever, (4) pleuritic chest pain, (5) new or increased physical findings on chest exam, (6) new/ increased shortness of breath or respiratory rate more than 25/ min or worsening mental or functional status.	n (%): 375 (58%)	[50]

NEGATIVE IMPACT OF ANTIMICROBIALS IN ELDERLY

Treating patients with antimicrobials is not without risk. Antimicrobials can have a negative impact on both individual users and the community. For individuals there is a risk of different adverse drug events (ADEs) including drug-drug interactions and drug-comorbidity interactions. In elderly patients, the age-related changes in pharmacokinetics, frequent concurrent use of medication, and higher prevalence of co-morbidities, all contribute to an increased risk of such events. Faulkner et al. gives an extensive review of antimicrobial ADEs in elderly[58]. In the community, there is a risk of

PART II | CHAPTER 8

INAPPROPRIATE USE OF ANTIMICROBIALS FOR LOWER RESPIRATORY TRACT INFECTIONS IN ELDERLY PATIENTS:
PATIENT- AND COMMUNITY-RELATED IMPLICATIONS AND POSSIBLE INTERVENTIONS

spread of *Clostridium difficile* infections (CDIs) and antimicrobial resistance.

ADVERSE DRUG EVENTS (ADE)

In a cohort study evaluating ADEs in hospitalized patients (N=1488; median age 59 years) 20% experienced at least one ADE, with gastro-intestinal, renal, and hematologic abnormalities being the most frequent. Furthermore, 20% of the reported antimicrobial-related ADEs were due to unnecessary antimicrobial use[59]. In another cohort evaluating ADEs in nursing home residents, antimicrobials were the second most often reported drug-class to cause an ADE (20%) after antipsychotics. The majority of observed ADEs were rashes and CDIs[60].

Certain ADEs are serious enough that elderly patients have to visit the emergency room (ER). 10.6% of elderly patient ER visits was due to an ADE, with antimicrobials as one of the most frequently implicated medication classes (16.7% of all ADEs and 25% of definite or probable ADEs)[58, 61]. The most frequent (serious) ADEs in elderly who use beta-lactam antimicrobials for LRTI are drug fever, interstitial nephritis, and blood dyscrasias. For macrolides it is QT prolongation, and for trimethoprim-sulfamethoxazole it is hyperkalemia, blood dyscrasias, and drug fever[58]. Certain very rare ADEs, such as tendinitis and tendon rupture during fluoroquinolone use, are relatively more frequent in elderly and can be further increased by concomitant glucocorticoid use or renal failure[62, 63], although the absolute risks remain low[62].

DRUG-DRUG INTERACTIONS

Elderly have an increased risk for drug-drug interactions since polypharmacy is more frequent[4]. It is estimated that 51% (1380/2707) of elderly patients receive ≥ 6 drugs per day[64]. The most serious drug-drug interactions in elderly using antimicrobials for LRTIs are: trimethoprim-sulfamethoxazole in combination with vitamin K antagonists (increases anticoagulant effect), trimethoprim-sulfamethoxazole with potassium-sparing diuretics (risk of hospitalization due to hyperkaliemia), and clarithromycin/erythromycin with drugs which are deactivated by Cytochrome P450 3A4 enzymes (e.g. risk of rhabdomyolysis by increased concentrations of atorvastatin)[58].

DRUG-COMORBIDITY INTERACTIONS

Macrolides are associated with an increased risk for cardiovascular events and deaths; especially in (elderly) patients with a higher baseline risk for cardiovascular events[65-70]. These associations should caution the use of macrolides for LRTIs unless strictly necessary[9]. Renal insufficiency is more frequent in elderly, which results in a decreased elimination of certain (hydrophilic) antimicrobials (e.g. cephalosporin's, fluoroquinolones) and increases the risk of ADEs. Therefore, a dose adjustment is recommended for patients with impaired renal function[9, 71].

CLOSTRIDIUM DIFFICILE INFECTION

Broad-spectrum antimicrobials disturb the gastrointestinal flora, contributing to an increased risk of CDI. Increased age, recent hospitalization, immune suppression, malignancy, chronic renal failure, and use of proton pump inhibitors have been identified as independent risk factors of CDI, and are highly prevalent in elderly[71, 72]. Macrolides are more strongly associated with CDI than doxycycline, physicians may therefore choose for the latter when atypical coverage is deemed necessary[71]. Still, the risk of developing CDI with macrolides appears smaller than with fluoroquinolones, clindamycin, or broad-spectrum beta-lactams[73].

ANTIMICROBIAL RESISTANCE

There is clear consensus that inappropriate antimicrobial use contributes to antimicrobial resistance, potentially leading to increased morbidity, mortality, and health-care costs[74, 75]. Antimicrobial resistance rates may increase with age, as reported in a Canadian surveillance study, further increasing risk for the elderly population[76]. This could be explained by elderly people more often residing in LTCFs, needing hospitalization, receiving healthcare at home, and receiving antimicrobials, which are all risk factors for developing antimicrobial resistance[71, 74, 77-79].

OPTIMIZING APPROPRIATE ANTIMICROBIAL USE IN ELDERLY PATIENTS

ANTIMICROBIAL STEWARDSHIP INTERVENTIONS

Antimicrobial stewardship interventions aim at reducing inappropriate use of antimicrobials while maintaining good clinical outcomes. Elderly patients might especially benefit from antimicrobial stewardship as they may have the highest risk for worse clinical outcomes due to both over-treatment (e.g. antimicrobial side-effects, CDI) or under-treatment (e.g. infectious complications). The risk for and possible harm due to treatment differs per patient and depends on patient factors such as co-morbidities, immunological status, co-medication, previous antibiotic treatment, recent hospitalizations and the severity of the LRTI. Therefore, an individualized approach where individual patient risks are balanced to possible collateral damage in the form of selecting for antimicrobial resistance is recommended.

PART II | CHAPTER 8

INAPPROPRIATE USE OF ANTIMICROBIALS FOR LOWER RESPIRATORY TRACT INFECTIONS IN ELDERLY PATIENTS:
PATIENT- AND COMMUNITY-RELATED IMPLICATIONS AND POSSIBLE INTERVENTIONS

Commonly used antimicrobial stewardship interventions include periodic or individual patient audit and feedback, decision support, education (educational meetings, educational materials), and formulary restriction of antimicrobials[74, 80].

To date, the majority of studies on stewardship interventions are performed on adult or pediatric patients, or on specific hospital settings e.g. ICU. In a recent systematic review of stewardship interventions in hospitalized patients, only 1.8% (4/211) of studies specifically targeted elderly patients[74, 81-84]. Two controlled before-after studies showed a reduction in 30-day mortality after introducing a pneumonia guideline with clinical decision support[81, 82]. Two interrupted-time series showed a reduction in the incidence of CDI after implementation of an audit and feedback program and a restrictive antimicrobial policy[83, 84]. Although promising, these four studies used non-randomized designs, risking confounding bias. There is a great need for high quality studies in elderly patients. Currently, the effects of antimicrobial stewardship interventions in elderly patients may be over- or underestimated.

As stated earlier the appropriate prescription of antimicrobials in LTCF is often impeded by the frailty of elderly residents, limited clinical evaluation –sometimes only consisting of nurse's assessment and telephone supervision by a physician, and lack of diagnostic testing. In nursing homes, the majority of antimicrobials are prescribed after telephone contact with nursing staff, highlighting the importance of involving nurses in antimicrobial stewardship programs[85]. In a recent review of stewardship interventions in LTCF, the following approaches were identified to be the most effective: multidisciplinary education; pre-prescriptive data collection tools; post-prescriptive prescriber recommendations; and introducing consultation by infectious diseases experts[86]. When designing antimicrobial stewardship interventions, it is important to consider the setting and corresponding prescription process. This was illustrated in nursing home interventions designed to improve communication which were only effective in reducing antimicrobial use when nursing staff was involved[87, 88]. Appealing antimicrobial stewardship targets to improve antimicrobial use for LRTI in LTCF include improving the indication for starting antimicrobials, optimizing the use of available diagnostics, limiting the use of fluoroquinolones given their strong association with CDI, and ensuring proper dosing and duration of antimicrobial therapy[87, 89].

NEW DIAGNOSTIC TECHNIQUES
Imaging
As infections can be difficult to diagnose in elderly patients, they may be initially treated with broad spectrum antimicrobials. Novel diagnostic techniques to establish the diagnosis of LRTI may particularly reduce unnecessary antimicrobials in elderly patients. The current cornerstone for the radiological diagnosis of pneumonia is the demonstration of an infiltrate by conventional chest X-ray. However, the estimated sensitivity of a chest X-ray is only 60-70% in patients with a clinical diagnosis of CAP[31, 90, 91]. In addition, certain co-morbidities that are common in elderly patients (e.g. heart failure and chronic obstructive pulmonary disease) may impede the detection of pulmonary infiltrates[92]. The increasing availability of CT-scans and possibility of low- or ultra-low-dose scanning seems a promising alternative. Recently, CT-scans changed the diagnosis of 58.6% (95% confidence interval, 53.2−64.0%) of consecutive CAP patients, potentially leading to optimization of management in 25% of patients[93]. These results need to be validated in clinical practice to demonstrate improved patient outcomes while reducing antimicrobial use for non-infectious alternative diagnoses. In settings where CT-scans are not readily available, chest ultrasonography may be a valuable alternative. It has high diagnostic accuracy for pneumonia, can be done at the bedside, is highly reproducible in trained professionals, and costs are relatively low[94].

Microbiological testing
Established microbiological techniques to determine the causative pathogen include respiratory cultures or PCR, blood cultures, and urinary antigen tests for Legionella and pneumococcus. However, the sensitivity of these tests are limited and in 60-70% of patients suspected of CAP no causative pathogen will be identified[20, 31].

When viral CAP is suspected, point-of-care PCR (PoC-PCR) for respiratory viruses might play a role in optimizing antimicrobial therapy. Previous studies using regular respiratory PCR have failed to show an effect on antimicrobial use, possibly because it is difficult to stop antimicrobials after they have been given for 1-2 days when PCR results become available[95, 96]. PoC-PCR may allow withholding or rapid discontinuation of antimicrobials if a viral pathogen is identified, as results can be available in 1-2 hours. The effects and cost-effectiveness of PoC-PCR on antimicrobial use and patient outcome have not yet been investigated, but a cluster-randomized clinical trial evaluating the clinical effect of both low-dose CT and PoC-PCR in patients admitted to non-ICU ward with CAP (ClinicalTrials.gov number, NCT01660204) is underway.

CONCLUSION

We have addressed several issues on the appropriate use of antimicrobials in elderly patients with LRTIs. As the microbial etiology is only slightly different compared to the younger population, the mainstay of treatment should consist of beta-lactam monotherapy. Extended coverage of gram-negatives could be considered in LTCF or

PART II | CHAPTER 8

INAPPROPRIATE USE OF ANTIMICROBIALS FOR LOWER RESPIRATORY TRACT INFECTIONS IN ELDERLY PATIENTS:
PATIENT- AND COMMUNITY-RELATED IMPLICATIONS AND POSSIBLE INTERVENTIONS

nursing home residents. Quinolones or macrolides should be restricted to selected cases empirically given the low incidence of atypical pathogens in elderly patients and higher risks of ADEs and drug-drug interactions. Next to these adverse drug events, inappropriate use of antimicrobials contributes to CDI and antimicrobial resistance. A reference standard for measuring inappropriate antimicrobial use is currently lacking. Future work on definitions and standardization will hopefully increase validity and generalizability of reports on (in)appropriate antimicrobial therapy. Well-designed antimicrobial stewardship interventions could improve antimicrobial prescribing, but studies specifically targeting elderly patients are needed. These programs should generally consist of multiple components depending on the specific clinical setting, such as improving diagnostics (i.e. indication for starting antimicrobials) and ensuring proper dosing and duration of therapy. We argue that new diagnostic techniques such as low-dose CT-scanning or PoC-PCR testing for viral pathogens could potentially reduce inappropriate use of antimicrobials. Such techniques and interventions can hopefully decrease the inappropriate use of antimicrobials in the near future.

REFERENCES

1. United Nations, Department of Economic and Social Affairs, Population Division. World Population Ageing 2015. 2015;(ST/ESA/SER.A/390).

2. World Bank Group. Development Knowledge: Data on population ages 65 and above . In: . 2016. Accessed November 29 2017.

3. Meyer K. The role of immunity and inflammation in lung senescence and susceptibility to infection in the elderly. Semin Respir Crit Care Med. 2010; doi: 10.1055/s-0030-1265897.

4. Yoshikawa TT. Epidemiology and unique aspects of aging and infectious diseases. Clin Infect Dis. 2000; doi: 10.1086/313792.

5. Millett ERC, Quint J, Smeeth L, Daniel R, Thomas S. Incidence of community-acquired lower respiratory tract infections and pneumonia among older adults in the United Kingdom: a population-based study. PLoS ONE. 2013; doi: 10.1371/journal.pone.0075131.

6. Welte T, Torres A, Nathwani D. Clinical and economic burden of community-acquired pneumonia among adults in Europe. Thorax. 2012; doi: 10.1136/thx.2009.129502.

7. Janssens J, Krause K. Pneumonia in the very old. Lancet Infect Dis. 2004; doi: 10.1016/S1473-3099(04)00931-4.

8. Pawelec G. Hallmarks of human "immunosenescence": adaptation or dysregulation?. Immun Ageing. 2012; doi: 10.1186/1742-4933-9-15.

9. Sera L, McPherson M. Pharmacokinetics and pharmacodynamic changes associated with aging and implications for drug therapy. Clin Geriatr Med. 2012; doi: 10.1016/j.cger.2012.01.007.

10. Fernández Sabé N, Carratalà J, Rosón B, Dorca J, Verdaguer R, Manresa F, Gudiol F. Community-acquired pneumonia in very elderly patients: causative organisms, clinical characteristics, and outcomes. Medicine (Baltimore). 2003; doi: 10.1097/01.md.0000076005.64510.87.

11. van Vught LA, Endeman H, Meijvis SC, Zwinderman AH, Scicluna BP, Biesma DH, van der Poll T. The effect of age on the systemic inflammatory response in patients with community-acquired pneumonia. Clin Microbiol Infect. 2014; doi: 10.1111/1469-0691.12717.

12. Kothe H, Bauer T, Marre R, Suttorp N, Welte T, Dalhoff K. Outcome of community-acquired pneumonia: influence of age, residence status and antimicrobial treatment. Eur Respir J. 2008; doi: 10.1183/09031936.00092507.

13. Sahuquillo Arce J, Menéndez R, Méndez R, Amara Elori I, Zalacain R, Capelastegui A, Aspa J, Borderías L, Martín Villasclaras J, Bello S, Alfageme I, de Castro F, Rello J, Molinos L, Ruiz Manzano J, Torres A. Age-related risk factors for bacterial aetiology in community-acquired pneumonia. Respirology. 2016; doi: 10.1111/resp.12851.

14. Capelastegui A, España P, Bilbao A, Gamazo J, Medel F, Salgado J, Gorostiaga I, Lopez de Goicoechea, Maria Jose, Gorordo I, Esteban C, Altube L, Quintana J. Etiology of community-acquired pneumonia in a population-based study: link between etiology and patients characteristics, process-of-care, clinical evolution and outcomes. BMC Infect Dis. 2012; doi: 10.1186/1471-2334-12-134.

15. Gutiérrez F, Masiá M, Rodríguez JC, Mirete C, Soldán B, Padilla S, Hernández I, De Ory F, Royo G, Hidalgo

PART II | CHAPTER 8

INAPPROPRIATE USE OF ANTIMICROBIALS FOR LOWER RESPIRATORY TRACT INFECTIONS IN ELDERLY PATIENTS: PATIENT- AND COMMUNITY-RELATED IMPLICATIONS AND POSSIBLE INTERVENTIONS

AM. Epidemiology of community-acquired pneumonia in adult patients at the dawn of the 21st century: a prospective study on the Mediterranean coast of Spain. Clin Microbiol Infect. 2005; doi: 10.1111/j.1469-0691.2005.01226.x.

16. Klapdor B, Ewig S, Pletz M, Rohde G, Schütte H, Schaberg T, Welte T. Community-acquired pneumonia in younger patients is an entity on its own. Eur Respir J. 2012; doi: 10.1183/09031936.00110911.

17. Marik PE. Aspiration pneumonitis and aspiration pneumonia. N Engl J Med. 2001; doi: 10.1056/NEJM200103013440908.

18. Reza Shariatzadeh M, Huang J, Marrie T. Differences in the features of aspiration pneumonia according to site of acquisition: community or continuing care facility. J Am Geriatr Soc. 2006; doi: 10.1111/j.1532-5415.2005.00608.x.

19. Moons KGM, Royston P, Vergouwe Y, Grobbee D, Altman D. Prognosis and prognostic research: what, why, and how?. BMJ. 2009; doi: 10.1136/bmj.b375.

20. Jain S, Self W, Wunderink R, Fakhran S, Balk R, Bramley A, Reed C, Grijalva C, Anderson E, Courtney DM, Chappell J, Qi C, Hart E, Carroll F, Trabue C, Donnelly H, Williams D, Zhu Y, Arnold S, Ampofo K, Waterer G, Levine M, Lindstrom S, Winchell J, Katz J, Erdman D, Schneider E, Hicks L, McCullers J, Pavia A, Edwards K, Finelli L. Community-Acquired Pneumonia Requiring Hospitalization among U.S. Adults. N Engl J Med. 2015; doi: 10.1056/NEJMoa1500245.

21. Huijskens EGW, Koopmans M, Palmen FMH, van Erkel, Adriana J M, Mulder PGH, Rossen JWA. The value of signs and symptoms in differentiating between bacterial, viral and mixed aetiology in patients with community-acquired pneumonia. J Med Microbiol. 2014; doi: 10.1099/jmm.0.067108-0.

22. Chalmers J, Rother C, Salih W, Ewig S. Healthcare-associated pneumonia does not accurately identify potentially resistant pathogens: a systematic review and meta-analysis. Clin Infect Dis. 2014; doi: 10.1093/cid/cit734.

23. José R, Periselneris J, Brown J. Community-acquired pneumonia. Curr Opin Pulm Med. 2015; doi: 10.1097/MCP.0000000000000150.

24. Montravers P, Harpan A, Guivarch E. Current and Future Considerations for the Treatment of Hospital-Acquired Pneumonia. Adv Ther. 2016; doi: 10.1007/s12325-016-0293-x.

25. Lopez A, Amaro R, Polverino E. Does health care associated pneumonia really exist?. Eur J Intern Med. 2012; doi: 10.1016/j.ejim.2012.05.006.

26. Ewig S, Klapdor B, Pletz M, Rohde G, Schütte H, Schaberg T, Bauer T, Welte T. Nursing-home-acquired pneumonia in Germany: an 8-year prospective multicentre study. Thorax. 2012; doi: 10.1136/thoraxjnl-2011-200630.

27. Lim WS, Macfarlane JT. A prospective comparison of nursing home acquired pneumonia with community acquired pneumonia. Eur Respir J. 2001.

28. Polverino E, Dambrava P, Cillóniz C, Balasso V, Marcos MA, Esquinas C, Mensa J, Ewig S, Torres A. Nursing home-acquired pneumonia: a 10 year single-centre experience. Thorax. 2010; doi: 10.1136/thx.2009.124776.

29. Eliakim Raz N, Robenshtok E, Shefet D, Gafter Gvili A, Vidal L, Paul M, Leibovici L. Empiric antibiotic coverage of atypical pathogens for community-acquired pneumonia in hospitalized adults. Cochrane Database Syst Rev. 2012; doi: 10.1002/14651858.CD004418.pub4.

30. Mills G, Oehley M, Arrol B. Effectiveness of beta lactam antibiotics compared with antibiotics active against atypical pathogens in non-severe community acquired pneumonia: meta-analysis. BMJ. 2005; doi: 10.1136/bmj.38334.591586.82.

31. Postma D, van Werkhoven C, van Elden, Leontine J R, Thijsen SFT, Hoepelman AIM, Kluytmans, Jan A J W, Boersma W, Compaijen C, van der Wall E, Prins J, Oosterheert J, Bonten MJM. Antibiotic treatment strategies for community-acquired pneumonia in adults. N Engl J Med. 2015; doi: 10.1056/NEJMoa1406330.

32. Kollef MH, Sherman G, Ward S, Fraser VJ. Inadequate antimicrobial treatment of infections: a risk factor for hospital mortality among critically ill patients. Chest. 1999.

33. Micek S, Lloyd A, Ritchie D, Reichley R, Fraser V, Kollef M. Pseudomonas aeruginosa bloodstream infection: importance of appropriate initial antimicrobial treatment. Antimicrob Agents Chemother. 2005; doi: 10.1128/AAC.49.4.1306-1311.2005.

34. Fraser A, Paul M, Almanasreh N, Tacconelli E, Frank U, Cauda R, Borok S, Cohen M, Andreassen S, Nielsen A, Leibovici L. Benefit of appropriate empirical antibiotic treatment: thirty-day mortality and duration of hospital stay. Am J Med. 2006; doi: 10.1016/j.amjmed.2006.03.034.

35. Ciccolini M, Spoorenberg V, Geerlings SE, Prins JM, Grundmann H. Using an index-based approach to assess the population-level appropriateness of empirical antibiotic therapy. J Antimicrob Chemother. 2015; doi: 10.1093/jac/dku336.

36. Seaton RA, Nathwani D, Burton P, McLaughlin C, MacKenzie AR, Dundas S, Ziglam H, Gourlay Y, Beard K, Douglas E. Point prevalence survey of antibiotic use in Scottish hospitals utilising the Glasgow Antimicrobial Audit Tool (GAAT). Int J Antimicrob Agents. 2007; doi: 10.1016/j.ijantimicag.2006.10.020.

37. Kumar A, Ellis P, Arabi Y, Roberts D, Light B, Parrillo J, Dodek P, Wood G, Simon D, Peters C, Ahsan M, Chateau D. Initiation of inappropriate antimicrobial therapy results in a fivefold reduction of survival in human septic shock. Chest. 2009; doi: 10.1378/chest.09-0087.

38. Raveh D, Levy Y, Schlesinger Y, Greenberg A, Rudensky B, Yinnon AM. Longitudinal surveillance of antibiotic use in the hospital. QJM. 2001.

39. James R, Upjohn L, Cotta M, Luu S, Marshall C, Buising K, Thursky K. Measuring antimicrobial prescribing quality in Australian hospitals: development and evaluation of a national antimicrobial prescribing survey tool. J Antimicrob Chemother. 2015; doi: 10.1093/jac/dkv047.

40. Vlahovic Palcevski V, Francetic I, Palcevski G, Novak S, Abram M, Bergman U. Antimicrobial use at a university hospital: appropriate or misused? A qualitative study. Int J Clin Pharmacol Ther. 2007.

41. Peron E, Hirsch A, Jury L, Jump RLP, Donskey C. Another setting for stewardship: high rate of unnecessary antimicrobial use in a veterans affairs long-term care facility. J Am Geriatr Soc. 2013; doi: 10.1111/jgs.12099.

42. van Buul L, Veenhuizen R, Achterberg W, Schellevis F, Essink, Rob T G M, de Greeff S, Natsch S, van der Steen, Jenny T, Hertogh CMPM. Antibiotic prescribing in Dutch nursing homes: how appropriate is it?. J Am Med Dir Assoc. 2015; doi: 10.1016/j.jamda.2014.10.003.

43. Schultz L, Lowe T, Srinivasan A, Neilson D, Pugliese G. Economic impact of redundant antimicrobial therapy in US hospitals. Infect Control Hosp Epidemiol. 2014; doi: 10.1086/678066.

44. Willemsen I, van der Kooij T, van Benthem B, Wille J, Kluytmans J. Appropriateness of antimicrobial therapy:

PART II | CHAPTER 8

INAPPROPRIATE USE OF ANTIMICROBIALS FOR LOWER RESPIRATORY TRACT INFECTIONS IN ELDERLY PATIENTS:
PATIENT- AND COMMUNITY-RELATED IMPLICATIONS AND POSSIBLE INTERVENTIONS

a multicentre prevalence survey in the Netherlands, 2008-2009. Euro Surveill. 2010.

45.   Hajjar E, Cafiero A, Hanlon J. Polypharmacy in elderly patients. Am J Geriatr Pharmacother. 2007; doi: 10.1016/j.amjopharm.2007.12.002.

46.   Warren JW, Palumbo FB, Fitterman L, Speedie SM. Incidence and characteristics of antibiotic use in aged nursing home patients. J Am Geriatr Soc. 1991.

47.   Rotjanapan P, Dosa D, Thomas K. Potentially inappropriate treatment of urinary tract infections in two Rhode Island nursing homes. Arch Intern Med. 2011; doi: 10.1001/archinternmed.2011.13.

48.   Zimmer JG, Bentley DW, Valenti WM, Watson NM. Systemic antibiotic use in nursing homes. A quality assessment. J Am Geriatr Soc. 1986.

49.   Pickering TD, Gurwitz JH, Zaleznik D, Noonan JP, Avorn J. The appropriateness of oral fluoroquinolone-prescribing in the long-term care setting. J Am Geriatr Soc. 1994.

50.   Loeb M, Simor AE, Landry L, Walter S, McArthur M, Duffy J, Kwan D, McGeer A. Antibiotic use in Ontario facilities that provide chronic care. J Gen Intern Med. 2001.

51.   Vergidis P, Hamer D, Meydani S, Dallal G, Barlam T. Patterns of antimicrobial use for respiratory tract infections in older residents of long-term care facilities. J Am Geriatr Soc. 2011; doi: 10.1111/j.1532-5415.2011.03406.x.

52.   Spivak E, Cosgrove S, Srinivasan A. Measuring Appropriate Antimicrobial Use: Attempts at Opening the Black Box. Clin Infect Dis. 2016; doi: 10.1093/cid/ciw658.

53.   Casaroto E, Marra A, Camargo TZS, de Souza, Ana Rita Araújo, de Almeida, Carlos Eduardo Saldanha, Pedroti E, Victor EdS, dos Santos, Oscar Fernando Pavão, Edmond M, Campos A. Agreement on the prescription of antimicrobial drugs. BMC Infect Dis. 2015; doi: 10.1186/s12879-015-0992-y.

54.   Baclet N, Ficheur G, Alfandari S, Ferret L, Senneville E, Chazard E, Beuscart J. Explicit definitions of potentially inappropriate prescriptions of antibiotics in older patients: a compilation derived from a systematic review. Int J Antimicrob Agents. 2017; doi: 10.1016/j.ijantimicag.2017.08.011.

55.   Sogn DD, Evans R, Shepherd GM, Casale TB, Condemi J, Greenberger PA, Kohler PF, Saxon A, Summers RJ, VanArsdel PP. Results of the National Institute of Allergy and Infectious Diseases Collaborative Clinical Trial to test the predictive value of skin testing with major and minor penicillin derivatives in hospitalized adults. Arch Intern Med. 1992.

56.   Barnes S, Rock C, Harris A, Cosgrove S, Morgan D, Thom K. The Impact of Reducing Antibiotics on the Transmission of Multidrug-Resistant Organisms. Infect Control Hosp Epidemiol. 2017; doi: 10.1017/ice.2017.34.

57.   Barlam T, Cosgrove S, Abbo L, MacDougall C, Schuetz A, Septimus E, Srinivasan A, Dellit T, Falck Ytter Y, Fishman N, Hamilton C, Jenkins T, Lipsett P, Malani P, May L, Moran G, Neuhauser M, Newland J, Ohl C, Samore M, Seo S, Trivedi K. Implementing an Antibiotic Stewardship Program: Guidelines by the Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America. Clin Infect Dis. 2016; doi: 10.1093/cid/ciw118.

58.   Faulkner C, Cox H, Williamson J. Unique aspects of antimicrobial use in older adults. Clin Infect Dis. 2005; doi: 10.1086/428125.

59.   Tamma P, Avdic E, Li D, Dzintars K, Cosgrove S. Association of Adverse Events With Antibiotic Use in Hospitalized Patients. JAMA Intern Med. 2017; doi: 10.1001/jamainternmed.2017.1938.

60.   Gurwitz JH, Field TS, Avorn J, McCormick D, Jain S, Eckler M, Benser M, Edmondson AC, Bates DW. Incidence and preventability of adverse drug events in nursing homes. Am J Med. 2000.

61.   Hohl CM, Dankoff J, Colacone A, Afilalo M. Polypharmacy, adverse drug-related events, and potential adverse drug interactions in elderly patients presenting to an emergency department. Ann Emerg Med. 2001; doi: 10.1067/mem.2001.119456.

62.   van der Linden, Paul D, Sturkenboom, Miriam C J M, Herings RMC, Leufkens HMG, Rowlands S, Stricker BHC. Increased risk of achilles tendon rupture with quinolone antibacterial use, especially in elderly patients taking oral corticosteroids. Arch Intern Med. 2003; doi: 10.1001/archinte.163.15.1801.

63.   Wise B, Peloquin C, Choi H, Lane N, Zhang Y. Impact of age, sex, obesity, and steroid use on quinolone-associated tendon disorders. Am J Med. 2012; doi: 10.1016/j.amjmed.2012.05.027.

64.   Fialová D, Topinková E, Gambassi G, Finne Soveri H, Jónsson P, Carpenter I, Schroll M, Onder G, Sørbye L, Wagner C, Reissigová J, Bernabei R. Potentially inappropriate medication use among elderly home care patients in Europe. JAMA. 2005; doi: 10.1001/jama.293.11.1348.

65.   Ray W, Murray K, Hall K, Arbogast P, Stein CM. Azithromycin and the risk of cardiovascular death. N Engl J Med. 2012; doi: 10.1056/NEJMoa1003833.

66.   Ray W, Murray K, Meredith S, Narasimhulu S, Hall K, Stein CM. Oral erythromycin and the risk of sudden death from cardiac causes. N Engl J Med. 2004; doi: 10.1056/NEJMoa040582.

67.   Mortensen E, Halm E, Pugh M, Copeland L, Metersky M, Fine M, Johnson C, Alvarez C, Frei C, Good C, Restrepo M, Downs J, Anzueto A. Association of azithromycin with mortality and cardiovascular events among older patients hospitalized with pneumonia. JAMA. 2014; doi: 10.1001/jama.2014.4304.

68.   Schembri S, Williamson P, Short P, Akram A, Taylor J, Singanayagam A, Hill A, Chalmers J. Cardiovascular events after clarithromycin use in lower respiratory tract infections: analysis of two prospective cohort studies. BMJ. 2013; doi: 10.1136/bmj.f1235.

69.   Chou H, Wang J, Chang C, Lai C, Lai M, Chan KA. Risks of cardiac arrhythmia and mortality among patients using new-generation macrolides, fluoroquinolones, and ß-lactam/ß-lactamase inhibitors: a Taiwanese nationwide study. Clin Infect Dis. 2015; doi: 10.1093/cid/ciu914.

70.   Svanström H, Pasternak B, Hviid A. Use of azithromycin and death from cardiovascular causes. N Engl J Med. 2013; doi: 10.1056/NEJMoa1300799.

71.   Petrosillo N, Cataldo M, Pea F. Treatment options for community-acquired pneumonia in the elderly people. Expert Rev Anti Infect Ther. 2015; doi: 10.1586/14787210.2015.1021783.

72.   Asempa T, Nicolau D. Clostridium difficile infection in the elderly: an update on management. Clin Interv Aging. 2017; doi: 10.2147/CIA.S149089.

73.   UpToDate. Antimicrobial agents that may induce Clostridium difficile diarrhea and colitis. In: . 2017. Accessed 12/23 2017.

74.   Davey P, Marwick C, Scott C, Charani E, McNeil K, Brown E, Gould I, Ramsay C, Michie S. Interventions to improve antibiotic prescribing practices for hospital inpatients. Cochrane Database Syst Rev. 2017; doi: 10.1002/14651858.CD003543.pub4.

75.   Dellit T, Owens R, McGowan J, Gerding D, Weinstein R, Burke J, Huskins WC, Paterson D, Fishman N,

PART II | CHAPTER 8

INAPPROPRIATE USE OF ANTIMICROBIALS FOR LOWER RESPIRATORY TRACT INFECTIONS IN ELDERLY PATIENTS:
PATIENT- AND COMMUNITY-RELATED IMPLICATIONS AND POSSIBLE INTERVENTIONS

Carpenter C, Brennan PJ, Billeter M, Hooton T. Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America guidelines for developing an institutional program to enhance antimicrobial stewardship. Clin Infect Dis. 2007; doi: 10.1086/510393.

76. Adam H, Baxter M, Davidson R, Rubinstein E, Fanella S, Karlowsky J, Lagacé-Wiens PRS, Hoban D, Zhanel G. Comparison of pathogens and their antimicrobial resistance patterns in paediatric, adult and elderly patients in Canadian hospitals. J Antimicrob Chemother. 2013; doi: 10.1093/jac/dkt024.

77. Mandell LA, Wunderink RG, Anzueto A, Bartlett JG, Campbell GD, Dean NC, Dowell SF, File TM,Jr, Musher DM, Niederman MS, Torres A, Whitney CG, Infectious Diseases Society of America, American Thoracic Society. Infectious Diseases Society of America/American Thoracic Society consensus guidelines on the management of community-acquired pneumonia in adults. Clin Infect Dis. 2007; doi: CID41620 [pii].

78. Prina E, Ranzani O, Torres A. Community-acquired pneumonia. Lancet. 2015; doi: 10.1016/S0140-6736(15)60733-4.

79. Amalakuhan B, Echevarria KL, Restrepo MI. Managing community acquired pneumonia in the elderly - the next generation of pharmacotherapy on the horizon. Expert Opin Pharmacother. 2017; doi: 10.1080/14656566.2017.1340937.

80. Dyar OJ, Huttner B, Schouten J, Pulcini C. What is antimicrobial stewardship?. Clin Microbiol Infect. 2017; doi: 10.1016/j.cmi.2017.08.026.

81. Dean NC, Silver MP, Bateman KA, James B, Hadlock CJ, Hale D. Decreased mortality after implementation of a treatment guideline for community-acquired pneumonia. Am J Med. 2001.

82. Dean N, Bateman K, Donnelly S, Silver M, Snow G, Hale D. Improved clinical outcomes with utilization of a community-acquired pneumonia guideline. Chest. 2006; doi: 10.1378/chest.130.3.794.

83. Fowler S, Webber A, Cooper BS, Phimister A, Price K, Carter Y, Kibbler CC, Simpson AJH, Stone SP. Successful use of feedback to improve antibiotic prescribing and reduce Clostridium difficile infection: a controlled interrupted time series. J Antimicrob Chemother. 2007; doi: 10.1093/jac/dkm014.

84. McNulty C, Logan M, Donald IP, Ennis D, Taylor D, Baldwin RN, Bannerjee M, Cartwright KA. Successful control of Clostridium difficile infection in an elderly care unit through use of a restrictive antibiotic policy. J Antimicrob Chemother. 1997.

85. Katz PR, Beam TR, Brand F, Boyce K. Antibiotic use in the nursing home. Physician practice patterns. Arch Intern Med. 1990.

86. Katz M, Gurses A, Tamma P, Cosgrove S, Miller M, Jump RLP. Implementing Antimicrobial Stewardship in Long-term Care Settings: An Integrative Review Using a Human Factors Approach. Clin Infect Dis. 2017; doi: 10.1093/cid/cix566.

87. Crnich C, Jump R, Trautner B, Sloane P, Mody L. Optimizing Antibiotic Stewardship in Nursing Homes: A Narrative Review and Recommendations for Improvement. Drugs Aging. 2015; doi: 10.1007/s40266-015-0292-7.

88. Naughton BJ, Mylotte JM, Ramadan F, Karuza J, Priore RL. Antibiotic use, hospital admissions, and mortality before and after implementing guidelines for nursing home-acquired pneumonia. J Am Geriatr Soc. 2001.

89. Wenisch J, Equiluz Bruck S, Fudel M, Reiter I, Schmid A, Singer E, Chott A. Decreasing Clostridium difficile infections by an antimicrobial stewardship program that reduces moxifloxacin use. Antimicrob Agents Chemother. 2014; doi: 10.1128/AAC.03006-14.

90. Syrjälä H, Broas M, Suramo I, Ojala A, Lähde S. High-resolution computed tomography for the diagnosis of community-acquired pneumonia. Clin Infect Dis. 1998.

91. Hagaman J, Rouan G, Shipley R, Panos R. Admission chest radiograph lacks sensitivity in the diagnosis of community-acquired pneumonia. Am J Med Sci. 2009; doi: 10.1097/MAJ.0b013e31818ad805.

92. Wunderink R, Waterer G. Advances in the causes and management of community acquired pneumonia in adults. BMJ. 2017; doi: 10.1136/bmj.j2471.

93. Claessens Y, Debray M, Tubach F, Brun A, Rammaert B, Hausfater P, Naccache J, Ray P, Choquet C, Carette M, Mayaud C, Leport C, Duval X. Early Chest Computed Tomography Scan to Assist Diagnosis and Guide Treatment Decision for Suspected Community-acquired Pneumonia. Am J Respir Crit Care Med. 2015; doi: 10.1164/rccm.201501-0017OC.

94. Llamas Álvarez A, Tenza Lozano E, Latour Pérez J. Accuracy of Lung Ultrasonography in the Diagnosis of Pneumonia in Adults: Systematic Review and Meta-Analysis. Chest. 2017; doi: 10.1016/j.chest.2016.10.039.

95. Oosterheert J, van Loon A, Schuurman R, Hoepelman AIM, Hak E, Thijsen S, Nossent G, Schneider MME, Hustinx WMN, Bonten MJM. Impact of rapid detection of viral and atypical bacterial pathogens by real-time polymerase chain reaction for patients with lower respiratory tract infection. Clin Infect Dis. 2005; doi: 10.1086/497134.

96. Hernes SS, Hagen E, Quarsten H, Bjorvatn B, Bakke PS. No impact of early real-time PCR screening for respiratory viruses on length of stay and use of antibiotics in elderly patients hospitalized with symptoms of a respiratory tract infection in a single center in Norway. Eur J Clin Microbiol Infect Dis. 2014; doi: 10.1007/s10096-013-1963-0.

97. Tobia C, Aspinall S, Good C, Fine M, Hanlon J. Appropriateness of antibiotic prescribing in veterans with community-acquired pneumonia, sinusitis, or acute exacerbations of chronic bronchitis: a cross-sectional study. Clin Ther. 2008; doi: 10.1016/j.clinthera.2008.06.009.

9

NARROW-SPECTRUM ANTIBIOTICS FOR COMMUNITY-ACQUIRED PNEUMONIA IN ADULTS: A STEPPED-WEDGE CLUSTER RANDOMISED ANTIMICROBIAL STEWARDSHIP TRIAL

Valentijn A. Schweitzer*, Inger van Heijl*, Cornelis H. van Werkhoven, Wim G. Boersma, Wouter Rozemeijer, Kees Verduin, Sanjay U.C. Sankatsing, Akke K. van der Bij, Heidi S.M. Ammerlaan, Ilse Overdevest, Milena Roorda-van der Vegt, Elske Engel-Dettmers, Dirk R.A.J. de Munck, Florence Ayuketah-Ekokobe, Michiel B. Haeseker, Wendelien Dorigo-Zetsma, Paul D. van der Linden, C.H. Edwin Boel, Jan J. Oosterheert, Marc J.M. Bonten
*Contributed equally

PART II | CHAPTER 9

NARROW-SPECTRUM ANTIBIOTICS FOR COMMUNITY-ACQUIRED PNEUMONIA IN ADULTS: A STEPPED-WEDGE CLUSTER RANDOMISED ANTIMICROBIAL STEWARDSHIP TRIAL

ABSTRACT

BACKGROUND

Adults hospitalized with moderately-severe community-acquired pneumonia (CAP) are often treated with broad-spectrum antibiotics, despite guidelines recommending narrow-spectrum antibiotics. We determined safety and effectiveness of a multifaceted antibiotic stewardship intervention to increase the use of narrow-spectrum antibiotics.

METHODS

In a stepped-wedge cluster-randomized trial in twelve hospitals we implemented an intervention consisting of education, motivating opinion leaders, and audit and feedback and determined its effect on the days of therapy (DOT) with narrow-spectrum antibiotics (penicillin, amoxicillin or doxycycline) and all-cause 90-day mortality. In an intention-to-treat analysis we used 90-day all-cause mortality as co-primary safety outcome (non-inferiority margin 3% and one-sided alpha of 0.05) and reduction in DOT with narrow-spectrum antibiotics as co-primary effectiveness outcome.

RESULTS

From November 2015 till November 2017 4,084 patients were included; 2,235 in the control and 1,849 in the intervention period. Median age was 73 (range 18 – 101) years, 53.0% (2163/4084) were male and mean Pneumonia Severity Index score was 91.3 (SD ±31.4). The adjusted relative reduction in broad-spectrum DOT during intervention was 26.9% (95% CI: 15.4%-37.4%) with averages of 6.6 and 4.8 days in the control and intervention period, respectively. Crude 90-day mortality was 10.9% (242/2228) and 10.8% (199/1841) during control and intervention period, yielding an adjusted absolute difference of 0.4% (90% CI: -2.7 to 2.4) for the control versus intervention period, indicating non-inferiority for all-cause mortality.

CONCLUSIONS

In patients hospitalized with moderately-severe CAP a multifaceted antibiotic stewardship intervention safely reduced the days of broad-spectrum antibiotic use with 27 %. (CAP-PACT ClinicalTrials.gov number, NCT02604628.)

INTRODUCTION

Community-acquired pneumonia (CAP) is associated with high morbidity and mortality[1,2]. International guidelines recommend to base empiric antibiotic treatment of patients with CAP on the severity of disease. Accordingly, CAP may be classified as moderately-severe based on the need of hospital admission in a non-intensive care unit (ICU) ward (pragmatic classification); on a pneumonia severity index (PSI) score of 3-4; or on a CURB-65 score of 2[3]. Recommended empirical antibiotic treatments of moderate-severe CAP differ and include narrow-spectrum beta-lactam monotherapy (such as penicillin G or amoxicillin)[4,5], beta-lactam macrolide combination therapy[6,7] or fluoroquinolone monotherapy[7]. Empiric use of narrow-spectrum beta-lactams exerts less selective pressure for antibiotic resistance, compared to more broad-spectrum antibiotics. However, high quality evidence for equal clinical effectiveness of narrow-spectrum and more broad-spectrum beta-lactam monotherapy is limited to two randomized trials comparing amoxicillin monotherapy to fluoroquinolones. One study used a fluoroquinolone no longer in use (sparfloxacin) and the other study only included patients with mild pneumonia[10,11]. Adherence to guideline recommendations for using narrow-spectrum beta-lactams as empiric treatment in patients with moderate-severe CAP is low[12]. Antimicrobial stewardship may improve guideline adherence and appropriateness of empiric antibiotic therapy[13]. Furthermore, high-quality evidence on safety and effectiveness of narrow-spectrum beta-lactam monotherapy in patients with moderate-severe CAP could augment confidence in adaptation of such treatment strategies.

We, therefore, investigated whether implementation of a multifaceted antimicrobial stewardship intervention reduced broad-spectrum antibiotic use without compromising patients safety.

METHODS

The "Community-Acquired Pneumonia increasing Protocol adherence by Antibiotic stewardship in a stepped-wedge Cluster-randomized Trial" (CAP-PACT) trial was an investigator-initiated stepped wedge cluster randomized quality improvement trial. The study was reviewed by the ethics review board of the University Medical Centre Utrecht (reference number 15/100), and local antimicrobial stewardship teams consented to participate in the study. As the implemented stewardship intervention promoted best-practice antibiotic therapy as described in the guidelines, as the intervention was aimed at healthcare providers rather than individual patients, and as anonymised routinely

PART II | CHAPTER 9

NARROW-SPECTRUM ANTIBIOTICS FOR COMMUNITY-ACQUIRED PNEUMONIA IN ADULTS: A STEPPED-WEDGE CLUSTER
RANDOMISED ANTIMICROBIAL STEWARDSHIP TRIAL

collected data were used, the need to obtain individual informed consent was waived. Data are reported according to the CONSORT guidelines of stepped wedge cluster randomised trials and non-inferiority trials[14]. The trial was prospectively registered at clinicaltrials.gov (identifier NCT02604628).

PATIENTS AND STUDY SETTING

The trial was performed from November 2015 till November 2017 in 12 hospitals in the Netherlands. All Dutch hospitals were eligible for study participation. Participating hospitals were university hospitals (n=2), teaching hospitals (n=7), and non-teaching hospitals (n=3). Consecutive adult patients of 18 years or older receiving antibiotic therapy for a working diagnosis of CAP and admitted to a non-ICU ward were enrolled. Patients were not eligible if they: recently (≤14 days) resided in a nursing home or long-term care facility; were recently (≤14 days) admitted to an acute care hospital for two or more days; were known to have cystic fibrosis; or were immunocompromised. Immunocompromised was defined as: having a human immunodeficiency virus infection with a last CD4 cell count of <300 cells/μL; having received cytotoxic chemotherapy or radiotherapy in the previous 3 months; being on chronic (>3 months) haemodialysis; having received a solid organ or bone marrow transplant; or receiving immunosuppressive therapy. In the Netherlands every hospital has a specialised antimicrobial stewardship team responsible for implementing antimicrobial stewardship interventions, consisting of at least a clinical microbiologist, an infectious disease specialist and a hospital pharmacist. Three out of twelve hospitals stopped participation before they entered the intervention period, either because the principal investigator stopped research activities (n=2) or because they were no longer able to collect study data (n=1). Data from these hospitals was not used for analysis.

ANTIMICROBIAL STEWARDSHIP INTERVENTION

The intervention was a multifaceted bundle based on proven effective stewardship interventions[15]. The bundle consisted of (1) education, (2) activating local opinion leaders, and (3) prospective audit and feedback of antibiotic use. Educational activities were targeted at physicians of pulmonary and internal medicine departments and consisted of clinical lessons, e-learning and educational attributes. Clinical lessons were given in regular intervals of 6 months in which the CAP guidelines were addressed using case-based discussions with feedback of antibiotic prescribing data of the respective hospital, which were anonymously benchmarked against other participating hospitals. At the start of the intervention period physicians of the participating hospitals were invited to complete the e-learning, containing case-based questions about the CAP guideline. Invites to complete the e-learning were send periodically to reach new employees. In addition, educational attributes in the form of posters and pocket cards

summarizing the CAP guidelines were distributed at the start of the intervention. At intervention introduction, local opinion leaders were identified in collaboration with the local antibiotic stewardship team, and actively involved in the study and stewardship activities. They were asked to stimulate guideline adherent treatment throughout the intervention period, e.g. during hand-over meetings. Thirdly, prospective audit and feedback was implemented during the intervention period by the local antimicrobial stewardship team. On week days all patients admitted with moderate-severe CAP were actively identified and responsible physicians were contacted by a member of the local antimicrobial stewardship team to switch treatment to penicillin or amoxicillin monotherapy if treatment was not according to the guideline recommendation. If, for any reason, treatment could not be switched it was recommended to perform a pneumococcal urine antigen test to facilitate de-escalation if the test result was positive. Recommendations were done by telephone and were registered in electronic health records, as were reasons for not accepting recommendations.

OUTCOMES

The study had two co-primary outcomes: broad-spectrum days of therapy (DOT) per patient and all-cause 90-day mortality. Days that patients received antibiotic treatment were classified as narrow-spectrum DOT if amoxicillin, penicillin or doxycycline monotherapy was given and as broad-spectrum DOT if any other antibiotic regimen was administered. Doxycycline monotherapy was defined as narrow-spectrum therapy as it is recommended in our national guideline as equivalent to amoxicillin for mild CAP. Vital status at day 90 was derived from the municipal records database if not evident from the medical records. Secondary outcomes were narrow-spectrum DOT, total DOT, 30-day mortality, length of hospital stay, hospital readmissions within 30-days of hospital admission, intensive care unit admissions, complications, *Clostridioides difficile* associated disease and antibiotic switches.

RANDOMISATION

The unit of randomization was the hospital and comprised the sequence in which they would crossover from control to intervention period (Figure S1). Randomisation was performed electronically by a data manager blinded to the intervention after recruitment of all hospitals. All the participating hospitals started with a control period and every 3 months a block of two hospitals transitioned from control to intervention, with all hospitals eventually ending in the intervention period. Allocation to the time of intervention implementation was concealed for treating physicians.

STATISTICAL ANALYSIS

The detailed statistical analysis plan was published at clinicaltrials.gov (NCT02604628)

PART II | CHAPTER 9

NARROW-SPECTRUM ANTIBIOTICS FOR COMMUNITY-ACQUIRED PNEUMONIA IN ADULTS: A STEPPED-WEDGE CLUSTER
RANDOMISED ANTIMICROBIAL STEWARDSHIP TRIAL

before database lock. The sample size calculation was determined by non-inferiority for 90-day mortality. Assuming a 90-day mortality of 10%, a non-inferiority margin of 3%, and a one-sided alpha 0.05, a total of 4464 patients were required for 80% power. Primary intention-to-treat (ITT) analyses were performed using mixed effect models including random intercept and fixed slope per hospital, time as fixed effect, adjusted for the following potential confounders as fixed effects: PSI-score, smoking status, COPD, diabetes mellitus, and antibiotic pre-treatment. For 90-day mortality we used a mixed effects logistic regression, recalculated to risk differences[16]. As secondary analyses we performed as-treated analysis and a Complier Average Causal Effect (CACE) analysis, both aiming to estimate the difference in mortality between patients empirically treated with narrow-spectrum versus broad-spectrum antibiotics. The CACE was used with randomisation as an instrumental variable to estimate the ITT adjusted for non-compliance[17]. DOTs were analysed using mixed effects negative binomial models with robust standard errors. Length of hospital stay was analysed using mixed effects Cox proportional hazards models using in-hospital mortality as a competing event, assuming maximum follow-up duration when censored. ICU admissions, hospital readmissions and antibiotic switches were analysed using mixed effects logistic regression. Missing data were imputed by multiple imputation, with the exception of data on respiratory rate, heart rate, and confusion at admission which were assumed to be normal when not documented in the medical charts. No transition period was used because the treatment effect of the audit and feedback was assumed to be immediate. All analyses were performed using R statistical software version 3.5.1.

RESULTS

A total of 4084 patients with CAP admitted to a non-ICU ward were included during the study period; 2,235 during the control period and 1,849 during the intervention period (Figure 1). The baseline characteristics of patients in the control period were similar to patients in the intervention period (Table 1). The most commonly identified pathogen was *Streptococcus pneumoniae* 488 (11.9%), followed by *Haemophilus influenza* 285 (7.0%), *and Staphylococcus aureus* 97 (2.4%), with no differences between the control and intervention period (Table S1).
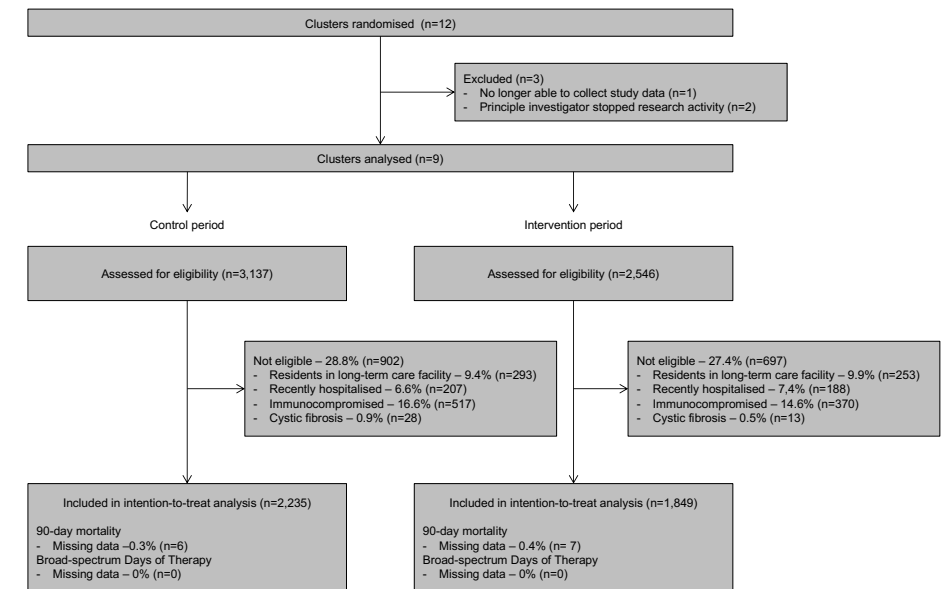
Figure 1 Flowchart of patient inclusion during the study period.

ANTIMICROBIAL STEWARDSHIP INTERVENTION IMPLEMENTATION

In the intervention periods, a total of 54 clinical lessons were given, with an average interval of 90.0 days (3 months) and 235 e-learnings were completed (Table S2). The local stewardship teams offered 330 recommendations for changing antibiotics, of which 197 (59.7%) were accepted. Most common reasons for not accepting recommendations were patients having COPD (11, 8.3%), and patients with severe pneumonia according to PSI or CURB score or admitted with clinical deterioration (9; 6.8%).

PRIMARY OUTCOMES

The median broad-spectrum DOT/patient was 6 (interquartile range (IQR) 2-9) in the control and 3 (IQR 0-8) in the intervention period, yielding an adjusted relative reduction in broad-spectrum DOTs of 26.9% (95% confidence interval (CI) 15.4%-37.4%) (Table S3). Adjusted relative reductions in broad-spectrum DOT per hospital ranged from 16.7% to 39.3% (Table S4). The median narrow-spectrum DOT/patient were 0 (IQR 0-6) and 5 (IQR 0-8), while the median total DOT/patient were 8 (IQR 7-10) in the control and 8 (IQR 7-11) in the intervention period (Figure 2). Most common prescribed empiric antibiotic regimens in the control and intervention period respectively were narrow-spectrum antibiotics (28.5% vs. 45.1%), broad-spectrum beta-lactam monotherapy (33.8% vs. 26.9%), and beta-lactam fluoroquinolone combination therapy (27.3% vs. 23.2%) (Table

PART II | CHAPTER 9

NARROW-SPECTRUM ANTIBIOTICS FOR COMMUNITY-ACQUIRED PNEUMONIA IN ADULTS: A STEPPED-WEDGE CLUSTER
RANDOMISED ANTIMICROBIAL STEWARDSHIP TRIAL

S5). Overall narrow-spectrum antibiotic use mostly consisted of amoxicillin (1303/1409, 92.5%), followed by penicillin (48/1409, 3.4%), and doxycycline (139/1409, 9.9%).

Table 1. Baseline characteristics of patients in the control period compared to patients in the intervention period

	Control (n=2,235)	Intervention (n=1,849)
Age (years, median, IQR)	73 (63-81)	74 (64-82)
Male gender (n, %)	1,188 (53.2)	975 (52.7)
Antibiotic use before admission (n, %)	742 (33.2)	569 (30.8)
Smoke behaviour (n,%)		
Active smoker	51 2(27.2)	412 (27.6)
Past smoker	725 (38.5)	626 (41.9)
Never smoked	396 (21.1)	264 (17.7)
Medical speciality (n,%)		
Internal medicine	416 (18.6)	349 (18.9)
Pulmonology	1731 (77.4)	1426 (77.1)
Co-morbid conditions (n, %)		
COPD or asthma	962 (43.0)	880 (47.6)
Cardiovascular disease	300 (13.4)	259 (14.0)
Diabetes mellitus	38 9(17.4)	31 5(17.0)
Malignancy	239 (10.7)	185 (10.0)
PSI score (median, IQR)	89 (70-112)	91 (72-113)
PSI risk class I (n, %)	101 (4.5)	73 (3.9)
PSI risk class II (n, %)	473 (21.2)	357 (19.3)
PSI risk class III (n, %)	581 (26.0)	493 (26.7)
PSI risk class IV (n, %)	823 (36.8)	722 (39.0)
PSI risk class V (n, %)	257 (11.5)	204 (11.0)
CURB-65 score (median, IQR)	2 (1-2)	2 (1-2)
Had radiologically confirmed CAP (n, %)	1689 (75.6)	1377 (74.5)
Blood culture obtained (n, %)	1602 (71.7)	1387 (75.0)
Sputum culture obtained (n, %)	888 (39.7)	784 (42.4)
PUAT performed (n, %)	965 (43.2)	1173 (63.4)
LUAT performed (n, %)	1297 (58.0)	1255 (67.9)

COPD chronic obstructive pulmonary disease, PUAT pneumococcal urinary antigen test , LUAT legionella urinary antigen test, PSI pneumonia severity index, IQR interquartile range, CURB-65 pneumonia severity score based on presence of confusion, blood ureum levels, respiratory rate, blood pressure, and age, CAP community-acquired pneumonia

All-cause 90-day mortality was 10.9% (242/2228) in the control period and 10.8% (199/1841) in the intervention period (Figure S2). In the ITT analysis, the adjusted risk difference in 90-day mortality was 0.4% (90% CI -2.7% to 2.4%). As the confidence interval does not contain the pre-specified non-inferiority margin of 3% we conclude that the intervention is non-inferior to the control period (Figure 3). Sensitivity analyses for patients with radiologically confirmed CAP yielded similar results. Results were comparable in the as-treated and CACE analysis (Table S6).
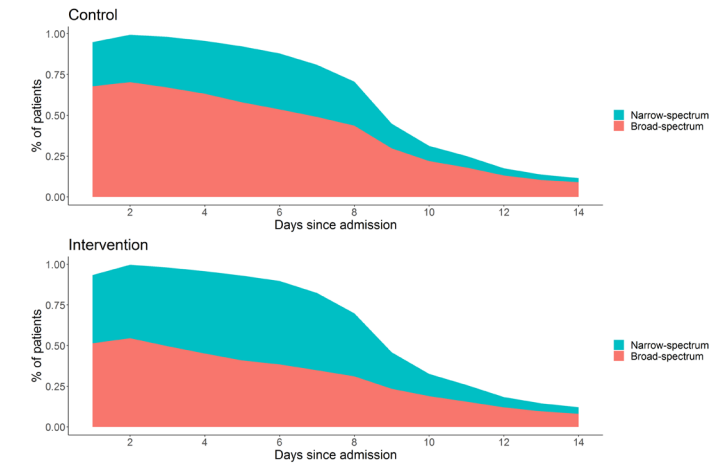
Figure 2 Proportion of patient receiving narrow-spectrum antibiotics (green) or broad-spectrum antibiotics (red) in days since hospital admission.
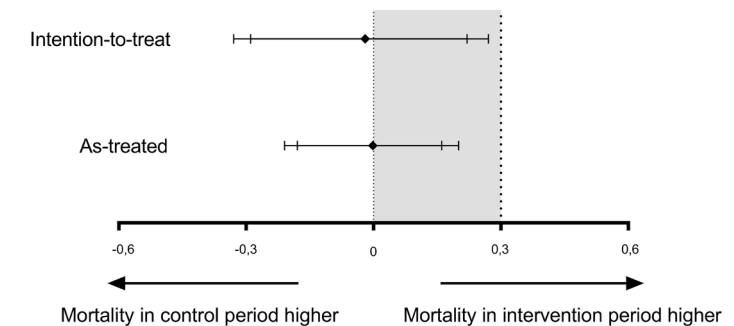
Figure 3 Non-inferiority plots

SECONDARY OUTCOMES

All-cause 30-day mortality was 6.9% (n=154) in the control and 6.7% (n=123) in the intervention period (Table 2). Results were comparable in the as-treated and CACE analysis (Table S7). Median length of hospital stay was 5 (IQR 3-8) days in the control and 5 (IQR 3-8) days in the intervention period. Comparable rates were observed for ICU-admissions (94, 4.2% control vs 38, 2.1% intervention), hospital readmissions (243, 11.3% control vs 203, 11.4% intervention), complications (380, 17.0% control vs 332, 18.0% intervention) (Table S8), *Clostridioides difficile* associated disease (2, 0.1% control vs 6, 0.3% intervention), and i.v. to oral switches (1469, 80.8% control vs 1262. 83.8% intervention). Switches from broad to narrow-spectrum occurred in 26.4% (n=421) in the control and 41.3% (n=413) in the intervention period.

PART II | CHAPTER 9

NARROW-SPECTRUM ANTIBIOTICS FOR COMMUNITY-ACQUIRED PNEUMONIA IN ADULTS: A STEPPED-WEDGE CLUSTER
RANDOMISED ANTIMICROBIAL STEWARDSHIP TRIAL

Table 2 Secondary outcomes

	Control (n=2,240)	Intervention (n=1,844)	Crude estimates (95% CI)	Adjusted estimates (95% CI)
30-day mortality (n, %)	154 (6.9%)	123 (6.7%)	RD -0.3 (-1.6-1.1)	RD -1.1 (-3.1 – 0.7)
Length of hospital stay (days, median, IQR)	5 (3-8)	5 (3-8)	HR 1.0 (1.0 – 1.1)	HR 1.1 (1.0 – 1.3)
Intensive care unit admissions (n, %)	94 (4.2%)	38 (2.1%)	OR 0.5 (0.3 – 0.7)	OR 0.3 (0.2 – 0.6)
Hospital readmissions (n, %)	243 (11.3%)	203 (11.4%)	OR 1.0 (0.8 – 1.2)	OR 1.2 (0.8 – 1.7)
Antibiotic switches				
Switch from BS to NS (n, %)	421 (26.4%)	413 (41.3%	OR 2.0 (1.7 – 2.3)	OR 2.1 (1.5 – 28)
Switch from NS to BS (n, %)	148 (23.0%)	195 (23.0%)	OR 1.0 (0.8 – 1.3)	OR 1.0 (0.7 – 1.5)
Switch from i.v. to oral (n, %)	1469 (80.8%)	1262 (83.8%)	OR 1.2 (1.0 – 1.5)	OR 1.3 (0.9 – 1.8)
Switch from oral to i.v. (n, %)	44 (10.6%)	26 (7.6%)	OR 0.7 (0.4 – 1.2)	OR 0.7 (0.3 – 1.8)

BS broad-spectrum antibiotics, NS narrow-spectrum antibiotics, i.v. intravenous, RD risk difference, HR hazard ratio, OR odds ratio

DISCUSSION

In this stepped wedge cluster randomised trial a multifaceted antimicrobial stewardship intervention based on education and audit and feedback resulted in a 27% reduction in broad-spectrum antibiotic use in patients hospitalized with moderate-severe CAP without compromising patient outcome at day 90 after hospital admission.

The stewardship intervention bundle contained elements that are considered effective in optimizing antibiotic use[13,15]. In meta-analysis the risk difference for guideline adherent prescription after implementation of antimicrobial stewardship interventions was 15% (95% CI 14%-16%)[13]. Compared to this, our bundle was more effective. However, heterogeneity was large between hospitals, with reductions ranging from 21% to 37%. Heterogeneity might have resulted from different barriers (reasons) for not prescribing narrow-spectrum antibiotics between hospitals. Therefore, even though the components of our bundle are evidence based, it might be a more optimal use of resources to first identify barriers for non-adherent prescribing and tailor the intervention to the specific setting based on behaviour change theory[18].

The optimal empiric treatment for moderate-severe CAP is still subject to debate. Thus far, two high quality RCTs have investigated the effects of empirical coverage for atypical pathogens on patient outcome[8,19]. In a multicentre non-inferiority RCT beta-lactam monotherapy was not non-inferior to beta-lactam macrolide combination

therapy for the time till reaching clinical stability. Patients randomized to beta-lactam monotherapy had a 7.6% higher absolute risk for not being clinically stable at day 7. The one-sided 90% CI was 13% which crossed the pre-defined non-inferiority margin of 8%. In a multicentre cluster randomised cross-over trial a strategy of preferred empiric treatment with beta-lactam monotherapy was compared to strategies with beta-lactam macrolide combination therapy and fluoroquinolone monotherapy. In this study, the beta-lactam monotherapy strategy was non-inferior to the broader regimens regarding 90-day mortality (2 sided 90% CI not overlapping the non-inferiority margin of 3%).

The equipoise is also reflected in international guideline recommendations for empiric treatment of moderate-severe CAP. American and British guidelines recommend beta-lactam macrolide combination therapy or fluoroquinolone monotherapy[6,7] while Swedish, Danish and Dutch guidelines recommend narrow-spectrum beta-lactam monotherapy[3-5]. The rationale for empiric treatment with narrow-spectrum beta-lactam monotherapy is, first, that the most common causative pathogen in CAP, *Streptococcus pneumoniae*, is susceptible to these antibiotics, and second that the severity of disease allows escalation within 48 hours to broader antibiotic therapy based on diagnostic testing or lack of clinical improvement. Thus far, two RCTs have evaluated the efficacy of narrow-spectrum beta-lactams in moderate-severe CAP patients[10,11]. Petitpretz et al. compared moxifloxacin to amoxicillin in patients with mild-to-moderate suspected pneumococcal CAP and clinical success rates were 86.5% (173/200) and 82.2% (171/208) in the moxifloxacin and amoxicillin treated patients, respectively. In the other RCT clinical cure rates were 83.6% (133/159) and 84.7% (144/170) in adult patients with community-acquired suspected pneumococcal pneumonia treated with sparfloxacin and amoxicillin, respectively. Yet, in clinical practice physicians apparently are reluctant to use narrow-spectrum beta-lactams, as demonstrated by the low adherence to current guideline recommendations in the control period of this study. Our findings provide further evidence that more patients with moderate-severe CAP can be safely empirically treated with narrow-spectrum antibiotics, which contributes to more prudent use of antibiotics.

Our study has several limitations. First, because we implemented a multifaceted bundle of multiple stewardship interventions it is not possible to estimate the effect of the individual components. However, the approach reflects clinical practice, where stewardship interventions are usually implemented as bundles. Second, the nature of the intervention precluded a blinded evaluation, and, therefore, information bias cannot be excluded. To minimize the impact of information bias we chose objective primary outcomes and used trained research nurses and standardized methodology for data collection. Third, in cluster randomized stepped wedge designs selection

PART II | CHAPTER 9

NARROW-SPECTRUM ANTIBIOTICS FOR COMMUNITY-ACQUIRED PNEUMONIA IN ADULTS: A STEPPED-WEDGE CLUSTER
RANDOMISED ANTIMICROBIAL STEWARDSHIP TRIAL

bias may occur if different patients are included in both study periods. Yet, baseline characteristics of patients in the control versus intervention period were similar and we adjusted for important prognostic factors. The cluster randomized design has major advantages because it reduces contamination of the intervention and better reflects clinical practice where a stewardship bundle gets implemented, than individual patient randomization. As a result, the study design as used has high generalizability. Fourth, the inclusion of patients without radiologically confirmed CAP may have diluted the effect to non-inferiority for mortality. However, as the study focusses on effectiveness rather than efficacy, the chosen study population (i.e., those being treated for presumed CAP) closely represents clinical practice. In addition, studies using low-dose CT scanning in addition to chest X-ray have demonstrated that radiological infiltrates are not apparent on chest radiographs in approximately 30% of patients with CAP based on CT findings[20]. This implies that many patients with CAP would be excluded if enrolment was based on presence of chest X-ray findings. Moreover, in the current study similar results were obtained in the subset of patients with radiologically confirmed CAP. Fifth, the study was performed in a setting with low prevalence of antimicrobial resistance and atypical pathogens, which may limit generalisability to settings with more resistance and atypical pathogens. However, *Streptococcus pneumoniae* with reduced susceptibility to penicillin can still be successfully treated with high dose penicillin's[21].

To conclude, a multifaceted antimicrobial stewardship intervention focused on education and audit and feedback reduced broad-spectrum antibiotic use by 27% in immunocompetent patients hospitalized with moderate-severe CAP and was non-inferior in terms of all-cause 90-day mortality. These results indicate that more patients with moderate-severe CAP can be safely treated with narrow-spectrum antibiotics.

LIST OF ABBREVIATIONS
BS; broad-spectrum antibiotics, CACE; Complier Average Causal Effect , CAP; Community-acquired pneumonia , CAP-PACT; Community-Acquired Pneumonia increasing Protocol adherence by Antibiotic stewardship in a stepped-wedge Cluster-randomized Trial, DOT; days of therapy, HR; hazard ratio, ICU; intensive care unit, IQT; interquartile range, ITT; intention-to-treat, i.v.; intravenous, NS; narrow-spectrum antibiotics, OR; odds ratio, PSI; pneumonia severity index, RD; risk difference.

CONFLICT OF INTERESTS
No conflict of interest to disclose.

PART II | CHAPTER 9

NARROW-SPECTRUM ANTIBIOTICS FOR COMMUNITY-ACQUIRED PNEUMONIA IN ADULTS: A STEPPED-WEDGE CLUSTER
RANDOMISED ANTIMICROBIAL STEWARDSHIP TRIAL

REFERENCES

1. Prina E, Ranzani OT, Torres A. Community-acquired pneumonia. *Lancet (London, England)*, 386(9998), 1097-1108 (2015).

2. Wunderink RG, Waterer GW. Clinical practice. Community-acquired pneumonia. *N Engl J Med*, 370(6), 543-551 (2014).

3. Wiersinga WJ, Bonten MJ, Boersma WG *et al.* Management of community-acquired pneumonia in adults: 2016 guideline update from the Dutch Working Party on Antibiotic Policy (SWAB) and Dutch Association of Chest Physicians (NVALT). *Neth J Med*, 76(1), 4-13 (2018).

4. Athlin S, Lidman C, Lundqvist A *et al.* Management of community-acquired pneumonia in immunocompetent adults: updated Swedish guidelines 2017. *Infect Dis (Lond)*, 50(4), 247-272 (2018).

5. Fally MM-W, U.; Lynghøj-Nielsen, T.; Stæhr Jensen, JU. Pneumoni – initial undersøgelse og behandling. (2017). https://www.lungemedicin.dk/fagligt/71-pneumoni.html

6. Lim WS, Baudouin SV, George RC *et al.* BTS guidelines for the management of community acquired pneumonia in adults: update 2009. *Thorax*, 64 Suppl 3, iii1-55 (2009).

7. Metlay JP, Waterer GW, Long AC *et al.* Diagnosis and Treatment of Adults with Community-acquired Pneumonia. An Official Clinical Practice Guideline of the American Thoracic Society and Infectious Diseases Society of America. *Am J Respir Crit Care Med*, 200(7), e45-e67 (2019).

8. Postma DF, van Werkhoven CH, van Elden LJ *et al.* Antibiotic treatment strategies for community-acquired pneumonia in adults. *N Engl J Med*, 372(14), 1312-1323 (2015).

9. Garau J. Treatment of drug-resistant pneumococcal pneumonia. *The Lancet. Infectious diseases*, 2(7), 404-415 (2002).

10. Aubier M, Verster R, Regamey C, Geslin P, Vercken JB. Once-daily sparfloxacin versus high-dosage amoxicillin in the treatment of community-acquired, suspected pneumococcal pneumonia in adults. Sparfloxacin European Study Group. *Clin Infect Dis*, 26(6), 1312-1320 (1998).

11. Petitpretz P, Arvis P, Marel M, Moita J, Urueta J. Oral moxifloxacin vs high-dosage amoxicillin in the treatment of mild-to-moderate, community-acquired, suspected pneumococcal pneumonia in adults. *Chest*, 119(1), 185-195 (2001).

12. Huijts SM, van Werkhoven CH, Boersma WG *et al.* Guideline adherence for empirical treatment of pneumonia and patient outcome. Treating pneumonia in the Netherlands. *Neth J Med*, 71(10), 502-507 (2013).

13. Davey P, Marwick CA, Scott CL *et al.* Interventions to improve antibiotic prescribing practices for hospital inpatients. *Cochrane Database Syst Rev*, 2, Cd003543 (2017).

14. Hemming K, Taljaard M, McKenzie JE *et al.* Reporting of stepped wedge cluster randomised trials: extension of the CONSORT 2010 statement with explanation and elaboration. *BMJ (Clinical research ed.)*, 363, k1614 (2018).

15. Dellit TH, Owens RC, McGowan JE, Jr. *et al.* Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America guidelines for developing an institutional program to enhance antimicrobial stewardship. *Clin Infect Dis*, 44(2), 159-177 (2007).

16. Kleinman LC, Norton EC. What's the Risk? A simple approach for estimating adjusted risk measures from nonlinear models including logistic regression. *Health Serv Res*, 44(1), 288-302 (2009).

17. Dunn G, Maracy M, Tomenson B. Estimating treatment effects from randomized clinical trials with noncompliance and loss to follow-up: the role of instrumental variable methods. *Stat Methods Med Res*, 14(4), 369-395 (2005).

18. Hulscher M, Prins JM. Antibiotic stewardship: does it work in hospital practice? A review of the evidence base. *Clin Microbiol Infect*, 23(11), 799-805 (2017).

19. Garin N, Genne D, Carballo S *et al.* beta-Lactam monotherapy vs beta-lactam-macrolide combination treatment in moderately severe community-acquired pneumonia: a randomized noninferiority trial. *JAMA internal medicine*, 174(12), 1894-1901 (2014).

20. Claessens YE, Debray MP, Tubach F *et al.* Early Chest Computed Tomography Scan to Assist Diagnosis and Guide Treatment Decision for Suspected Community-acquired Pneumonia. *Am J Respir Crit Care Med*, 192(8), 974-982 (2015).

21. Jacobs MR. Antimicrobial-resistant Streptococcus pneumoniae: trends and management. *Expert Rev Anti Infect Ther*, 6(5), 619-635 (2008).

PART II | CHAPTER 9

NARROW-SPECTRUM ANTIBIOTICS FOR COMMUNITY-ACQUIRED PNEUMONIA IN ADULTS: A STEPPED-WEDGE CLUSTER
RANDOMISED ANTIMICROBIAL STEWARDSHIP TRIAL

SUPPLEMENTARY MATERIAL

Table S1 Bacterial pathogens

	Control (N = 2235)		Intervention (N = 1849)	
	Proven	Possible	Proven	Possible
Streptococcus pneumoniae	205 (9.1%)	46 (2.0%)	188 (10.2%)	49 (2.7%)
Staphylococcus aureus	9 (0.4%)	32 (1.4%)	11 (0.6%)	45 (2.4%)
Other gram-positives	19 (0.9%)	18 (0.8%)	24 (1.3%)	5 (0.3%)
Haemophilus influenzae	4 (0.2%)	137 (6.1%)	7 (0.4%)	137 (7.4%)
Moraxella catarrhalis	-	20 (0.9%)	-	29 (1.6%)
Escherichia coli	17 (0.8%)	29 (1.3%)	13 (0.7%)	22 (1.2%)
Klebsiella pneumoniae	3 (0.1%)	11 (0.5%)	1 (0.1%)	18 (1.0%)
Pseudomonas aeruginosa	2 (0.1%)	35 (1.6%)	-	29 (1.6%)
Other gram-negatives	8 (0.4%)	65 (2.9%)	13 (0.7%)	78 (4.2%)
Legionella pneumophila	28 (1.2%)	-	22 (1.2%)	-
Mycoplasma pneumoniae	-	-	-	2 (0.1%)
Mycobacteria	-	2 (0.1%)	-	2 (0.1%)
Aspergillus species	-	16 (0.7%)	-	22 (1.2%)
Candida species	-	17 (0.8%)	-	25 (1.4%)
Other fungi / yeast	-	2 (0.1%)	1 (0.1%)	3 (0.2%)
No pathogen	811 (36.3%)	519 (23.2%)	650 (35.2%)	419 (22.7%)

Proven pathogens: based on pathogens detected in blood cultures, pleural fluid cultures, and urinary antigen tests (BINAX Now for
S. pneumoniae and L. pneumophila). Possible pathogens: based on pathogens detected in sputum cultures, bronchoalveolar lavage
fluid cultures, and serology.

Table S2 Stewardship implementation

	Intervention (n=1,849)
Education	
E-learnings completed	235 (45.2%)
Clinical lessons performed	54
Audit and feedback	
Patients eligible for feedback	591 (32.0 %)
Recommendations given	330 (55.8%)
Accepted	197 (59.7%)
Rejected	133 (40.3%)
Reasons for rejection advice:	
Antibiotics discontinued	5 (3.8 %)
Patient discharged or deceased	4 (3.0%)
Penicillin allergy	2 (1.5 %)
Legionella risk factors	2 (1.5 %)
Advice of microbiologist	1 (0.8%)
Severe pneumonia by PSI/CURB score / clinical deterioration	9 (6.8%)
COPD*	11 (8.3%)
Suspected resistant pathogen	9 (6.8%)
Treatment based on resistant pathogen in new culture	6 (4.5%)
Treatment based on resistant pathogen in old culture	2 (1.5 %)
Pneumococcal urine antigen test is negative	8 (6.0%)
Pneumococcal urine antigen test forgotten	5 (3.8%)
Supervisor wants to continue antibiotics	7 (5.3%)
Reason not clear	52 (39.0%)
Other reasons[1]	10 (7.5%)
Recommendations given by	
Telephone	244 (74.0%)
Medical record	17 (5.1%)
Both	69 (20.9%)
No recommendations given	261 (44.2%)
Reasons for no advice given	
Severe pneumonia by PSI/CURB score / clinical deterioration	56 (21.5%)
COPD*	12 (4.6%)
Suspected resistant pathogen	21 (8.0%)
Treatment based on resistant pathogen in old culture	39 (14.9%)
No time to give advice / missed	42 (16.1%)
Reason not clear	28 (10.7%)
Other reasons[2]	63 (24.1%)
Patients not eligible for feedback	1258 (64.0%)
Reasons for not eligible	
Started with narrow-spectrum antibiotics	866 (68.8%)
Switched to narrow-spectrum antibiotics	89 (7.1%)
Antibiotics discontinued	11 (0.9%)
Patient discharged or deceased	31 (2.5%)
Penicillin allergy	125 (9.9%)
Legionella risk factors	107 (8.5%)
Advice of microbiologist	29 (2.3%)
Second recommendations given	13 (3.9%)
Accepted	3 (0.9%)
Rejected	10 (3.0%)
Recommendations given by	
Telephone	6 (46.2%)
Medical record	5 (38.5%)
Both	2 (15.4%)

* Chronic Obstructive Pulmonary Disease, 1: Other reasons for rejecting advice; due to hospital-acquired pneumonia (not according
to official risk factors), due to possible other focus, due to recurrent pneumonia, post-obstructive pneumonia, bronchiectasis or
awaiting culture results.
2: Other reasons for no advice given: hospital-acquired pneumonia (not according to guideline risk factors), patient
immunocompromised (not according to our definitions), suspected empyema, patient is agitated, possible other focus, awaiting
culture results, post-obstructive pneumonia, due to legionella risk factors (not according to guideline risk factors), all cultures
negative, possible abscess, continuing treatment of general practitioner.

PART II | CHAPTER 9

NARROW-SPECTRUM ANTIBIOTICS FOR COMMUNITY-ACQUIRED PNEUMONIA IN ADULTS: A STEPPED-WEDGE CLUSTER
RANDOMISED ANTIMICROBIAL STEWARDSHIP TRIAL

Table S3 Crude, adjusted for design and fully adjusted relative risk reductions in broad-spectrum DOT

	Control (n=2,240) Median (IQR)	Intervention (n=1,844) Median (IQR)	Crude RRR (95% CI)	Adjusted for design RRR* (95% CI)	Fully adjusted** RRR (95% CI)
Broad-spectrum DOT	6 (2-9)	3 (0-8)	20.2% (9.2% - 29.8%)	28.1% (15.5% - 38.9%)	28.4% (16.2% - 38.9%)

* Adjusted for design and time, ** also adjusted for possible confounders.

Table S4 Crude intervention effect stratified per hospital

	Broad-spectrum DOT (median, IQR)			Empirical narrow-spectrum (proportion)	
	Control	Intervention	Adjusted risk difference	Control	Intervention
Hospital					
A	4 (1-9)	4 (1-9)	0,225	29,2	28,1
B	4 (0-8)	3 (0-8)	0,170	47,9	52,0
C	2 (0-7)	0 (0-5)	0,167	59,2	69,1
D	8 (3-9)	5 (2-9)	0,286	14,7	27,3
E	8 (3-11)	3 (0-8)	0,393	16,2	43,6
F	4 (0-8)	3 (0-8)	0,219	39,1	49,5
G	8 (4-11)	2 (0-8)	0.376	20,5	48,3
H	6 (0-9)	3 (0-8)	0,259	31,8	60,0
I	7 (2-9)	2 (0-8)	0,284	26,7	63,6
Total	6 (2-9)	3 (0-8)	0,284	28,8	45,9

Table S5 Specification of antibiotics given as empirical treatment

	Control (n=2,235)	Intervention (n=1,849)
Beta-lactams		
Penicillin	19 (0.9%)	29 (1.6%)
Amoxicillin	548 (24.5%)	745 (40.3%)
Amoxicillin – clavulanic acid	426 (19.0%)	244 (13.2%)
Ceftriaxone	114 (5.1%)	111 (6.0%)
Cefuroxime	149 (6.7%)	91 (4.9%)
Cefotaxime	-	-
Ceftazidime	12 (0.5%)	6 (0.3%)
Cefazolin	1 (0.0%)	-
Piperacillin/tazobactam	10 (0.4%)	2 (0.1%)
Macrolides		
Azithromycin	9 (0.4%)	3 (0.2%)
Clarithromycin	10 (0.4%)	-
Erythromycin	-	1 (0.1%)
Fluoroquinolones		
Moxifloxacin	62 (2.8%)	22 (1.2%)
Levofloxacin	4 (0.2%)	-
Ciprofloxacin	20 (0.9%)	13 (0.7%)
Tetracyclines		
Doxycycline	73 (3.3%)	66 (3.6%)
Other		
Co-trimoxazole	10 (0.4%)	12 (0.6%)
Clindamycin	2 (0.1%)	-
Combination therapy		
Amoxicillin + ciprofloxacin	153 (6.8%)	122 (6.6%)
Penicillin + ciprofloxacin	211 (9.4%)	159 (8.6%)
Amoxicillin– clavulanic acid + ciprofloxacin	99 (4.4%)	60 (3.2%)
Ceftriaxone + ciprofloxacin	40 (1.8%)	20 (1.1%)
Cefuroxime + ciprofloxacin	33 (1.5%)	11 (0.6%)
Cefuroxime + erythromycin	25 (1.1%)	5 (0.3%)
Cefuroxime + clarithromycin	23 (1.0%)	1 (0.1%)
Amoxicillin + cefuroxime	14 (0.6%)	14 (0.8%)
Amoxicillin + penicillin + ciprofloxacin	12 (0.5%)	9 (0.5%)
Amoxicillin– clavulanic acid + penicillin + ciprofloxacin	13 (0.6%)	12 (0.6%)
Other*		103 (5.6%)

*All other combination therapies below N = 10.

PART II | CHAPTER 9

NARROW-SPECTRUM ANTIBIOTICS FOR COMMUNITY-ACQUIRED PNEUMONIA IN ADULTS: A STEPPED-WEDGE CLUSTER
RANDOMISED ANTIMICROBIAL STEWARDSHIP TRIAL

Table S6 Crude, adjusted for design and fully adjusted risk differences in 90-day mortality

90-day mortality	Control (n=2,240)	Intervention (n=1,844)	Crude RD	Adjusted for design RD*	Fully adjusted** RD
N (%)	242 (10.9%)	199 (10.8%)			
Intention-to-treat					
90%CI			-0.09% (-1.7% ; 1.5%)	0.9% (-2.3% ; 3.2%)	0.4% (-2.7% ; 2.4%)
95%CI			-0.09% (-2.0% ; 1.8%)	0.9% (-2.9% ; 3.7%)	0.4% (-3.3% ; 2.8%)
As-treated					
90%CI			-3.1% (-4.6% ; -1.4%)	-3.1% (-4.8% ; -1.4%)	0.07% (-1.8% ; 1.7%)
95%CI			-3.1% (-4.9% ; -1.1%)	-3.1% (-5.1% ; -1.1%)	0.07% (-2.1% ; 2.1%)
CACE***					
90%CI			-0.5% (-12.6% ; 10.1%)	5.4% (-13.9% ; 19.1%)	2.2% (-15.8% ; 13.7%)
95%CI			-0.5% (-15.0% ; 12.1%)	5.4% (-17.0% ; 22.7%)	2.2% (-19.2% ; 16.4%)

* Adjusted for design and time, ** also adjusted for possible confounders, *** Complier Average Causal Effect.

Table S7 Crude, adjusted for design and fully adjusted risk differences in 30-day mortality

| 30-day mortality	Control (n=2,240)	Intervention (n=1,844)	Crude RD	Adjusted for design RD*	Fully adjusted** RD
N (%)	154 (6.9%)	123 (6.7%)			
Intention-to-treat					
90%CI			-0.3% (-1.6% ; 1.1%)	-0.7% (-2.9% ; 1.2%)	-1.1% (-3.1% ; 0.7%)
95%CI			-0.3% (-1.8% ; 1.3%)	-0.7% (-3.3% ; 1.6%)	-1.1% (-3.5% ; 1.1%)
As-treated					
90%CI			-2.6% (-4.0% ; -1.4%)	-2.6% (-4.0% ; -1.4%)	-0.3% (-1.8% ; 1.1%)
95%CI			-2.6% (-4.2% ; -1.1%)	-2.6% (-4.2% ; -1.1%)	-0.3% (-2.0% ; 1.4%)
CACE***					
90%CI			-1.5% (-9.1% ; 6.1%)	-3.9% (-17.0% ; 7.6%)	-6.4% (-18.5% ; 4.2%)
95%CI			-1.5% (-10.6% ; 7.5%)	-3.9% (-19.4% ; 10.0%)	-6.4% (-20.8% ; 6.3%)

* Adjusted for design and time, ** also adjusted for possible confounders, *** Complier Average Causal Effect.

Table S8 Complications

	Control (n=2,235)	Intervention (n=1,849)
Pleural effusion	398 (17.8)	326 (17.6)
Organ failure	78 (3.5)	87 (4.7)
Empyema	22 (1.0)	23 (1.2)
Septic shock	12 (0.5)	4 (0.2)
Acute respiratory distress syndrome	10 (0.5)	7 (0.4)
Pneumothorax	9 (0.4)	3 (0.2)
Lung abscess	9 (0.4)	1 (0.1)
Other	244 (10.9)	205 (11.1)
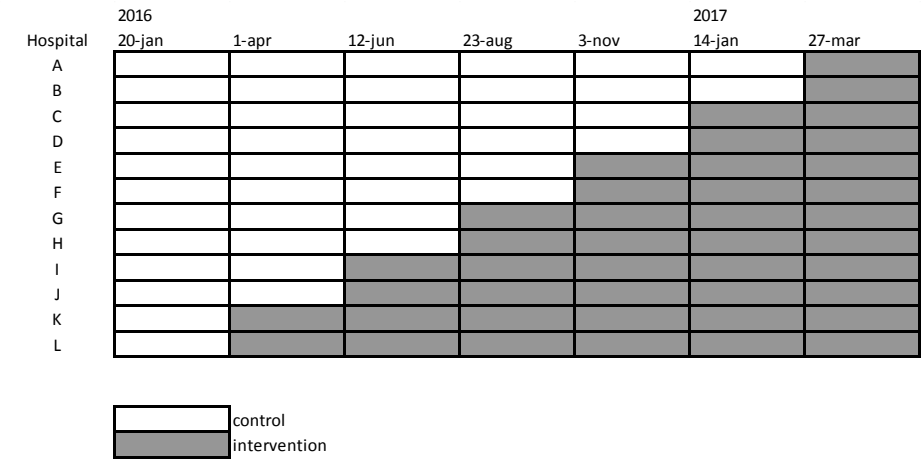
Figure S1 Stepped-wedge randomisation scheme
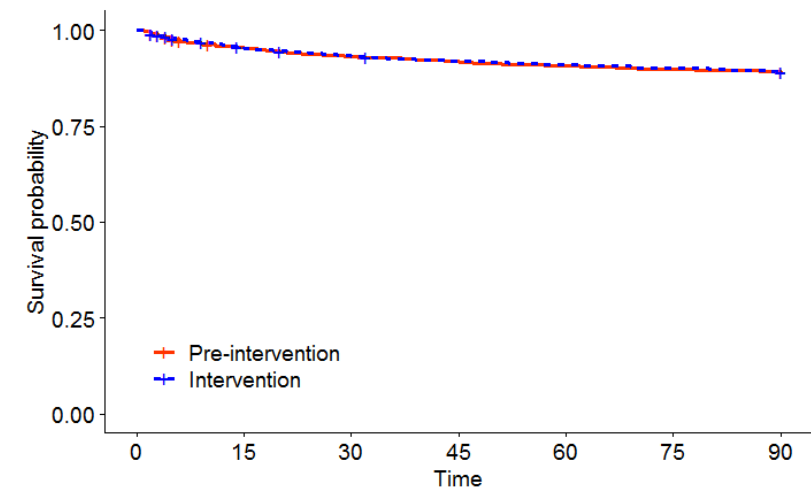
Figure S2 Kaplan Meier Curve

10

GENERAL DISCUSSION

GENERAL DISCUSSION

The studies presented in this thesis focus on the methodology of antimicrobial stewardship interventions and on the implementation of antimicrobial stewardship on optimising antibiotic use in moderate-severe community-acquired pneumonia patients (CAP). In the general discussion I would like to focus on the implications and future direction of antimicrobial stewardship and the management of moderate-severe CAP patients.

PART I: ANTIMICROBIAL STEWARDSHIP

Many antimicrobial stewardships studies that are being performed are of insufficient methodological quality to inform clinical practice. Systematic reviews on the effect of stewardship interventions recognize this problem[1,2]. Still, systematic reviews often select studies based on a limited number of pre-specified quality criteria, which are only fulfilled by <50% of published studies[1]. The quality of studies not fulfilling these quality criteria has never been described in detail. In chapter 2, we are the first to perform a comprehensive systematic review including all stewardship studies to assess their design quality[3]. This is important in order to identify areas of antimicrobial stewardship in which there is most need of improvement. The majority of studies were single center (63%) and used uncontrolled before-after study designs (50%). Studies fulfilled a median of 3 of the 10 (IQR 2-5) quality features and none of the studies fulfilled all 10. Only 44% of the studies reported clinical and 26% reported microbiological outcomes. Many of these studies are of insufficient methodological quality and therefore do not inform clinical practice. In fact, poorly conducted studies lead to considerable research waste[4]. This suboptimal use of valuable resources could be reduced by improving research design and scientific rigor[5]. A recent international working group identified optimizing the design of antimicrobial stewardship a top research priority[6]. Specifically, conducting robust evaluations of antimicrobial stewardship interventions and defining a balanced set of outcomes to measure the impact of antimicrobial stewardship interventions were identified as top priorities. Therefore, it is clear that guidance on how to optimize the design of stewardship studies is urgently needed. For this purpose, we formed an international working group containing clinical and academic specialists in antimicrobial stewardship and clinical trial design with the goal to provide recommendations about the optimal design for antimicrobial stewardship intervention studies (chapter 3)[7]. In a consensus procedure recommendations were formulated in three main domains: outcomes, objectives and study design. An important design aspect to consider is to where the study to be conducted fits on the efficacy-effectiveness-implementation spectrum. On this spectrum, efficacy and effectiveness studies investigate 'what' antibiotic prescribing practices should be adopted and implementation studies

investigate 'how' these practices should be implemented[8]. This consideration is important because it determines how outcomes are selected and prioritized. In general, research on healthcare interventions starts with proving efficacy, followed by effectiveness and ultimately implementation of the intervention[9]. Similarly, in antimicrobial stewardship, efficacy/effectiveness of the antimicrobial has to be established (e.g. antibiotic A is effective for disease B – the 'what') before interventions that improve adherence to this antimicrobial can be evaluated (e.g. stewardship intervention to improve adherence to antibiotic A in disease B – the 'how'). However, in practice the certainty of the efficacy of the antimicrobial, measured in clinical outcome of the patient (i.e. mortality, clinical cure), depends on the type of antimicrobial, infection, and specific patient setting under evaluation. Accordingly, the importance of evaluating efficacy or safety of the antimicrobial stewardship intervention is depend on 'what' the antimicrobial stewardship is set out to achieve. For example: if an antimicrobial stewardship intervention is aiming to increase compliance to amoxicillin use in moderate-severe CAP patients, and the evidence supporting the efficacy amoxicillin in this particular patient group is not well established, it is important to evaluate non-inferiority regarding safety in clinical outcomes. In contrast, if the antimicrobial stewardship intervention is aimed to increase adherence to a treatment that is clearly established as safe by high quality evidence in a similar patient population it might be justified to focus on process measures as the primary outcome. This is dependent on what the antimicrobial stewardship intervention aims to achieve. For example: the possibility for harm of an intervention aimed to increase adherence to stopping treatment or narrowing antibiotic spectrum might be higher than an intervention aimed to increase adherence to targeted treatment based on culture results. For this reason, meta-analyses that pool the results of clinical outcomes of different antimicrobial stewardship interventions with different aims, targeting adherence to different antibiotics, in different patients, diseases, and clinical settings are misleading[1].

The reproducibility crisis shows us that many results from randomised controlled trials (RCTs) can not be reproduced[10]. Therefore, it is important to continue collecting information and evidence on the efficacy and effectiveness of treatments. It is important to realize that efficacy studies (RCTs) on antimicrobials are generally performed in highly selected patient populations while antimicrobial stewardship interventions are generally performed in broader patient populations. Also, the sample size of RCTs is often not large enough to have sufficient power for important subgroup analyses. Non-inferiority trials require even larger sample sizes and novel methodologies such as the Response Adjusted for Days of Antibiotic Risk (RADAR) to increase the efficiency of these trials are not successful (chapter 4)[11]. As a result, the possibility of harm in the broader population, or in subgroups (for example immunocompromised patients), needs to be considered

even when efficacy of the antibiotic was demonstrated in RCTs. Studies that focus both on the efficacy/effectiveness of the antibiotic therapy and implementation of the antimicrobial stewardship intervention are called hybrid designs[9]. The major advantage of hybrid antimicrobial stewardship designs is that besides provdiing information on implementation, establishing safety of the antimicrobial stewardship intervention provides additional evidence of the efficacy of the antibiotic treatment that is implemented. Normally, estimating the treatment effect in observational studies is problematic due to residual confounding by indication. However, estimating unbiased treatment effects is possible with methods using instrumental variables. Instrumental variables are variables that are associated with the exposure (in our case what we are trying to achieve with our antimicrobial stewardship intervention, i.e. adherence to antibiotic A) but not associated with the outcome, except through its association with the exposure[12]. In designs where implementation of the antimicrobial stewardship intervention is randomised, the randomisation variable itself may be used as an instrumental variable to estimate the efficacy of the antibiotic therapy being implemented using a method called the complier average causal effect (CACE)[13,14]. The efficiency of this method depends on the amount of change in antibiotic therapy achieved by the antimicrobial stewardship intervention. For example: if an antimicrobial stewardship intervention changes the targeted antibiotic use from 20% to 40% the estimated CACE is not precise (resulting in wide 95% confidence intervals), while if the antibiotic use changed from 10% to 90% the estimated CACE is precise (narrow 95% confidence intervals). Yet, even if the change in antibiotic use was minimal, precise estimates can be obtained with large sample sizes. In the era of big data, where large quantities of routinely collected healthcare data can be automatically extracted, increasing the sample size should become less of a problem. Therefore, using appropriate methodology, there is great opportunity of future antimicrobial stewardship implementation hybrid studies to not only contribute knowledge on optimizing implementation but also build evidence on efficacy of what the antimicrobial stewardship intervention is set out to achieve (i,e, antibiotic therapies), both in broader patient populations and specific subgroups where randomized controlled trials are not feasible.

PART II: COMMUNITY-ACQUIRED PNEUMONIA

The optimal treatment of patients with moderate-severe CAP is subject to discussion. The choice of preferred empirical therapy varies globally. Guidelines in The United States of America recommend beta-lactam macrolide combination therapy or a respiratory fluoroquinolone[15], British guidelines recommend amoxicillin macrolide combination therapy[16], and Dutch, Swedish and Danish guidelines recommend amoxicillin/penicillin monotherapy[17,18]. Two high quality RCTs compare beta-lactam monotherapy with beta-lactam macrolide combination therapy providing different conclusions.

One randomised controlled non-inferiority trial failed to show non-inferiority with a nonsignificant increase in 30-day mortality of 1.4% (p=.42), but with more rapid clinical stability in beta-lactam macrolide combination therapy (4.5 days) compared to beta-lactam monotherapy (5 days)[19]. The other, called the CAP-START trial, was a cluster randomised non-inferiority trial that showed non-inferiority of a strategy of preferred empirical treatment with beta-lactam monotherapy compared to beta-lactam macrolide combination therapy and fluoroquinolone monotherapy regarding 90-day mortality[20]. If we focus on mortality as an outcome, one study provides evidence for the safety of beta-lactam monotherapy while the other is inconclusive. Surprisingly, a prominent international systematic review published after these two studies concluded that empirical treatment for moderate-severe CAP should consist of beta-lactam macrolide combination therapy or fluoroquinolone monotherapy[21]. This conclusion was largely based on observational studies and therefore discarded the only high quality conclusive trial available, which is the CAP-START trial. In addition, the international community has expressed several points of critique on the CAP-START trial[22]. First, in the beta-lactam monotherapy group broader-spectrum antibiotics with atypical coverage were used in 27% of the patients. However, the reduction in the use of antibiotics with atypical coverage by 57% to 62% compared to the other strategies was not appreciated. In addition, the trial pragmatically evaluated an empirical strategy where deviation from the strategy was possible. This closely resembles clinical practice, where guidelines are not followed blindly but deviations are possible. This point is therefore not a limitation but increases generalizability of the study results. Second, the chosen non-inferiority of 3% was considered to be inappropriately large for an event rate of 10%. However, compared to other recent clinical trials where a non-inferiority margin of 10% on an event rate of 10%, and a non-inferiority margin of 5% on an event rate of 5% were used, the non-inferiority margin was actually quite stringent[23,24]. Thirdly, the study included patients with a clinical diagnosis of CAP and this included patients without radiological evidence of pneumonia on chest X-ray. However, recent findings from a study using low-dose CT scanning in addition to chest X-ray demonstrated that in ~30% of patients with CAP radiological infiltrates are not apparent on chest radiographs, meaning that many CAP patients may be missed if only radiologically confirmed CAP patients are enrolled[25]. Lastly, the choice for 90-day mortality as the primary endpoint was deemed inappropriate because for non-severe CAP mortality due to uncontrolled infection is unlikely and probably caused by acute cardiovascular events or chronic comorbidities. Yet, the 90-day mortality was 10% which is not negligible and it is possible that acute cardiovascular events and exacerbation of chronic comorbidities are triggered by the underlying infection. Also, in cluster randomised trials where blinding is not feasible, selecting objective outcomes such as all-cause mortality is required to avoid information bias and therefore more subjective outcomes such as clinical cure are not appealing.

Ultimately, these misconceptions about the CAP-START trial led to the study not even being mentioned in the latest update of the CAP guideline by IDSA[15].

The CAP-START trial showed that beta-lactam monotherapy is non-inferior to beta-lactam macrolide combination therapy and fluoroquinolone monotherapy. However, the recommended empirical therapy in Dutch, Swedish and Danish guidelines is even more narrow, namely narrow-spectrum beta-lactam monotherapy (amoxicillin or penicillin). Even though this was also the recommended empirical therapy during the CAP-START beta-lactam monotherapy periods, only 22% of the patients were actually treated with narrow-spectrum beta-lactam monotherapy (unpublished). Most patient in the beta-lactam monotherapy period received amoxicillin-clavulanic acid or cephalosporins. Therefore, the CAP-START trial teaches us nothing about the efficacy of narrow-spectrum beta-lactams. Thus far, two RCTs investigated the efficacy of narrow-spectrum beta-lactams compared to respiratory fluoroquinolones[26,27]. One RCT compared moxifloxacin to amoxicillin in patients with mild-to-moderate suspected pneumococcal CAP, which showed similar clinical success rates of 86.5% (173/200) and 82.2% (171/208) in the moxifloxacin and amoxicillin treated patients, respectively[26]. In the other RCT, sparfloxacin showed clinical cure rates of 83.6% (133/159) compared to 84.7% (144/170) with amoxicillin in similar patients[27]. However, both studies included only patients with mild disease and one used a fluoroquinolone no longer used in clinical practice. Therefore, it is unclear how these results translate to clinical practice and the wider population of moderate-severe CAP patients. The rationale to recommend narrow-spectrum beta-lactam monotherapy as empirical therapy is twofold. First, the most commonly identified pathogen in moderate-severe CAP patients is *Streptococcus pneumoniae*, which is virtually always susceptible to narrow-spectrum beta-lactams. In addition, it is assumed that a large proportion of culture negative CAPs is also caused by *Streptococcus pneumoniae.* Second, the severity of illness in these patients allows to start with narrow-spectrum therapy and evaluate the clinical response and escalate when patients do not improve or change to targeted therapy based on diagnostic results. However, the confidence of clinical doctors in the recommendation to start with narrow-spectrum beta-lactam is low, which is reflected by low guideline adherence[28]. We performed a stepped-wedge cluster randomised antimicrobial stewardship intervention trial in 12 hospitals in the Netherlands with the aim to reduce broad-spectrum antibiotic use (Days of Therapy (DOT)) and show that this reduction is safe for patients (90-day mortality)(**chapter 9**). In this trial, we included immunocompetent adult patients with moderate-severe CAP. After a baseline period where standard care was measured, we implemented a multifaceted antimicrobial stewardship intervention bundle in a stepwise fashion. The multifaceted intervention consisted of (1) education, (2) motivating opinion leaders, and (3) audit and feedback. The intervention reduced the median broad-

spectrum DOT from 6 (interquartile range (IQR) 2 – 9) in the control period to 3 (IQR 0 – 8) in the intervention period. The adjusted relative reduction in broad-spectrum DOT during intervention was 26.9% (95% CI: 15.4%-37.4%). The 90-day mortality was similar during control and intervention period, with an adjusted absolute difference of 0.4% (90% CI: -2.7 to 2.4) for the control versus intervention period, indicating non-inferiority for all-cause mortality and safety of the intervention. These results indicate that doctors may safely prescribe more narrow-spectrum beta-lactam therapy in patients with moderate-severe CAP. However, how many patients can be empirically treated with narrow-spectrum beta-lactam antibiotics before safety becomes an issue? Guidelines are not made to be strictly followed without deviations. In CAP patients there are certain medical reasons where deviation from narrow-spectrum beta-lactam is appropriate, such as risk factors for Legionella and not showing clinical improvement on earlier beta-lactam therapy. Future studies on the treatment of moderate-severe CAP patients should therefore focus on clearly defining in which patients narrow-spectrum beta-lactam is safe and which patients require broader-spectrum antibiotics, similar to the principles of personalized medicine.

One patient group that has been investigated for possibly benefitting of broader-spectrum treatment are patients with recent contact to healthcare. Pneumonia in these patients is described as Healthcare-associated pneumonia (HCAP). Due to their recent contact with healthcare, patients with HCAP may be at increased risk for being colonized with antibiotic-resistant and healthcare associated pathogens. To investigate this, we performed a post-hoc analysis of a cohort of patients with moderate-severe CAP in which we compared the *in vitro* susceptibility of causative pathogens in CAP patients versus HCAP patients for amoxicillin and broad-spectrum antibiotics (**chapter 6**)[29]. Indeed, HCAP predicts for higher non-susceptibility to amoxicillin but we considered the difference too little to justify treating all these patients with broad-spectrum antibiotics. In another explorative post-hoc analysis of three international prospective cohorts of non-severe CAP patients we investigated whether various clinical variables could predict better treatment effects for empirical treatment with either beta-lactam monotherapy, beta-lactam macrolide combination therapy or fluoroquinolone-based therapy **(chapter 7)**[30]. After correcting for confounders, older age was associated with lower effectiveness of fluoroquinolones for 30-day mortality (interaction OR 1.67, 95% CI 1.23-2.29) and beta-lactam macrolide combination therapy with lower effectiveness for length of stay (interaction effect ratio 1.14, 95% CI 1.06-1.22). Current smoking was associated with lower effectiveness of fluoroquinolones for 30-day mortality (interaction OR 2.36, 95% CI 1.34-4.17). These results indicate that older age and smoking could possibly modify the effect of empirical antibiotic treatment. Future RCTs are needed to investigate whether these predictors could be used to individualize antibiotic treatment

in patients with CAP.

Improvements in diagnostics will change the way future CAP patients will be treated. Faster identification of causative pathogens by Point-of-Care (PoC) diagnostics means that antibiotic treatment can be switched from empirical to targeted, often more narrow, treatment earlier. Rapid tests can therefore make an important contribution to antimicrobial stewardship. Surprisingly, the use of Legionella and Pneumococcal urine antigen testing for moderate-severe CAP is discouraged in the latest IDSA CAP guideline[15]. They argue that urine antigen tests do not provide a direct benefit on clinical outcomes such as mortality. However, they fail to recognize the antimicrobial stewardship opportunity that antigen testing provides, as often therapy can be narrowed significantly if the test results are positive. In CAP patients, there is still much diagnostic uncertainty. In up to 50% of patients a causal pathogen is not identified[20]. New PoC Multiplex Polymerase Chain Reaction (PCR) tests are able to quickly identify bacteria and viruses in respiratory samples. This may allow for rapid targeted therapy or discontinuation of antibiotic therapy if a viral pathogen is identified and therefore viral pneumonia is suspected. In addition, as mentioned previously, with the introduction of novel imaging techniques such as low-dose CT, we observe that many CAP patients are misdiagnosed using conventional chest X-ray[25]. In a recent study, low-dose CT scans revealed no infiltrate in 30% of patients with an infiltrate on chest X-ray, while an infiltrate was detected in 33% of patients without an infiltrate on chest X-ray[25]. As a consequence, it is possible that implementation of low-dose CT instead of chest X-ray will have consequences for the antibiotic therapy of patients as alternative diagnoses need to be considered. How PoC PCR testing and low-dose CT scans will influence the antibiotic management and clinical outcomes of patients with moderate-severe CAP patients is currently being investigated (NCT03360851).

CONCLUSION

To conclude, the methodology to evaluate the efficacy of antimicrobial stewardship interventions needs to be improved. With concrete recommendations and appropriate methodology, implementation of antimicrobial stewardship has the opportunity to generate high quality evidence for implementation and for the efficacy of antibiotic strategies. Antimicrobial stewardship is an effective tool to optimize antibiotic treatment in moderate-severe CAP patients. More patients with moderate-severe CAP can be treated with narrow-spectrum beta-lactam monotherapy without compromising patient outcomes. Future studies should focus on determining which patients can be safely treated with narrow-spectrum beta-lactam and which patients would benefit

from broader treatment. New diagnostic strategies such as PoC PCR and low-dose CT can help to optimize management of moderate-severe CAP patients and move to a more individualized approach.

REFERENCES

1. Davey P, Marwick CA, Scott CL, et al. Interventions to improve antibiotic prescribing practices for hospital inpatients. *Cochrane Database Syst Rev*. 2017;2017(2). doi:10.1002/14651858.CD003543.pub4

2. Schuts EC, Hulscher MEJL, Mouton JW, et al. Current evidence on hospital antimicrobial stewardship objectives: A systematic review and meta-analysis. *Lancet Infect Dis*. 2016;16(7):847-856. doi:10.1016/S1473-3099(16)00065-7

3. Schweitzer VA, van Heijl I, van Werkhoven CH, et al. The quality of studies evaluating antimicrobial stewardship interventions: a systematic review. *Clin Microbiol Infect*. 2019. doi:10.1016/j.cmi.2018.11.002

4. Macleod MR, Michie S, Roberts I, et al. Biomedical research: Increasing value, reducing waste. *Lancet*. 2014. doi:10.1016/S0140-6736(13)62329-6

5. Ioannidis JPA, Greenland S, Hlatky MA, et al. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet*. 2014. doi:10.1016/S0140-6736(13)62227-8

6. Rzewuska M, Charani E, Clarkson JE, et al. Prioritizing research areas for antibiotic stewardship programmes in hospitals: a behavioural perspective consensus paper. *Clin Microbiol Infect*. 2019. doi:10.1016/j.cmi.2018.08.020

7. Schweitzer VA, van Werkhoven CH, Rodríguez Baño J, et al. Optimizing design of research to evaluate antibiotic stewardship interventions: consensus recommendations of a multinational working group. *Clin Microbiol Infect*. 2019. doi:10.1016/j.cmi.2019.08.017

8. Hulscher MEJL, Prins JM. Antibiotic stewardship: does it work in hospital practice? A review of the evidence base. *Clin Microbiol Infect*. 2017. doi:10.1016/j.cmi.2017.07.017

9. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: Combining elements of clinical effectiveness and implementation research to enhance public health impact. *Med Care*. 2012. doi:10.1097/MLR.0b013e3182408812

10. Krauss A. Why all randomised controlled trials produce biased results. *Ann Med*. 2018. doi:10.1080/07853890.2018.1453233

11. Schweitzer VA, van Smeden M, Postma DF, Oosterheert JJ, Bonten MJM, van Werkhoven CH. Response Adjusted for Days of Antibiotic Risk (RADAR): evaluation of a novel method to compare strategies to optimize antibiotic use. *Clin Microbiol Infect*. 2017. doi:10.1016/j.cmi.2017.05.003

12. Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol*. 2000. doi:10.1093/ije/29.4.722

13. Hewitt CE, Torgerson DJ, Miles JNV. Is there another way to take account of noncompliance in randomized controlled trials? *CMAJ*. 2006. doi:10.1503/cmaj.051625

14. Dunn G, Maracy M, Tomenson B. Estimating treatment effects from randomized clinical trials with noncompliance and loss to follow-up: The role of instrumental variable methods. *Stat Methods Med Res*. 2005. doi:10.1191/0962280205sm403oa

15. Metlay JP, Waterer GW, Long AC, et al. Diagnosis and Treatment of Adults with Community-acquired Pneumonia. An Official Clinical Practice Guideline of the American Thoracic Society and Infectious Diseases Society of America. *Am J Respir Crit Care Med*. 2019. doi:10.1164/rccm.201908-1581st

16. Lim WS, Baudouin S V, George RC, et al. BTS guidelines for the management of community acquired pneumonia in adults: update 2009. *Thorax*. 2009;64(Suppl 3):iii1-iii55. doi:10.1136/thx.2009.121434

17. Athlin S, Lidman C, Lundqvist A, et al. Management of community-acquired pneumonia in immunocompetent adults: updated Swedish guidelines 2017. *Infect Dis (Auckl)*. 2018. doi:10.1080/23744235.2017.1399316

18. Wiersinga WJ, Bonten MJ, Boersma WG, et al. Management of community-acquired pneumonia in adults: 2016 guideline update from the dutch working party on antibiotic policy (SWAB) and dutch association of chest physicians (NVALT). *Neth J Med*. 2018.

19. Garin N, Genné D, Carballo S, et al. -Lactam Monotherapy vs -Lactam-Macrolide Combination Treatment in Moderately Severe Community-Acquired Pneumonia: A Randomized Noninferiority Trial. *JAMA Intern Med*. 2014;174(12):1894-1901. doi:10.1001/jamainternmed.2014.4887

20. Postma DF, van Werkhoven CH, van Elden LJR, et al. Antibiotic Treatment Strategies for Community-Acquired Pneumonia in Adults. *N Engl J Med*. 2015;372(14):1312-1323. doi:10.1056/NEJMoa1406330

21. Lee JS, Giesler DL, Gellad WF, Fine MJ. Antibiotic Therapy for Adults Hospitalized With Community-Acquired Pneumonia A Systematic Review. *Jama*. 2016;315(6):593-602. doi:10.1001/jama.2016.0115

22. Wunderink RG, Waterer G. Advances in the causes and management of community acquired pneumonia in adults. *BMJ*. 2017. doi:10.1136/bmj.j2471

23. Iversen K, Ihlemann N, Gill SU, et al. Partial oral versus intravenous antibiotic treatment of endocarditis. *N Engl J Med*. 2019. doi:10.1056/NEJMoa1808312

24. Li HK, Rombach I, Zambellas R, et al. Oral versus intravenous antibiotics for bone and joint infection. *N Engl J Med*. 2019. doi:10.1056/NEJMoa1710926

25. Claessens YE, Debray MP, Tubach F, et al. Early chest computed tomography scan to assist diagnosis and guide treatment decision for suspected community-acquired pneumonia. *Am J Respir Crit Care Med*. 2015;192(8):974-982. doi:10.1164/rccm.201501-0017OC

26. Petitpretz P, Arvis P, Marel M, et al. Oral moxifloxacin vs high-dosage amoxicillin in the treatment of mild-to-moderate, community-acquired, suspected pneumococcal pneumonia in adults. *Chest*. 2001;119(1):185-195. doi:10.1378/chest.119.1.185

27. Aubier M, Verster R, Regamey C, Geslin P, Vercken JB. Once-daily sparfloxacin versus high-dosage amoxicillin in the treatment of community-acquired, suspected pneumococcal pneumonia in adults. *Clin Infect Dis*. 1312;26(6):1312-1323.

28. Huijts SM, Van Werkhoven CH, Boersma WG, et al. Guideline adherence for empirical treatment of pneumonia and patient outcome treating pneumonia in the netherlands. *Neth J Med*. 2013.

29. Schweitzer VA, van Werkhoven CH, van Heijl I, et al. Relevance of healthcare-associated pneumonia for empirical antibiotic therapy in the Netherlands. *Neth J Med*. 2018.

30. Simonetti AF, van Werkhoven CH, Schweitzer VA, et al. Predictors for individual patient antibiotic treatment effect in hospitalized community-acquired pneumonia patients. *Clin Microbiol Infect*. 2017. doi:10.1016/j.cmi.2017.03.010

11

NEDERLANDSE SAMENVATTING

Infecties door bacteriën kunnen behandeld worden met antibiotica. Indien bacteriën in contact komen met antibiotica dan bestaat er een kans dat deze bacteriën resistent worden voor deze antibiotica. Wanneer de antibiotica niet meer werkt, noemt men dit *antibioticaresistentie*. De mate van antibioticaresistentie is de afgelopen jaren sterk toegenomen. Eén van de redenen van de toename van antibioticaresistentie is het overmatig en onjuist gebruik van antibiotica. *Overmatig antibioticagebruik* kan teveel antibiotica betekenen of antibiotica met een onnodige brede werking (spectrum). Onder *juist antibioticagebruik* verstaan we dat er een balans moet zijn tussen het geven van te weinig antibiotica en het geven van teveel antibiotica. Bij het geven van te weinig antibiotica wordt de onderliggende infectie niet goed behandeld, terwijl bij het geven van teveel antibiotica onnodige resistentie ontstaat. Het nastreven van dit passende antibioticagebruik wordt *antimicrobial stewardship* genoemd. In de praktijk zijn er verschillende antimicrobial stewardship methoden om antibioticagebruik te verbeteren. De meest gebruikte en bewezen effectieve methoden zijn: educatie voor dokters die antibiotica voorschrijven, en terugkoppeling geven aan voorschrijvende dokters in het geval dat ze een verkeerd antibioticum voorschrijven. In het eerste gedeelte (**hoofdstuk 2, 3 en 4**) van dit proefschrift richten we ons op antimicrobial stewardship. Door de toename in antibioticaresistentie wordt het uitvoeren van antimicrobial stewardship steeds belangrijker. Het is dan ook niet verwonderlijk dat er de afgelopen jaren veel wetenschappelijke studies zijn verschenen over de effectiviteit van antimicrobial stewardship. In veel van deze studies wordt gekeken hoe effectief bepaalde antimicrobial stewardship interventies zijn (zoals bijvoorbeeld educatie) in het verbeteren van het antibioticagebruik. Veel van deze studies zijn echter van lage wetenschappelijke kwaliteit. In **hoofdstuk 2** beschrijven we de wetenschappelijke kwaliteit van studies over antimicrobial stewardship interventies en identificeren we de tekortkomingen in de wetenschappelijke kwaliteit. Deze tekortkomingen vormen de basis voor de samenkomst van een internationale groep experts op het gebied van antimicrobial stewardship, onderzoeksmethodologie en statistiek. Het doel van deze expertgroep is om een consensus document te ontwikkelen waarin concrete aanbevelingen staan om kwalitatief goede wetenschappelijk studies te doen naar antimicrobial stewardship interventies. Deze aanbevelingen worden beschreven in **hoofdstuk 3**. Eén van de meest opvallende tekortkomingen is dat veel wetenschappelijke studies naar antimicrobial stewardship interventies niet kijken of het verminderen van antibioticagebruik ook daadwerkelijk veilig is voor patiënten. Het is namelijk mogelijk dat door het verminderen van antibioticagebruik er uiteindelijk niet goed genoeg behandeld wordt. Hierdoor kan de onvoldoende behandelde infectie zorgen voor nadelige gevolgen voor de patiënt, bijvoorbeeld een langere ziekenhuisopname of zelfs overlijden ten gevolge van de infectie. Een studie kan succesvol zijn in het verminderen van antibioticagebruik, maar als er twijfel is of dit veilig is voor patiënten dan zullen dergelijke antimicrobial stewardship interventies niet snel worden toegepast in de klinische praktijk. Eén van de aanbevelingen is dan ook dat er van tevoren goed moet worden nagedacht over mogelijk nadelige gevolgen voor patiënten, en indien hier twijfel over bestaat altijd data te verzamelen over klinische uitkomsten van patiënten. In **hoofdstuk 4** onderzoeken we een gepubliceerde methode die de uitkomstmaten van minder antibioticagebruik weet te combineren. Bij deze nieuwe methode wordt een samengestelde uitkomstmaat gemaakt die zowel antibioticagebruik als patiënten uitkomsten bevat. We passen deze nieuwe methode toe op data van een eigen antimicrobial stewardship interventie onderzoek om te zien of de methode in staat is om slechtere uitkomsten van patiënten te identificeren. Onze conclusie is dat de nieuwe methode dit niet goed kan en we adviseren daarom deze methode niet als belangrijkste uitkomst te gebruiken in antimicrobial stewardship interventie onderzoek.

Het tweede gedeelte (**hoofdstuk 5, 6, 7, 8, 9**) van dit proefschrift gaat over het toepassen van antimicrobial stewardship bij patiënten met longontsteking. *Longontsteking* is een ontsteking van de lagere luchtwegen die zowel mild als ernstig kan verlopen. Indien patiënten buiten het ziekenhuis een longontsteking oplopen noemen we dat een *thuis-opgelopen longontsteking*. Dit in tegenstelling tot de *ziekenhuis-opgelopen longontsteking*. Dit onderscheid is belangrijk omdat de verschillende soorten longontsteking doorgaans door verschillende bacteriën worden veroorzaakt. In ons onderzoek richten wij ons uitsluitend op mensen met een thuis-opgelopen longontsteking. Patiënten die in het ziekenhuis belanden met een thuis-opgelopen longontsteking hebben een aanzienlijke kans om aan deze aandoening te overlijden. De hoeksteen van de behandeling van patiënten met thuis-opgelopen longontsteking is antibiotica. Meestal weten we nog niet door welke bacterie de longontsteking veroorzaakt wordt wanneer een patiënt wordt opgenomen met longontsteking. De antibiotica die in eerste instantie wordt gegeven is daarom altijd gericht op de groep bacteriën die het vaakst voorkomen, dit zijn vaak *breed-spectrum antibiotica* (werkzaam tegen een breed spectrum van bacteriën). Tijdens de opname wordt er vervolgens onderzoek verricht, zoals het kweken van de bacterie, om erachter te komen welke bacterie de daadwerkelijke veroorzaker is. Als er een bacterie gevonden wordt, dan zal het antibioticum aangepast worden naar een antibioticum die specifiek gericht is op deze bacterie, dit zijn vaak *smal-spectrum antibiotica* (werkzaam tegen een smal spectrum van bacteriën). Dit toespitsen van de antibiotica op de gevonden bacterie noemen we het *de-escaleren van de antibiotica*. Er worden veel wetenschappelijke onderzoeken gepubliceerd die onderzoeken of dit de-escaleren veilig is voor patiënten. In **hoofdstuk 5** van dit proefschrift bespreken we wederom de wetenschappelijke

tekortkomingen van gepubliceerde studies, dit keer van studies die de veiligheid van de-escalatie onderzoeken. In 2005 werd er in Amerika een nieuwe categorie van longontstekingen benoemd, een tussenvorm van thuis-opgelopen longontsteking en ziekenhuis-opgelopen longontsteking. Dit betrof patiënten die de longontsteking wel thuis hadden opgelopen, maar frequent in aanraking zijn geweest met zorg. De gedachten hierachter was dat door het zorgcontact deze patiënten meer resistente bacteriën bij zich zouden dragen en dat deze resistente bacteriën de oorzaak zouden kunnen zijn van de longontsteking, waardoor er andere antibiotica nodig is. Deze categorie werd *zorg-gerelateerde-longontsteking* genoemd. In **hoofdstuk 6** onderzoeken we of in Nederland de longontsteking van patiënten die binnen de categorie zorg-gerelateerde-longontsteking vallen ook veroorzaakt wordt door meer resistente bacteriën en of er daarom andere antibiotica nodig is. Uit ons onderzoekt blijkt dat er kleine verschillen zijn in de bacteriën die zorg-gerelateerde-longontsteking veroorzaken, maar dat deze verschillen niet groot genoeg zijn om al deze patiënten anders te behandelen. In **hoofdstuk 7** onderzoeken we of er bepaalde patiëntkarakteristieken zijn die kunnen voorspellen of een patiënt beter zal reageren op een bepaald soort antibioticum. Uit dit onderzoek blijkt dat oudere patiënten en patiënten die roken wellicht baat zouden kunnen hebben van andere antibiotica in vergelijking met jongere patiënten en patiënten die niet roken. In **hoofdstuk 8** beschrijven we mogelijke manieren om door middel van antimicrobial stewardship het antibioticagebruik bij patiënten met thuis-opgelopen longontsteking te verbeteren. Tot slot, in **hoofdstuk 9**, presenteren we de resultaten van een studie waarbij we in 12 Nederlandse ziekenhuizen een antimicrobial stewardship interventie implementeren om het antibioticagebruik bij patiënten met thuis-opgelopen longontsteking te verbeteren. Het merendeel van de thuis-opgelopen longontsteking wordt veroorzaakt door de bacterie *Streptococcus pneumoniae*. Deze bacterie is goed te behandelen met de smal-spectrum antibiotica penicilline of amoxicilline. Dit is de reden dat Nederlandse richtlijnen adviseren om deze patiënten initieel met deze smal-spectrum antibiotica te behandelen, ook als we de daadwerkelijke veroorzakende bacterie van de longontsteking nog niet hebben aangetoond. In de praktijk blijkt echter dat dokters vaak breed-spectrum antibiotica geven, wat kan leiden tot onnodige vorming van antibioticaresistentie. In onze studie proberen we het breed-spectrum antibioticagebruik te verminderen door middel van educatie voor artsen en door terugkoppeling indien er onnodig breed-spectrum antibiotica is gegeven. Daarnaast willen we aantonen dat deze reductie in breed-spectrum antibiotica veilig is voor patiënten en dat we de patiënten hierbij niet onderbehandelen. Onze antimicrobial stewardship interventie was in staat om het breed-spectrum antibioticagebruik met 27% te verminderen, zonder dat dit nadelige effecten had voor patiënten. Deze resultaten kunnen ervoor zorgen dat dokters meer vertrouwen krijgen in het voorschrijven van smal-spectrum antibiotica bij patiënten met thuis-opgelopen

longontsteking, wat zal zorgen voor minder onnodige toename in antibioticaresistentie. Dit proefschrift laat zien dat antimicobial stewardship een effectieve manier is om antibioticagebruik te verbeteren. De wetenschappelijke kwaliteit van antimicrobial stewardship interventie studies kan echter beter. We hebben nu concrete aanbevelingen om deze kwaliteit te verbeteren. Antimicrobial stewardship interventies zijn in staat om bij patiënten met thuis-opgelopen longontsteking het gebruik van breed-spectrum antibiotica aanzienlijk te verminderen, zonder dat dit nadelige gevolgen heeft voor patiënten. Door verstandiger om te gaan met het gebruik van antibiotica kan de toename in antibioticaresistentie geremd worden.

LIST OF PUBLICATIONS

van Heijl I, **Schweitzer VA**, Boel CHE, Oosterheert JJ, Huijts SM, Dorigo-Zetsma W, van der Linden PD, Bonten MJM, van Werkhoven CH. Confounding by indication of the safety of de-escalation in community-acquired pneumonia: A simulation study embedded in a prospective cohort. PLoS ONE. 2019;14(9):e0218062.

Schweitzer VA, van Werkhoven CH, Rodríguez Baño J, Bielicki J, Harbarth S, Hulscher M, Huttner B, Islam J, Little P, Pulcini C, Savoldi A, Tacconelli E, Timsit JF, van Smeden M, Wolkewitz M, Bonten MJM, Walker AS, Llewelyn MJ; Joint Programming Initiative on Antimicrobial Resistance (JPIAMR) Working Group on Design of Antimicrobial Stewardship Evaluations. Optimizing design of research to evaluate antibiotic stewardship interventions: consensus recommendations of a multinational working group. Clin Microbiol Infect. 2020;26(1):41-50.

Burgmeijer EH, Duijkers R, Lutter R, Bonten MJM, **Schweitzer VA**, Boersma WG. Plasma cytokine profile on admission related to aetiology in community-acquired pneumonia. Clin Respir J. 2019;13(10):605-613.

Schweitzer VA, van Heijl I, van Werkhoven CH, Islam J, Hendriks-Spoor KD, Bielicki J, Bonten MJM, Walker AS, Llewelyn MJ; Consensus on Antimicrobial Stewardship Evaluations (CASE) study group. The quality of studies evaluating antimicrobial stewardship interventions: a systematic review. Clin Microbiol Infect. 2019;25(5):555-561.

Schweitzer VA, van Werkhoven CH, van Heijl I, Smits RF, Boel CHE, Bonten MJM, Postma DF, Oosterheert JJ. Relevance of healthcare-associated pneumonia for empirical antibiotic therapy in the Netherlands. Neth J Med. 2018;76(9):389-396.

van Heijl I, **Schweitzer VA**, Zhang L, van der Linden PD, van Werkhoven CH, Postma DF. Inappropriate Use of Antimicrobials for Lower Respiratory Tract Infections in Elderly Patients: Patient- and Community-Related Implications and Possible Interventions. Drugs Aging. 2018;35(5):389-398.

Schweitzer VA, van Smeden M, Postma DF, Oosterheert JJ, Bonten MJM, van Werkhoven CH. Response Adjusted for Days of Antibiotic Risk (RADAR): evaluation of a novel method to compare strategies to optimize antibiotic use. Clin Microbiol Infect. 2017;23(12):980-985.

Simonetti AF, van Werkhoven CH, **Schweitzer VA**, Viasus D, Carratalà J, Postma DF, Oosterheert JJ, Bonten MJM. Predictors for individual patient antibiotic treatment effect in hospitalized community-acquired pneumonia patients. Clin Microbiol Infect. 2017;23(10):774.e1-774.e7.

Schweitzer VA, van Dam AP, Hananta IP, Schuurman R, Kusters JG, Rentenaar RJ.. Identification of Neisseria gonorrhoeae by the Bruker Biotyper Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry System Is Improved by a Database Extension. J Clin Microbiol. 2016;54(4):1130-2.

Marbus SD, **Schweitzer VA**, Groeneveld GH, Oosterheert JJ, Schneeberger PM, van der Hoek W, van Dissel JT, van Gageldonk-Lafeber AB, Mangen MJ. Incidence and costs of hospitalized adult influenza patients in The Netherlands: a retrospective observational study. Eur J Health Econ. 2020;

van Heijl I, **Schweitzer VA**, van der Linden PD, Bonten MJM, van Werkhoven CH. Impact of antimicrobial de-escalation on mortality: a literature review of study methodology and recommendations for observational studies. Expert Rev Anti Infect Ther. 2020;

DANKWOORD

Het is eindelijk zover, het proefschrift is af. Dit was nooit gelukt zonder de bijdrage van veel mensen die ik graag in het bijzonder wil bedanken.

Beste **Marc**, bedankt voor jouw begeleiding de afgelopen jaren. In mijn ogen ben jij een echte en ontzettend goede wetenschapper. Ondanks jouw altijd drukke agenda, lukte het je altijd om tijd vrij te maken om te brainstormen voor een nieuwe blog of om een nieuw ingewikkeld idee of theorie te bespreken. Een van jouw eigenschappen die ik tijdens mijn promotie ontzettend kon waarderen is jouw vermogen om problemen snel te doorzien en om vervolgens knopen door te hakken als iedereen twijfelde over hoe verder te gaan.

Beste **Jan Jelrik**, de eerste keer dat we elkaar spraken was in 2010 tijdens mijn sollicitatiegesprek voor SUMMA. Destijds hadden we meteen een goede klik, want ons gesprek ging alleen over reizen en leuke dingen, terwijl andere sollicitanten aan een kruisverhoor werden onderworpen. De wekelijkse besprekingen op de woensdag gaven me altijd veel energie, waren laagdrempelig en gemoedelijk en de wetenschappelijke discussies waardenvol. Dank je wel voor je begeleiding.

Beste **Edwin**, ondanks dat de discussies over nieuwe ingewikkelde statistische technieken soms misschien een beetje te ver gingen, heb ik veel aan jouw begeleiding gehad. Als het even tegenzat met de studie kon ik altijd bij je terecht en was je bereid meteen de telefoon te pakken om dingen op te lossen. Tijdens de momenten op de woensdagmiddag waarbij we bij een kopje espresso de voortgang bespraken was je altijd in staat om zaken te relativeren en in perspectief te plaatsen, dank daarvoor.

Beste **Henri**, officieel ben jij niet mijn copromotor geweest maar onofficieel was je voor mij toch een van de beste copromotoren die een promovendus zich kan wensen. Jij bent een van de slimste en meest integere mensen die ik ken. Zelfs als je met 20 projecten tegelijk aan het jongleren was, had jij altijd tijd om laagdrempelig dingen te bespreken als dat nodig was. Ik heb erg veel van jou geleerd, dank je wel.

Beste **Inger**, ik ben heel blij dat we tijdens onze promotietrajecten veel projecten samen hebben kunnen doen. Niet alleen omdat het praktisch gezien bijna onmogelijk is om zo'n grote studie alleen uit te voeren, maar ook omdat we als lotgenoten makkelijk al onze 'promotie-frustraties' met elkaar konden bespreken. We hebben er gelukkig ook genoeg om kunnen lachen. Vooral de bijpraat momenten in de auto, als we weer eens voor de 100ste keer om half 8 s'ochtends een klinische les moesten geven aan de andere kant van het land, waren voor mij heel waardevol.

Beste **Sebas**, bedankt dat jij mijn vriend bent en mijn paranimf wilt zijn. Ons bijna-wekelijkse lunchmomentje was altijd een goed moment om alles weer even in perspectief te plaatsen in het soms absurde leven van een promovendus. Wij hebben dezelfde humor, dus als we tijdens de lunch na het kijken van een of ander filmpje de tranen over de wangen hadden lopen van het lachen, was dit weer een goede opfrisser om er daarna tegenaan te gaan. Ik hoop dat, ondanks dat je nu voor je opleiding kindergeneeskunde in Rotterdam woont, we elkaar nog vaak zullen zien.

Geachte leden van de leescommissie, prof. dr. **Jan Kluytmans**, prof. dr. **Pim de Jong**, prof. dr. **Theo Verheij**, prof. dr. **Marlies Hulscher**, prof. dr. **Harry Heijerman**, hartelijk dank voor uw bereidheid zitting te nemen in de beoordelingscommissie van dit proefschrift.

Beste **CAP-PACT onderzoekers** en **betrokken patiënten**, heel veel dank voor jullie bijdrage aan de onderzoeken.

Lieve collegae van de **epi infectieziekten groep, WMM'ers, en XEWMM'ers**. Elke woensdagmorgen stonden jullie garant voor de nodige dosis aan humor, scherpe inzichten en inhoudelijk goede discussies. Eén van de grootste voordelen van promoveren bij deze groep zijn jullie. Gezamenlijk wordt bijna al het onderzoek naar een hoger niveau getild. Daarnaast waren de borrels en ECCMID ook jaarlijks weer een hoogtepunt.

Lieve **AIOS Medische Microbiologie**, wat een leuke groep hebben wij. Ik ben ontzettend blij dat ik hier in opleiding ben gekomen. Vanaf dag één heb ik me bij jullie meteen welkom gevoeld.

Lieve collegae van de Jan Jelrik onderzoeksgroep, **Jan Jelrik, Sonja, Marieke, Laura, Bianca, Sonja, Henri, Lufang, Jan Willem en Thijs**. De combinatie van inzichten uit de klinische praktijk en epi- en statistiek kennis, maakt deze bespreking ontzettend waardevol. Daarnaast was er altijd een laagdrempelige sfeer en de ruimte voor humor, dank jullie wel.

Lieve kamergenoten van 5.122, **Darren, Stephanie, Loes, Carmen, Lufang, Marieke, Carla, Anne en Eveline** en later van 5.14, **Tim en Roland**. Dank voor de goede tijden.

Dear **Martin, Sarah, and Jasmin**. It was an honour to have worked with you on the JPIAMR project. Without your dedication and help this project, which is now an important part of my PhD thesis, would never have been successful. I learned a lot from you and I hope to have the opportunity to collaborate with you again in the future.

Beste **ADO'20 veteranen** en specifiek de **PS4 groep**. Dank voor de nodige afleiding in de afgelopen jaren. Zonder jullie was dit proefschrift waarschijnlijk vorig jaar al af geweest.

Lieve **pap** en **mam**, dank voor jullie onvoorwaardelijke steun in alles wat ik doe. Het had jullie helemaal niets uitgemaakt of ik nou brandweerman, schilder of dokter had willen worden. Door jullie steun en complete vrijheid in keuze heb ik alle kansen gehad om te doen wat ik wilde. Ik kan altijd op jullie rekenen, dank jullie wel. Lieve **Tho**, **Amber**, **Saar**, **Bram**, **Joop**, **Ilse**, **Sjuul**, dank voor jullie eeuwige steun en oprechte interesse.

Lieve **Lisa**, zonder jou was me dit zeker niet gelukt. Jij liet me op de juiste momenten zien dat er belangrijkere dingen zijn in het leven dan werk. Dank voor jouw geduld en steun, vooral op de momenten als ik soms 's avonds na het werk afwezig kon zijn, omdat ik in mijn hoofd weer met 1000 problemen tegelijk bezig was. Ik hoop dat we nog veel mooie dingen samen mee mogen maken. I love you.

CURRICULUM VITAE

Valentijn Schweitzer was born on the 27th of June 1987 in Beverwijk, the Netherlands. He followed secondary education at the Jac. P. Thijsse college in Castricum, from which he graduated in 2005. After that, he completed a bachelor's degree in Biomedical Sciences at the University of Amsterdam in 2009. After his bachelor in Biomedical Sciences, he was selected for the Selective Utrecht Medical Master (SUMMA) at the University of Utrecht and graduated in 2014 as a Medical Doctor. During his bachelor in Biomedical Sciences and medical training he developed an interest in Infectious Diseases and Medical Microbiology.

After completing his Medical degree in 2014, he started working as PhD-candidate in the Epidemiology of Infectious Diseases group of prof. dr. Marc Bonten in Utrecht. He worked on research methodology in antimicrobial stewardship and on improving antibiotic treatment for community-acquired pneumonia patients under the supervision of prof. dr. Marc Bonten, dr. Jan Jelrik Oosterheert, and dr. Edwin Boel. Next to his PhD work, he completed a postgraduate master in Epidemiology, with a specialisation of Infectious Disease Epidemiology, at the University of Utrecht.

In October 2018, he started his residency training in becoming a medical specialist in Medical Microbiology at the University Medical Centre Utrecht and the St. Antonius Hospital.