OHE

Proposals for a Novel UK Antimicrobial Subscription Model: How Will Antibiotic Innovation be Scored?

CONTRACT RESEARCH REPORT
NOVEMBER 2023

Matthias Hofer
Grace Hampson

ohe.org

OHE

NOVEMBER 2023

Proposals for a Novel UK Antimicrobial Subscription Model: How Will Antibiotic Innovation be Scored?

Matthias Hofer
Office of Health Economics, London

Grace Hampson
Office of Health Economics, London

OFFICE OF HEALTH ECONOMICS
CONTRACT RESEARCH

Please cite this report as:
Hofer MP., Hampson G., 2023. Proposals for a Novel UK Antimicrobial Subscription Model: How will antibiotic innovation be scored? OHE Contract Research Report, London: Office of Health Economics. Available at: www.ohe.org/publications/novel-uk-antimicrobial-subscription-model-how-will-innovation-be-scored/

Corresponding Author:
Grace Hampson
ghampson@ohe.org

For further information please contact:

Professor Graham Cookson
Chief Executive, OHE
Honorary Visiting Professor in Economics at City, University of London

Tel +44 (0)207 747 1408
Email gcookson@ohe.org

About OHE Contract Research Reports

Many of the studies OHE performs are proprietary and the results are not released publicly. Studies of interest to a wide audience, however, may be made available, in whole or in part, with the client's permission. They may be published by OHE alone, jointly with the client, or externally in scholarly publications. Publication is at the client's discretion.

Studies published by OHE as OHE Contract Research Reports are subject to internal quality assurance and undergo external review, usually by a member of OHE's Editorial Panel. Any views expressed are those of the authors and do not necessarily reflect the views of OHE as an organisation.

Funding and Acknowledgements

This consulting report was commissioned and funded by The Association of the British Pharmaceutical Industry (ABPI).

Table of Contents

Executive Summary

INTRODUCTION

NHS England held a public consultation on proposals for a subscription style revenue guarantee to stimulate research and development (R&D) in antimicrobials from July to October 2023. Under the scheme, eligible antimicrobials will be assessed against a multicriteria scoring system. If they achieve a sufficient score, they will be placed into one of four value bands and receive a corresponding annual payment to serve as a revenue guarantee.

The purpose of this research was to test how new antibiotics are likely to be valued under the new scoring system that has been proposed by NHS England. To do this, five realistic dummy antibiotics were created based on the current pipeline and scored against the proposed system. In doing this we sought to evaluate:

- how products are likely to score,
- how feasible it may be for products to reach each proposed value band,
- whether expert clinicians consider the scoring system, including criteria, levels, weights and scores to be appropriate.

METHODS

Dummy products were created based on assumptions regarding what could be achievable in terms of product development and evidence generation in the short, medium, and long-term future. Dummies were validated and scored via four steps to ensure that the products were realistic and could potentially represent the key characteristics of a large proportion of products in development.

1. **WHO pipeline extraction:** we used the most recent publication by the World Health Organization (WHO) (World Health Organization, 2022) extracting type of product (non-traditional and novel antimicrobials, innovative traditional antibiotics, or traditional antibiotics), method of administration, target pathogen(s), and clinical trials in resistant pathogens.
2. **Creation of preliminary product profiles:** We created five product profiles based on the key characteristics extracted from the WHO pipeline. Most profiles target gram-negative bacteria and their resistance mechanism, with the exception of dummy profile 3.
    - **Dummy 1** represents a highly innovative non-traditional antimicrobial (phage enzyme or antibody) that is administered intravenously (IV). It targets most WHO priority pathogens.
    - **Dummy 2** is an innovative traditional antibiotic that has both oral and IV formulations, targeting most priority pathogens.
    - **Dummy 3** concerns an innovative traditional antibiotic that is administered orally. This dummy targets gram-positive pathogens and resistance mechanisms, and therefore targets fewer priority pathogens.
    - **Dummy 4** is an oral product with no specific innovative features. This dummy targets many priority pathogens.
    - **Dummy 5** requires IV administration. As with dummy 4, it has no specific innovative features and targets many priority pathogens.

3. **Expert input to validate, refine and score preliminary product profiles**: we scored and validated the dummy products with the help of five leading clinical experts: four clinical microbiologists and one hospital pharmacist.
4. **Sensitivity analysis:** we varied the input across the full range of levels to test our assumptions.

Note that the scoring was based on assumptions around what might feasibly be achieved, rather than development of a mock evidence package. The scoring was also not developed with particular products or case studies in mind.

RESULTS

Dummy 2 scored the highest, narrowly missing out on the top value band. Dummies 4 and 5 received the lowest scores, reaching value band 4 and no value band respectively. The most novel product, dummy 1, achieved value band 2. The innovative antibiotic (dummy 3) targeting gram positive pathogens achieved value band 3.

COMMENTS ON THE SCORING SYSTEM

Most of the experts commented that they would have included different criteria and levels, for example, one suggested the inclusion of a criterion on microbiological cure (i.e. eradicating the pathogen). It is unsurprising that different experts would suggest different combinations, so divergence of opinion is not necessarily a cause for concern. What is important is that the development of the scoring system has captured a range of views, with the aim of reflecting clinical consensus as far as possible. However, the consultation documents are not clear as to how the criteria or levels were designed and selected.

NHS England provided an analysis of how the proposed scoring system aligns with the STEDI framework. We conducted a complementary analysis, in collaboration with the clinical experts, that determines if the scoring system considers direct evidence, indirect evidence, or no evidence to support STEDI values (Table 6). Overall, we found that the categories and criteria do broadly align with the STEDI framework. However, there is a gap in relation to transmission value, for which the scoring system only considers indirect evidence.

At least one expert felt there was too much focus on UK needs in terms of selection and weighing of WHO resistance pathogens and resistance mechanisms. The proposed scoring system favours products targeting gram-negative bacteria and carbapenem resistance, where the highest scores can be achieved. This reflects UK needs, at the expense of reflecting the wider global landscape of evolving international resistance. The latter must be considered if we are to effectively tackle this global problem.

Criterion 1D was particularly controversial amongst the experts. This criterion requires RCT evidence in resistant pathogens to achieve any score greater than 50. Current international regulatory standards for antibiotics do not require such evidence, and experts confirmed that this type of data is not required for clinical decision making, and thus this evidence would have no use outside of the proposed framework. The experts were concerned that the levels and scores as written will simply penalise all antibiotics, particularly in the short term. The impact of this criterion is substantial due to weight it carries (11.3% of the total score).

KEY TAKEAWAYS

Based on the results of the analysis and the discussions with clinical experts, we propose four key takeaways:

1. Realistic dummy products can be constructed from the pipeline that reach value bands 2-4, representing annual payments of £5-15m.

2. The highest scoring dummy product (dummy 2) achieved a score of 79.2, just missing the top value band (minimum score 80). It is an innovative antibiotic, targeting gram-negative pathogens and resistance mechanisms, with both oral and IV formulation.

3. Not all products in the pipeline will achieve sufficient scores to quality for a value band. Less innovative antibiotics, which are most likely to reach market in the short to medium term, do not score well.

4. It will be particularly challenging for products to achieve a good score on criterion 1D, related to evidence of clinical effectiveness. This criterion requires RCTs in resistant pathogens to achieve a score greater than 50. Given that such trials are almost impossible to conduct due to ethical, statistic, practical and commercial concerns, the demands of this criterion are not reasonable.

1. Introduction

Antibiotics, a sub-set of antimicrobials, are essential for modern medicine. Antibiotic resistance (AMR) occurs when bacteria become resistant to existing antibiotics, reducing the effectiveness of these antibiotics in treating infections. AMR is a serious issue worldwide, causing an estimated 1.27 million premature deaths in 2019 (Antimicrobial Resistance Collaborators, 2022). In Europe, 670,000 infections and 33,000 deaths annually are attributed to resistant bacteria (European Centre for Disease Prevention and Control and WHO Regional Office for Europe, 2022).

There is an urgent need for new antibiotics to address increasing AMR, but there is currently only a limited number of promising new antibiotics in late-stage development. The pipeline is underdeveloped in terms of both product originality and target pathogen(s) (Wellcome Trust, 2020; Butler et al., 2022).

This extent of underdevelopment is largely because the market for antibiotics is not commercially viable. Treatment durations are short, giving manufacturers limited opportunity to recoup development costs, and new products face fierce competition from existing, cheap, generic antibiotics that still work well for many patients. There is persistent market failure caused by (appropriate and necessary) restrictions on use in line with antimicrobial stewardship practises and a lack of adequate health technology assessment (HTA) methodologies to capture the full value of antibiotics (Prasad et al., 2022; Leonard et al., 2023). As a result, a number of international pharmaceutical companies have left the antibiotics market, while smaller players have filed for bankruptcy (Årdal et al., 2020).

To tackle the market failure, a combination of push and pull incentives have been advocated as policy options. Push incentives typically come via philanthropic or state funds that offer up-front money for research and development (R&D). Pull incentives, on the other hand, serve as rewards for goods when they first enter the market (Dutescu and Hillier, 2021). For antibiotics, a "volume-delinked" pull incentive has been suggested that could take the form of a subscription, allowing developers to be reimbursed independent of the volume of sales (Rex and Outterson, 2016).

In the UK, the NICE-NHS England AMR Pilot (NICE, 2023) tested the application of this type of pull incentive for two antibiotics. The funding arrangements under the pilot commenced in July 2022 and will last for three years, with the option of extending up to 10 years (Leonard et al., 2023). More recently, NHS England initiated a public consultation on new proposals for a more long-term follow-up scheme (NHS England, 2023). Under the proposed scheme, eligible antibiotics will be assessed against a multicriteria scoring system. If they achieve a sufficient score, they will be placed into one of four value bands and receive a corresponding annual payment to serve as a revenue guarantee. Further details are provided in chapter 2.

1.1. About this report

The purpose of this report is to explore how new antibiotics are likely to be valued under the new scoring system proposed by NHS England. To this end, five realistic dummy antibiotics were created based on the current pipeline and scored against the proposed system. In doing this we sought to evaluate:

- how products are likely to score,

- how feasible it may be for products to reach each proposed value band,
- whether expert clinicians consider the scoring system (including the proposed criteria, levels, weights and scores) to be appropriate.

This report gives an overview of the scoring system as outlined in NHS England's proposals, describes the process and methods used to create, validate and score the dummy products, outlines the results of the dummy product analysis, and summarises feedback on the proposals from the expert clinicians.

2. Overview of the UK proposals

2.1. Eligibility

The proposed process is split into two steps. In a first step, eligibility for the scheme will be assessed in an "administrative" procedure according to set criteria, with the eligibility criteria reviewed every 12 months. The eligibility criteria are detailed in Table 1.

Products that do not meet the eligibility criteria for a subscription style contract will still be able to access the NHS market via the standard route, i.e. Medicines and Healthcare products Regulatory Agency (MHRA) approval and subsequent listing through the Department for Health and Social Care (DHSC) in the same way as other medicines.

The proposal explains that the window for considering new products is likely to be open periodically (e.g. once a year).

TABLE 1: ELIGIBILITY CRITERIA FOR UK AMR SUBSCRIPTION MODEL PROPOSAL

Target pathogens	Products must be active against pathogens on the WHO Priority List
Agreement with contract terms on surety of supply, antimicrobial stewardship, and performance	Agreement to surety of supply, antimicrobial stewardship (including sales and promotional activities), key performance measures and payment terms
Environmental Standards	Companies must demonstrate compliance with specified antibiotic manufacturing standards
Economic and Financial Standing	Companies must demonstrate they have a sufficient economic and financial standing to justify award of the proposed contract
Probity	Companies must demonstrate they do not trigger any of the requirements that would make them ineligible to be awarded a public contract
Social Value	Companies must demonstrate their commitment to specified social value requirement, e.g. achieving net zero emissions

What is not clear from the information provided is whether products categorised as non-traditional and novel antimicrobials by the World Health Organization (WHO) will be eligible for the scheme. These products include antibodies, bacteriophages and phage-derived enzymes, immunomodulating agents, microbiome-modulating agents, and as such are not typical antibiotics. Clarification on whether these are eligible will be crucial for manufacturers and developers.

This report does not explore the eligibility criteria further. For a discussion on how the eligibility criteria may impact the pull incentive, see Hofer and Hampson (2023).

2.2. Scoring

In a second step, it is proposed that eligible antimicrobials will be assessed against 17 criteria, split into 3 categories: i) relative effectiveness and unmet clinical need, ii) pharmacological benefit, iii) health system benefit. Each criterion is broken down into a number of levels. Each level has an associated score, which will be multiplied by a weight allocated to the relevant criterion. The weighted score from each of the 17 criteria is then summed to generate the total score, between 0 and 100[1].

The consultation documents explain that the criteria have been developed to reflect broad values based on the experience from the UK pilot and following the STEDI framework, which emphasises antimicrobial values in the context of spectrum, transmission, enablement, diversity, and insurance (see section 4.1) (NHS England, 2023).

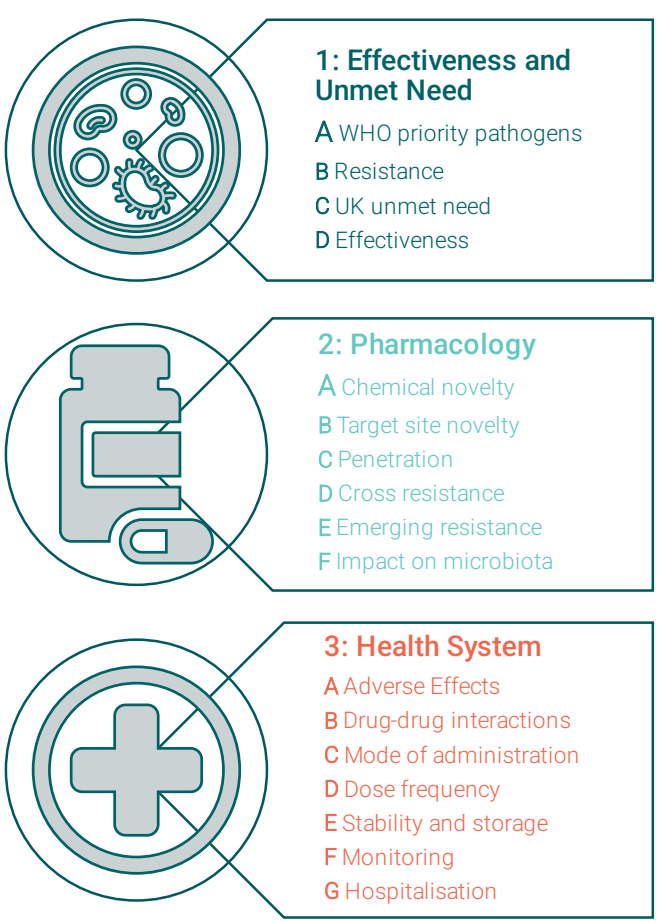
The categories, criteria, weights and levels are presented in Table 2. The relative weights of the criteria are presented visually in Figure 1. Full details of the levels and their scores can be found in the consultation documentation (NHS England, 2023).

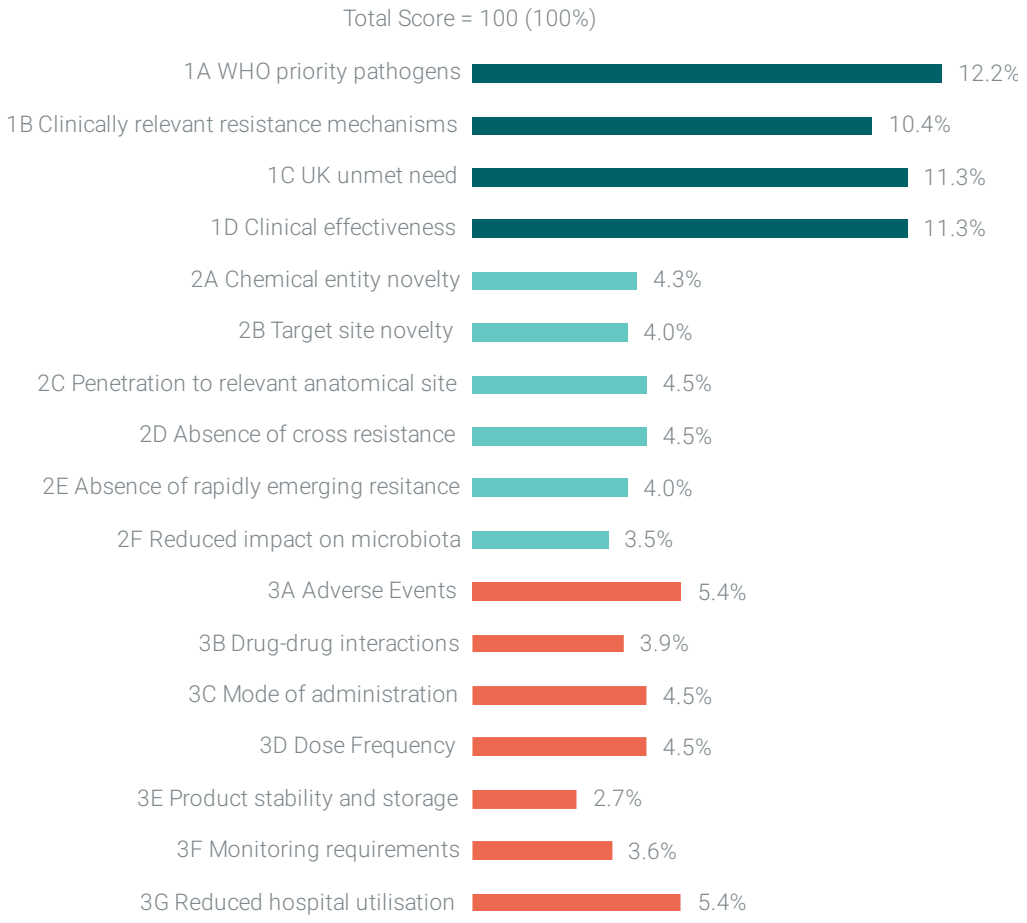
[1] *Note that in the proposals there is an error in the weighting of the pharmacological benefit category so the weights do not sum to 1. The maximum possible score is therefore currently 99.75.*

TABLE 2: PROPOSED CRITERIA WITH WEIGHTING AND LEVELS (NHS ENGLAND, 2023)

Category	Criteria		Category weight	Criterion weight	Number of levels
Relative Effectiveness and Unmet Clinical Need	1A	Activity against WHO priority pathogens	0.45	0.27	10
	1B	Activity against clinically relevant resistance mechanisms		0.23	9
	1C	Activity against UK unmet needs		0.25	3
	1D	Clinical effectiveness compared to best standard of care		0.25	6
Pharmacological Benefit	2A	Chemical entity novelty	0.25	0.17	4
	2B	Target site novelty		0.16	3
	2C	Penetration to relevant anatomical site		0.18	3
	2D	Absence of cross resistance		0.18	4
	2E	Absence of rapidly emerging resistance		0.16	4
	2F	Reduced impact on microbiota		0.14	3
Health System Benefit	3A	Adverse Events	0.3	0.18	4
	3B	Drug-drug interactions		0.13	4
	3C	Mode of administration		0.15	6
	3D	Dose Frequency		0.15	5
	3E	Product stability and storage		0.09	3
	3F	Monitoring requirements		0.12	3
	3G	Reduced hospital admissions or length of stay		0.18	4

FIGURE 1: CRITERIA AND WEIGHTS (ADAPTED FROM NHS ENGLAND, 2023)





2.3. Value bands

Assuming a minimum score (50 points) is achieved, the new antimicrobial will fall into one of four bands. Each band corresponds to a payment value. For England, the values are: £5 million, £10 million, £15 million, or £20 million per year respectively (Figure 2). The values are expected to be subject to review each year. NHS England and the devolved governments, at their discretion, will offer a contract in the relevant value band. If the minimum score is not achieved, the new antimicrobial will not qualify for a subscription incentive contract.

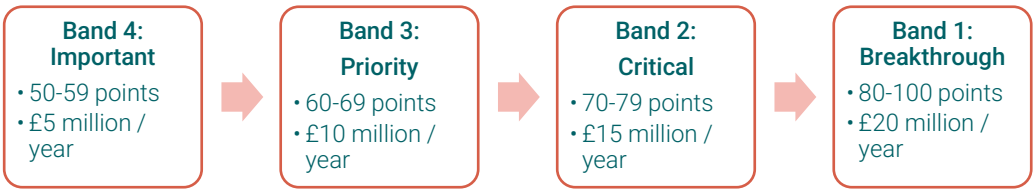


FIGURE 2: PROPOSED MONETARY VALUE BANDS FOR ENGLAND

2.4. Evidence requirements

The consultation by NICE and NHS England sets out four distinct types of evidence that are allowed to be submitted. All other data types (e.g. modelling approaches or health economic assessment) are not allowed. Permitted evidence includes:

- Evidence from the UK Summary of Product Characteristics (SmPC). A draft SmPC will be accepted if assessment is performed pre-approval. International marketing authorisation documentation may also be taken into account.
- Clinical evidence and systematic reviews of clinical evidence according to NICE Decision Support Unit guidance (NICE Decision Support Unit, 2022). This encompasses data from clinical trials, registry data analyses, and case series studies. For certain criteria, a preference for UK specific data is expressed, for example for criteria 1B, 1C, and 3G.
- Evidence from *in vitro* studies. This is generally data that has been generated by the applicant to support product development and assessment. As an exception, for the assessment of cross resistance criterion (2D), where *in vitro* evidence has to be generated externally by UK Health Security Agency (UKHSA).
- Evidence from pharmacokinetic and pharmacodynamic studies that are often part of the evidence package generated for regulatory approval.

Figure 3 shows the proportion of the criteria that require each type of evidence.

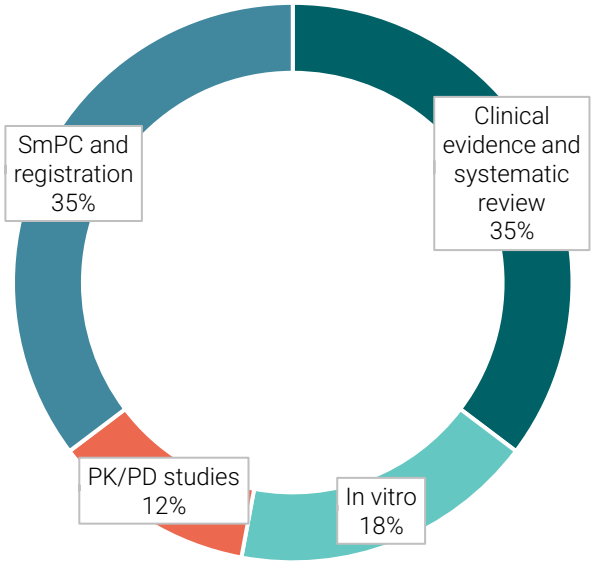


FIGURE 3: PROPORTION OF THE CRITERIA REQUIRING EACH TYPE OF EVIDENCE

Notes: a small number of criteria allow for multiple evidence types. For the purpose of this figure 1A is recorded as requiring in vitro evidence; 2E, 2F, and 3D as requiring clinical evidence and systematic review.

2.5. Contracting

Initially NHS England is expected to act as the lead authority running the procurement process and issuing the invitations to tender. The devolved administrations NHS Scotland, NHS Wales and Northern Ireland Department for Health will be responsible for the decision to issue a contract for their nation and for the payments within their respective jurisdictions.

The proposed subscription contract is for an initial three-year period and is extendable to cover the patent exclusivity period of the product (to a maximum of 15 years). The proposed contract includes key performance indicators on surety of supply.

The product could be moved between value bands over time if the value of the product to the NHS changes, and the NHS reserve the right to cancel the contract if it is determined that the product no longer warrants a subscription contract. For discussion around how the proposed contractual requirements will impact the pull incentive, see Hofer and Hampson (2023).

3. Dummy product analysis

3.1. Methods

Dummy products were created, validated and scored via four steps designed to ensure that the products were realistic and represented a large proportion of products in development.

1. **WHO pipeline extraction:** we used the most recent pipeline analysis from the WHO which includes a total of 80 antibacterial products in clinical development for priority pathogens (World Health Organization, 2022). We extracted the key characteristics that are listed for these 80 products including type of product (traditional/non-traditional; and novelty of traditional), method of administration, target pathogen(s), and clinical trials in resistant pathogens. In relation to type of product, the WHO categorise the pipeline according to three different product types:

    i. non-traditional and novel antimicrobials (i.e. antibodies, bacteriophages and phage-derived enzymes, immunomodulating agents, microbiome-modulating agents, and others)

    ii. innovative (traditional) antibiotics (with new chemical class, or new target, or new mode of action; and/or no cross resistance to other antibiotic classes)

    iii. novel (traditional) antibiotics.

    The key characteristics of these three product groups were collected and used in step 2.

2. **Creation of preliminary product profiles:** we created five product profiles that synthesise the core characteristics of the three WHO product types. Most profiles target gram-negative bacteria and their resistance mechanism, except for dummy profile 3.

    - **Dummy 1** represents a highly innovative non-traditional antimicrobial (phage enzyme or antibody) that is administered intravenously (IV).
    - **Dummy 2** is an innovative traditional antibiotic that has both oral and IV formulations.
    - **Dummy 3** concerns an innovative traditional antibiotic that is administered orally and targets gram-positive pathogens and resistance mechanisms.
    - **Dummy 4** is a novel oral antibiotic product with no specific innovative features.
    - **Dummy 5** is a novel antibiotic that requires IV administration, and as with dummy 4, has no specific innovative features.

    Based on type of product, novelty and type of administration, these dummy profiles collectively cover the key characteristics of approximately 66% of the WHO pipeline (Annex, A1). Table 3 provides more detail on each of these high-level profiles.

TABLE 3: OVERVIEW OF WHO PIPELINE PRODUCT CHARACTERISITCS AND RELATED DUMMY PROFILES

WHO Pipeline		Product Profiles
Non-traditional Antimicrobials	• Mostly oral and IV administration • Address more WHO priority pathogens • Some clinical developments in drug resistant pathogens • Examples include: phage endolysin, antibacterial antibodies, bacteriophages	**Dummy 1** → Targets most priority pathogens → Highly innovative → IV administration → Profile represents 11% of the WHO pipeline
Innovative Traditional Antibiotics	• Mostly oral, or oral & IV administration • Address less WHO priority pathogens (often targeting gram positive bacteria) • Some clinical developments in drug resistant pathogens • Examples include FtsZ inhibitors, FabI inhibitors, Triazaacenaphthylene, DprE1 inhibitors, Bis-benzimidazoles, Diaryldiamine	**Dummy 2** → Targets most priority pathogens → Innovative → Oral & IV admin → Profile represents 3% of the WHO pipeline **Dummy 3** → Targets gram-positive bacteria and resistance mechanisms -> less priority pathogens → Innovative → Oral administration → Profile represents 13% of the WHO pipeline
Traditional Antibiotics	• IV or oral administration • Address more WHO priority pathogens • Few clinical developments in drug resistant pathogens • Examples include oxazolidinones, diarylquinolines, combinations with β-lactam inhibitors, macrolides	**Dummy 4** → Targets many priority pathogens → Not innovative → Oral administration → Profile represents 12% of the WHO pipeline **Dummy 5** → Targets many priority pathogens → Not innovative → IV administration → Profile represents 14% of the WHO pipeline

Next, we generated preliminary scores across all 17 criteria, to provide a starting point for discussions with experts in step 3. The WHO pipeline did not provide granularity to inform target pathogen, resistance mechanism, or many of the other criteria, thus these were based on preliminary assumptions. To give an indication of the broad assumptions used to develop this starting point, we assumed that more novel products would score highly in category 2 (pharmacological benefit, with criteria such as chemical entity novelty, target site novelty, and absence of rapidly emerging resistance), whilst oral products would score highly in category 3 (health system benefit, with criteria such as mode of administration, dose frequency, and product stability and storage). All preliminary assumptions and each of the scores were subsequently discussed with and altered by the experts (see step 3).

Note that the dummies and scores were not developed with particular products or case studies in mind, but rather to reflect the breadth of the WHO pipeline and the diversity of products in development. The scores were also not constructed with any value band(s) in mind, but purely to represent realistic approximations of potential products. The scoring was based on assumptions around what could feasibly be achieved based on the products assumed characteristics, rather than development of a mock evidence package or any detailed assessment of the feasibility of the required evidence package for each score.

3. **Expert input to validate, refine and score preliminary product profiles:** we scored and validated the dummy products with the help of five leading clinical experts: four clinical microbiologists and one hospital pharmacist. The input was gathered by means of five individual virtual 'interactive exercises' in which we presented our methods and preliminary products, and asked the experts for their feedback and amendments on the selection/range of dummy profiles, the feasibility of these types of products reaching the market, the types of evidence that might be available, the scores for each product, and any additional general comments they had on the scoring system (including the criteria, weights, and levels). The process was iterative in that after each session we made amendments, with the revised set of dummies and scores presented to the next participant. Scores changed significantly from step 2 based on the feedback and expertise of the consulted experts. Table 4 provides a heatmap overview of the resulting scores for each dummy. The table reveals that even the most innovative products were not allocated top scores in many of the criteria. For example, Criterion 1D requires RCT evidence in resistant populations to achieve any score higher than 50, and such trials are not considered feasible (see section 4.1). As such, the score for this criterion was capped at 50 across all dummy products.

   Full scoring details are provided in Annex A2. Based on conversations with the experts, we suggest that the dummies could represent products which may emerge in the short (dummies 4&5), medium (dummies 2&3), and long term (dummy 1).

4. **Sensitivity analysis:** there is substantial uncertainty surrounding the potential scores that new antibiotics might achieve. As such, we conducted a number of sensitivity analyses to explore the impact of different scores on the results. Whilst uncertainty exists across all criteria, we focused sensitivity analysis on five that were considered to be key due to weighting, scale of uncertainty, and subjectivity of score:

   - **1C: Activity against UK unmet needs:** Chosen for sensitivity analysis as the scoring for this criterion appears to be subjective and includes only three levels, despite it representing a substantial proportion of the total score (>10%).
   - **1D: Clinical effectiveness:** Chosen as many of the experts felt the levels included under this criterion were wildly unrealistic and unattainable, particularly in the short term. Again this criterion holds substantial weight, representing over 10% of the total score.
   - **2E: Emerging resistance**: Included as it is challenging to infer the emerging resistance from the dummy profiles as created. Some stakeholders have expressed concern that the scores and levels will be extremely challenging to achieve.
   - **3A: Adverse effects:** Included as some stakeholders have expressed concern that the scores and levels will be extremely challenging to achieve.
   - **3G: Hospital utilisation:** Included as it was not possible to infer the impact on hospitalisation from the dummy profiles as created. As such this criterion was held constant in the base case and varied via sensitivity analysis. This criterion also holds the joint highest weight outside of category 1 (5.4%, see Figure 1).

All sensitivity analyses were one-way, which means they indicate what would happen to the total score when a dummy achieves a different score (relative to the base case) on one criterion in isolation. In reality, there are lots of moving parts, and thus these analyses give only a limited insight into how quickly the results can change if a product were to score differently on two or more criteria simultaneously. Results of the sensitivity analysis are presented in section 3.2 in relation to individual dummies, and in Chapter 4 in relation to what they tell us about the scoring system.

TABLE 4: HEATMAP OVERVIEW OF THE SCORING FOR THE FIVE DUMMY PRODUCTS

Scoring criteria	Dummy 1	Dummy 2	Dummy 3	Dummy 4	Dummy 5
1A Activity against WHO priority pathogens					
1B Activity against clinically relevant resistance mechanisms					
1C Activity against UK unmet needs					
1D Clinical effectiveness compared to best standard of care					
2A Chemical entity novelty					
2B Target site novelty					
2C Penetration to relevant anatomical site					
2D Absence of cross resistance					
2E Absence of rapidly emerging resistance					
2F Reduced impact on microbiota					
3A Adverse Events					
3B Drug-drug interactions					
3C Mode of administration					
3D Dose Frequency					
3E Product stability and storage					
3F Monitoring requirements					
3G Reduced hospital admissions or length of stay					

Legend: Score >90, Score 50-89, Score <50. Detailed scores are provided in the Annex.

3.2. Results

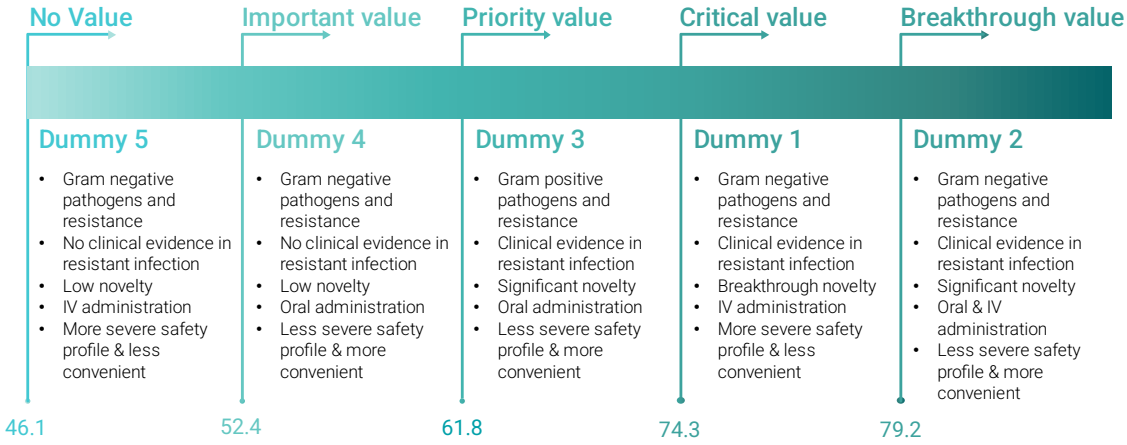
Figure 4 gives an overview of the scores. Dummy 2 scored the highest, narrowly missing out on the top value band[2]. Dummies 4 and 5 received the lowest scores, with Dummy 5 not qualifying for any value band.

Dummy 1
Score 74.3
Value band: 2 - Critical new antimicrobial
Dummy 1 targets gram-negative pathogens and resistance mechanisms and achieved high scores in category 2 due to its novelty and its more targeted approach reducing the effects on microbiota and preventing emerging resistance and cross resistance. However, the less convenient IV administration and a potentially more severe safety profile leads to lower scores in category 3. Sensitivity analyses demonstrated that dummy 1 could reach the highest value band if it fulfils a high UK unmet need

[2] *Value band 2: 'critical value' is defined as 70-79, so technically this dummy surpasses value band 2, but does not reach value band 1 which requires a minimum score of 80.*

(Table 5). Note that it is yet to be clarified that this highly innovative non-traditional type of antimicrobial is eligible for the scheme.

FIGURE 4: OVERVIEW OF SCORING AND VALUE OF DUMMY PRODUCTS



Dummy 2
Score: 79.2
Value band: 2 - Critical new antimicrobial
Dummy 2 also targets gram-negative pathogens and resistance mechanisms and achieved high scores in category 2 due to its novelty, with a better chance of preventing the emergence of resistance and cross resistance and impact on microbiota than less novel products. In contrast to dummy product 1, the oral and IV formulations provide flexibility and convenience leading to good scores in category 3. Sensitivity analyses reveal that this product could reach value band 1if it was supported by RCT evidence in resistant populations for criterion 1D, if it fulfils a high UK unmet need (criterion 1C), if it demonstrates no emerging resistance (criterion 2E), or if it reduces hospital utilisation or averts hospital admission compared to best standard care (criterion 3D) (Table 5).

Dummy 3
Score: 61.8
Value band: 3 – Priority new antimicrobial
Dummy 3 targets gram-positive pathogens and resistance mechanisms and hence achieves significantly lower scores in category 1 compared to all other dummies. It does however achieve medium to high scores in category 2 due to its novelty, with a better chance of preventing the emergence of resistance and cross resistance and impact on microbiota than less novel products. The oral formulation and convenience also lead to good scores in category 3. The substantially reduced score compared to equally innovative dummy 2 is largely driven by the lack of scores from criterion 1A and 1B, despite generally good scoring in relation to its novelty and type of administration. A product targeting gram-positive bacteria and resistance mechanisms without novel mechanism of action would receive even lower scores. Sensitivity analyses revealed that the product cannot reach value band 2 by varying, *ceteris paribus*, criteria 1D, 1C, 2E, 3A or 3G. If this product were to fulfil a high UK unmet need and achieve a higher score on the evidence of clinical effectiveness criterions (1D, minimum score 70), it would reach value band 2. This product would drop to value band 4 if it scores any lower on clinical effectiveness (1D), if it only meets a low unmet

need (criterion 1C), or if it cannot demonstrate equivalent hospital length of stay with current standard care (3G) (Table 5).

Dummy 4
Score: 52.4
Value band: 4 – Important new antimicrobial
Dummy 4 targets gram-negative pathogens and resistance mechanisms and achieved significantly lower scores in category 2 compared to dummies 1-3 due to a lack of novelty. This increases the chance for effects on microbiota and the emergence of resistance and cross-resistance. The oral formulation and convenience lead to good scores in category 3. Sensitivity analyses reveals that the product could reach value band 3 if it was to achieve a score of 70 for criterion 1D (see annex and Table 5). Presently this would require RCT evidence of non-inferiority in resistant pathogens, plus non-randomised evidence of effectiveness in resistant pathogens. This product would drop below the value bands if it scores any lower on UK unmet need (criterion 1C), emerging resistance (2E), adverse effects (3A), or if it cannot demonstrate equivalent hospital length of stay with current standard care (3G).

Dummy 5
Score: 46.1
Value band: None
Dummy 5 targets gram negative pathogens and resistance mechanisms and, like dummy 4, achieved significantly lower scores in category 2 due to a lack of novelty. This increases the chance for effects on microbiota and the emergence of resistance and cross-resistance. The less convenient IV administration and a potentially more severe safety profile leads to lower scores in category 3. Sensitivity analyses revealed that the product could however reach value band 4 if it fulfilled a high UK unmet need (criterion 1C), or is supported by clinical evidence (criterion 1D) (Table 5).

A summary of the sensitivity analyses are presented in Table 5, with full details in Annex A3.

TABLE 5: BASE CASE SCORES AND SENSITIVITY ANALYSES OF FIVE DUMMY PRODUCTS

Sensitivity analysis		Dummy 1	Dummy 2	Dummy 3	Dummy 4	Dummy 5
Base case	Score	74.3	79.2	61.8	52.4	46.1
	Value band	2	2	3	4	x
1C: UK unmet need	Score range	69.2 - 80.5	74.2 – 85.4	56.8 – 68.0	47.3 – 58.6	41.0 – 52.2
	Value bands	3 - 1	2 – 1	4 – 3	x – 4	x - 4
1D: evidence of effectiveness	Score range	68.7 - 79.9	73.6 – 84.8	56.2 – 67.5	52.4 – 63.6	46.1 – 57.3
	Value bands	3 – 2	2 – 1	4 – 3	4 – 3	x – 4
2E: emerging resistance	Score range	70.3 – 74.3	76.2 – 80.2	58.8 – 62.8	49.4 – 53.4	43.1 – 47.1
	Value bands	2	2 - 1	4 - 3	x - 4	x
3A: adverse events	Score range	70.0 – 75.4	73.8 – 79.2	56.4 – 61.8	47.0 – 52.4	41.7 – 47.1
	Value bands	2	2	4 - 3	x - 4	x
3G: hospital utilisation	Score range	71.0 – 76.4	76.0 – 81.4	58.6 – 64.0	49.2 – 54.6	42.8 – 48.2
	Value bands	2	2 - 1	4 - 3	x - 4	x

Note: The 'score range' is the range of total scores the dummy achieves when varying the score for that criterion across the full range of levels. Full results are provided in Annex A3.

Reaching value band 1: Breakthrough antimicrobials

To reach the top value band, products need to score well across the following key drivers:

- Target pathogen(s) and resistance mechanism(s), with higher scores given for broad spectrum products targeting gram-negative pathogens
- Novelty (product and target)
- Type of evidence for effectiveness, with RCT evidence in resistant populations required to achieve any score over 50 on this criterion
- Unmet need in the UK
- Type of administration and convenience.

This is challenging, hence none of the proposed dummy products reached the highest value band in the base case analysis.

Sensitivity analyses showed that dummy product 1 and dummy product 2 have the scope to reach value band 1. To achieve a score of 80, the products would need to:

- **Dummy 1:**
    - score higher in the UK unmet need category by showing that it can address a disease area of key importance with a high population mortality or morbidity burden, or
    - score higher on some combination of criteria, for example if it was to be administered orally or via inhalation, this may lead to higher scores across much of category 3 regarding type of administration and convenience. There are a small number of products in the pipeline which could potentially fit the characteristics of this dummy and are administered orally (n=3) or via inhalation (n=4) and thus could potentially fulfil this requirement.
- **Dummy 2:** demonstrate a small improvement across one of many different criteria, including but not limited to, unmet clinical need, evidence of clinical effectiveness, or hospital utilisation.

4. The Scoring System

Throughout the dummy product analyses, we took note of where the process seemed to work well, and where we encountered challenges. Experts also provided their comments and feedback on the process and the scoring system in varying levels of detail. What follows is not intended to be a comprehensive critique of the scoring system, but rather a note of some interesting points we encountered whilst constructing, validating and scoring the dummies.

4.1. Criteria and levels

Selection of the criteria and levels

Experts commented that they would have included different criteria and levels, for example, one suggested the inclusion of a criterion on microbiological cure (i.e. eradicating the pathogen). Its unsurprising that different experts would suggest different combinations, so divergence of opinion is not necessarily a cause for concern. What is important is that the development of the scoring system has captured a range of views, with the aim of reflecting clinical consensus as far as possible.

However, the consultation documents are not clear as to how the criteria or levels were designed and selected. NHS England state that the award criteria were created based on the eligibility criteria for the previous pilot and refined in consultation with clinical experts from the NHSE Antimicrobial Resistance (AMR) Programme and the UK Government's Advisory Committee on Antimicrobial Prescribing, Resistance and Healthcare Associated Infection (APRHAI), using the STEDI values as a conceptual basis. Further information is required (regarding sample size and methods) to allow a thorough critique.

Alignment with STEDI

STEDI is a conceptual framework to help capture the full societal value of an antibiotic, based on attributes of the antibiotics. It was originally proposed by Karlsberg Schaffer et al. (Karlsberg Schaffer et al., 2017), formalised by the Policy Research Unit in Economic Methods of Evaluation in Health and Social Care Interventions (Rothery et al., 2018), and named ´STEDI´ by (Outterson and Rex, 2020). It incorporates five additional elements of value over and above what is typically captured in value assessment (Outterson and Rex, 2020; Brassel, Al Taie and Steuten, 2023)

- **Spectrum value**: The benefit associated with the use of narrow(er)-spectrum antibiotics stemming from a reduction in collateral damage on the treated individuals' microbiome and the prevention of resistance selection in untargeted bacteria.
- **Transmission value**: The indirect benefit of reduced infection rates by avoiding the onward spread of pathogens to other individuals within a population using antibiotics.
- **Enablement value**: The benefit associated with enabling or improving the outcomes of other treatments or procedures where antibiotics are also needed.
- **Diversity value**: The indirect benefit stemming from preserving the activity of existing antibiotics for longer as they will be used less if the novel antibiotic is added to the treatment options (Brassel, Al Taie and Steuten, 2023).
- **Insurance value**: The indirect value associated with having an antibiotic treatment as a last line option for a patient if all other treatments fail, and for dealing better with (or completely avoiding) major catastrophic outbreaks of AMR in the future.

NHS England provided an analysis of how the proposed scoring system aligns with STEDI. We conducted a complementary analysis, in collaboration with the clinical experts, that determines if the scoring system considers direct evidence, indirect evidence, or no evidence to support STEDI values (Table 6). Overall, we found that the categories and criteria do broadly align with the STEDI framework. However, there is a gap in relation to transmission value, for which the scoring system only considers indirect evidence (via 1D, 2C, 3G).

Transmission value could be directly measured by considering microbiological eradication as an outcome measure with a value on its own, as a compliment to that on evidence of clinical effectiveness (1D). Microbiological cure is distinct from curing the patient and was highlighted as a key missing element from the scoring system by one of the clinical experts. Currently, microbiological eradication is only partially considered in criterion 1D, which mostly focuses on clinical outcomes for patients.

Note also the scoring system attributes higher scores for antibiotics with broader spectrum as the scores in criteria 1A and 1B are cumulative. This could be considered counter-intuitive in relation to aim for 'spectrum value' which values narrow spectrum antibiotics. The clinical experts noted this but did not express concern.

Finally, the scoring system intentionally incentivises the development of a specific subset of antibiotics that are scored as higher value than other antibiotics. This is appropriate and in line with value-based pricing, but it is critical that the antibiotics that score lower are still sufficiently incentivised due to the need for a diverse portfolio of new products (diversity value). For a discussion around the appropriateness of the size of the incentives offered alongside the scoring system, see Hofer and Hampson (2023).

TABLE 6: ANALYSIS OF ALIGNMENT OF REWARD CRITERIA WITH STEDI FRAMEWORK

Scoring criteria	S	T	E	D	I
1A Activity against WHO priority pathogens	×		×		×
1B Activity against clinically relevant resistance mechanisms	×		×	×	×
1C Activity against UK unmet needs			×	×	
1D Clinical effectiveness compared to best standard of care		×	×		
2A Chemical entity novelty					
2B Target site novelty					
2C Penetration to relevant anatomical site		×			
2D Absence of cross resistance					
2E Absence of rapidly emerging resistance					
2F Reduced impact on microbiota	×				
3A Adverse Events			×	×	
3B Drug-drug interactions			×		
3C Mode of administration			×		
3D Dose Frequency			×		
3E Product stability and storage			×		
3F Monitoring requirements			×		
3G Reduced hospital admissions or length of stay		×	×		

Legend: Direct evidence, Indirect Evidence, No evidence, x NHS-E/NICE analysis

Expert comments and reflections on the criteria and levels

Category 1: Relative Effectiveness and Unmet Clinical Need

At least one expert felt there was too much focus on UK needs in criteria 1A, 1B and 1C (also 2E which is discussed in the next category). The proposed scoring system favours products targeting gram-negative bacteria and carbapenem resistance, where the highest scores can be achieved. This reflects UK needs, potentially at the expense of reflecting the wider global landscape of evolving international resistance. The latter must be considered if we are to effectively tackle this global problem.

Criterion 1D was particularly controversial amongst the experts. This criterion requires RCT evidence in resistant pathogens to achieve any score greater than 50. The top score of 100 requires RCT evidence of superiority in resistant pathogens. Current international regulatory standards for antibiotics do not generally require this evidence, and as such this evidence will not be available for many products. There is also debate around whether or not RCTs in these small populations are possible due to ethical concerns and statistical limitations (insufficient sample size), or feasible due to practical or commercial considerations. Indeed, regulatory and HTA bodies typically relax RCT requirements for such small population groups (as seen for rare pathogens in the EMA's guidance on the evaluation of medicinal products for bacterial infections) (European Medicines Agency, 2022).

Adaptive trial designs or alternative statistical criteria could potentially be employed (Dane et al., 2022; Lanini et al., 2019), but still these studies would be highly unlikely to be available at launch. It is also unclear from the wording of the levels how evidence from pre-specified sub-groups of multi-drug resistant patients in larger RCTs would be scored.

Experts were concerned that rather than incentivise the collection of high-quality evidence (as we assume is intended by the inclusion of these levels) the levels and scores as written will simply penalise all antibiotics, particularly in the short term. Clinicians also suggested that this type of data is not required for clinical decision making, and thus the data has no use outside of the proposed framework.

Sensitivity analyses revealed that if the score for 1D was increased to 60 (from 50) for dummy 2 (e.g. if the requirements for the levels within this criterion were relaxed even slightly), this dummy would reach the top value band. The impact of this criterion is so substantial due to the weight it carries (11.3% of the total score) combined with tough requirements that may prevent most products from reaching anything other than the bottom levels.

Category 2: Pharmacological Benefit

Within this category, experts noted that criteria 2E: Absence of rapidly emerging resistance should explicitly consider the wider global landscape of evolving international resistance landscape without a particular UK focus.

Further, within 2C: Penetration to relevant anatomical sites of infection, there are only three levels:

- Effective penetration to relevant anatomical sites with drug concentrations reaching at least 4-fold above the resistance breakpoint for target pathogens (score = 100)
- Penetration to relevant anatomical sites with drug concentrations reaching above the resistance breakpoint for target pathogens (score = 85)
- None of the above (score =0).

One expert suggested that a product scoring 0 on this criterion is not of value and should not even be included in the scheme. As it stands, such a product could still obtain a reasonably high score overall.

Experts also noted that differentiation between products may be hampered by substantial overlaps across some criteria in category 2 and 3. For example there are overlaps in terms of mechanism of action, type of administration and product safety (2A, 2B, 3B, 3D). Furthermore, substantial overlaps exist across the criteria on type of administration, safety, monitoring and stability/storage (3A, 3B, 3C, 3D, 3E, 3F), and UK unmet needs and hospital utilisation (1C and 3G).

Category 3: Health system benefit

As proposed, the scoring system equally applies to all types of products independent of their target populations or intended use case. Experts highlighted that context and target population is critical when choosing between antibiotics, thus any scoring system must recognise that different products will be valued differently in different contexts.

Of note, the proposed scoring system places substantial emphasis throughout category 3 on products requiring oral administration at the expense of those with intravenous administration (IV). We assume this is intended to reflect the higher opportunity costs for products that need to be administered in an outpatient or secondary care setting. Nevertheless, there is an urgent need for the UK to have access to products that target multi-resistant strains in the hospital or intensive care unit (ICU), where oral medication may not be appropriate and the safety profile, monitoring and

convenience may be of lesser significance. To appropriately incentive products that will be critical in these contexts, points should not be deducted to the same extent as those which are intended to be most useful in the community setting.

As such, it may be more appropriate to tailor the criteria, weights, and/or levels to the circumstance(s) in which the product is most likely to be of high value. This, however, must be balanced against investor feedback (Hofer and Hampson, 2023) that the scoring system already appears complex. Complexity can be a deterrent for investment, thereby undermining the pull incentive. If the system is to be adapted to take context into account, the rules by which this will be determined should be made explicit and as clear as possible.

4.2. Weights and scores

The full set of weights as proposed are given in Figure 1.

The consultation documents provide some limited detail on how the weights and the scores for each level were determined. A swing weighting study was conducted with clinical experts from the NHSE AMR Programme and the APRHAI Advisory Committee, in which experts identify the most important criterion and stipulate how relatively important each of the remaining criteria are compared to the top criterion.

It is well-recognised that swing weighting and other rating-type tasks can allow respondents to avoid difficult choices or trade-offs, by assigning equal or similar weights that reflect a respondent's discomfort with trade-offs between attributes rather than genuine equality in terms of importance (i.e. equalizing bias) (Tervonen et al., 2017; Rezaei, Arab and Mehregan, 2022). Indeed, many (10) of the criteria included here present with similar scores, and as such it is unclear whether this is a valid finding. This said, methods which do force participants to make trade-offs (such as discrete choice experiments) would not be appropriate here due to the number of variables. Swing weighting may be the best available method, despite the known limitations.

The experts also highlighted the following concerns with the weighting and scores as proposed:

- Category 1 has a substantially greater weight than the others, which may not be appropriate. This will prevent the system from differentiating between products as many gram-negative organisms are likely to score similarly (this is also related to the small differences in points between the first three categories under 1A (carbapenem resistance)).
- There is a high weight and focus on gram-negative bacteria and their resistance mechanisms.
- There is little weight on emergence of resistance (only 4% of total score).

Further, Figure 1 shows that the weights for 1A-1D are all individually over 10%. This is a particularly relevant consideration for:

- Criterion 1C because there are only three levels (scores 100, 45, or 0). The analysis of dummy products assumed all five dummy products would score level 2: moderate unmet need in the base case. Given the substantial weight given to this criterion and that there are only three levels, all five products shift up or down a value band with an increase or decrease of just one (subjective) level.
- Criterion 1D because the experts consider the levels to be unrealistic (see section 4.1) and thus the higher scores to be unobtainable. Given the substantial weight given to this criterion, capping the score at 50 (which we understand may reflect reality) leads to a substantial dent in any product's ability to reach the higher value bands.

Importantly, only one of the experts explicitly stated that these weights (for 1C and 1D) were too high. For the others, the key point related to the high weights was the importance of getting the scores and levels within these criteria correct given their substantial impact on the total score.

More information on the methodology used to develop the weights and scores (including sample size and what the tasks and workshops entailed) to allow readers to thoroughly critique the methods used, or a validation exercise of the weights and scores, is crucial to determine whether the weights and scores (and therefore the scoring system as a whole) are valid.

4.3. Evidence requirements

We studied the evidence requirements across the criteria and find that only six of the 17 criteria (35%; 2A, 3A,3B,3C, 3E, 3F) are substantiated with data routinely collected as part of the regulatory package. When applying the weighting to the criteria, this data is expected to substantiate only 24% of the total score. Four criteria will certainly require new evidence outside of the regulatory package (24%; 1C, 2D, 2F, 3G). Seven criteria might be covered by the regulatory package (41%; 1A, 1B, 1D, 2B, 2C, 2E, 3D), depending on the type of product and the information that was collected through in vitro, PK/PD, or clinical studies (Figure 5). The implication is that the scheme requires a high degree of evidence that is not routinely collected for regulatory purposes.

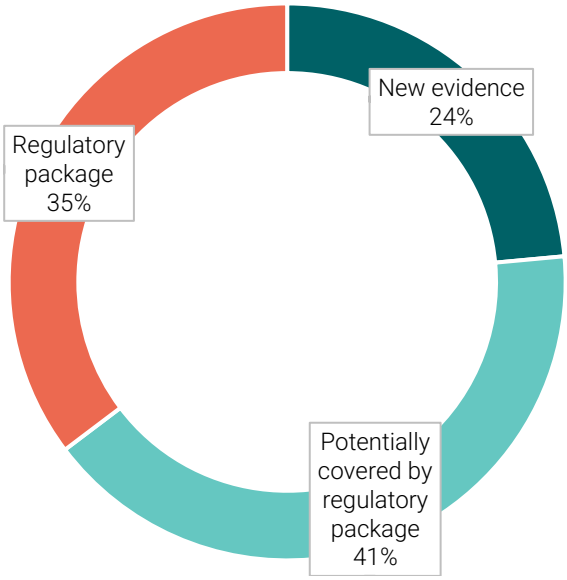


FIGURE 5: BURDEN OF EVIDENCE FOR APPLICANTS

4.4. Application of the scoring system

No information is provided on how the system could be applied to new indications. How new indications will be treated will be key, as further development of existing products in new indications can be a relatively efficient way of developing products, but without clear guidance the incentives for companies to invest in additional trials and seek license extensions may be limited.

It would be helpful to have clarification around whether the score will be considered absolute or whether it will be used to inform a committee deliberation, as per a more typical NICE assessment process. This may be particularly important in cases where products just miss out on a value band, e.g. they achieve a score between 79 and 80, as with our Dummy 2. In addition, some level of committee discretion could serve as an alternative solution to the contextual concerns raised in section 4.1.

5. Summary and key takeaways

To summarise, we found that the system generally rewards:

- **Products with broad spectrum.** Criteria 1A and 1B have high weights and reward antibiotics that target multiple resistant pathogens and resistance mechanisms.
- **Products against gram-negative priority pathogens.** Many of the criteria are geared towards incentivising gram-negative pathogens. The dummy product targeting gram-positive bacteria scored substantially lower.
- **Products that are more innovative in terms of chemical structure or mechanism of action**. Dummy products that reach band 2 (critical antibiotics) are those which are classified by the WHO as innovative or novel. Very few of these are considered likely to reach market in the short term. In contrast, dummies that represent products that are more likely to reach market, particularly in the short to medium term, do not score well. These are the less innovative antibiotic dummies, which demonstrated reasonably high scores on target pathogens and absence of resistance, but still have difficulties attaining sufficient points to qualify for the lowest value bands when they do not provide supporting clinical evidence (criterion 1D). These product profiles represent a significant proportion of the pipeline (>25%, Annex A1).
- **Products that have been tested in patients with resistant infections showing superiority over current standard of care**. The weight held by this criterion (1D) suggests that such evidence will be a key driver in attainment of higher value bands. This was supported by the results of the sensitivity analyses. Experts did not consider the higher levels which required RCT evidence to be obtainable, particularly in the short term.
- **Products that address an area of unmet need in the UK in terms of high morbidity/mortality, or where there are no treatment alternatives**. Again, the weight held by this criterion (1C) suggests that such evidence will be a key driver in attainment of higher value bands. This was supported by the results of the sensitivity analyses.
- **Products that are administered orally.** This is appropriate in some contexts, but not others.

We propose four key-takeaways from this analysis:

1. **Realistic dummy products can be constructed from the pipeline that reach value bands 2-4, representing annual payments of £5-15m.** Achievement of these scores is contingent on many assumptions, including meeting challenging evidence requirements.

2. **The highest scoring dummy product (dummy 2) achieved a score of 79.2, just missing the top value band (minimum score 80).** It had the following characteristics:
   - Targets gram negative pathogens and resistance mechanisms, leading to good scores in category 1 (particularly 1A and 1B).
   - Categorised as an innovative traditional antibiotic by WHO, leading to high scores in category 2, with a better chance of preventing the emergence of resistance and cross resistance and impact on microbiota than less novel products.
   - Available in oral and IV formulations, providing flexibility and convenience, leading to good scores in category 3.
   - Supported by non-randomised clinical evidence of effectiveness in resistant pathogens. If it was supported by a higher level of evidence, or the requirements for this criterion (1D) were relaxed even slightly, this dummy would reach value band 1.

3. **Not all products in the pipeline will achieve sufficient scores to qualify for a value band. Less innovative antibiotics, which are most likely to reach market in the short to medium term, do not score well.** These dummies demonstrate reasonably high scores regarding target pathogens and absence of resistance, but still have difficulties attaining sufficient points to qualify for the lowest value bands when they do not provide evidence in line with evidence requirements for effectiveness (criterion 1D). These product profiles represent a significant proportion of the pipeline (>25%).

4. **It will be particularly challenging for products to achieve a good score on criterion 1D, related to evidence of clinical effectiveness.** On this criterion, RCT evidence in resistant populations is required to obtain a score greater than 50. RCTs are not considered to be feasible in these small populations due to ethical concerns, statistical limitations and practical and commercial considerations. With this score effectively capped at 50, none of the dummy products reach the top value band. The impact of these demanding requirements is so substantial due to the weight carried by this criterion (11.3% of the total score).

Based on the insights summarised here and in our related work (Hofer and Hampson, 2023), we suggest that changes to the proposals are needed to ensure valuable and critically needed new antimicrobials are adequately incentivised. Updates to the proposals in line with the findings in this report will help NHS England and international policy makers to achieve the aim of the pull incentive in stimulating much needed investment in novel antimicrobials.

Annex

A1. WHO pipeline

The WHO pipeline 2022 encompasses 80 products across all phases of development. We extracted the following key variables: novelty/type of product, type of administration, coverage of WHO priority pathogen lists, links to clinical trials. Products were divided into non-traditional antimicrobial, innovative (traditional) antibiotic, (traditional) antibiotic. These three product categories were analysed in terms of the key variables. Links to clinical trials were used to establish if trials were specific to drug resistant pathogens.

Presented below (Table A1) is a heat map representing proportions of products in different product groups and key characteristic categories. Figures below 30% are highlighted in green, 30-65 are yellow and orange, and those over 65 are highlighted in red.

TABLE A1: WHO PIPELINE EXTRACTION OF KEY CHARACTERISTICS

	Non-traditional Antimicrobial	Innovative Antibiotic	Antibiotic
Number of products	N= 34	N=17	N=29
Inhalation administration	14.7%	0.0%	0.0%
IV administration	32.4%	5.9%	48.3%
IV & Oral administration	0.0%	17.6%	10.3%
Oral administration	47.1%	58.8%	41.4%
Covers all critical priority pathogens	61.8%	35.3%	75.9%
Covers other priority pathogens	61.8%	35.3%	79.3%
Clinical dev in drug resistant pathogens	5.9%	11.8%	3.4%

When taking into account types of product, novelty and types of administration, the dummy profiles collectively cover about 66% of the WHO pipeline (Figure A1). The other 34% represent mostly non-traditional antimicrobials, which have been excluded by design as it is unclear if these types of products will be eligible for the scheme.

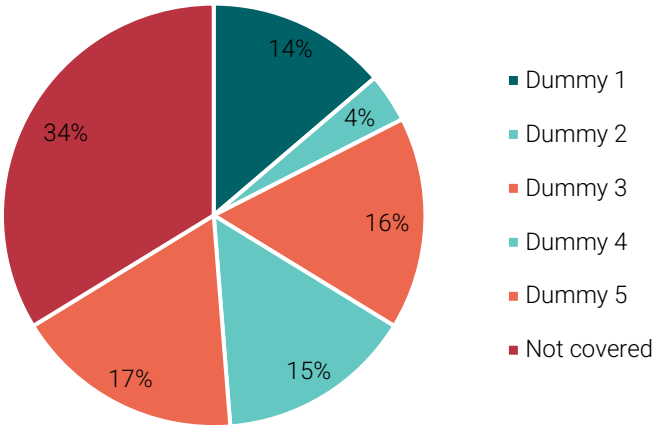


FIGURE A1: PERCENTAGE OF THE WHO PIPELINE COVERED BY DUMMY PRODUCT PROFILES, BASED ON TYPE OF PRODUCT, NOVELTY, AND TYPE(S) OF ADMINISTRATION (N=80)

A2. Dummy product scores

TABLE A2: SCORES FOR DUMMY PRODUCT 1

Category & Weight		Criteria & Weight		Levels & Points									
				L1	L2	L3	L4	L5	L6	L7	L8	L9	L10
Relative Effectiveness and Unmet Clinical Need	x0.45	1A Activity against WHO priority pathogens	x0.27	36	32	32	28	26	22	21	20	18	14
		1B Activity against clinically relevant resistance mechanisms	x0.23	29	26	24	21	19	17	17	15	14	
		1C Activity against UK unmet needs	x0.25	100	45	0							
		1D Clinical effectiveness compared to best standard of care	x0.25	100	80	70	60	50	0				
Pharmacological Benefit	x0.25	2A Chemical entity novelty	x0.17	100	75	45	0						
		2B Target site novelty	x0.16	100	80	0							
		2C Penetration to relevant anatomical site	x0.18	100	85	0							
		2D Absence of cross resistance	x0.18	100	70	30	0						
		2E Absence of rapidly emerging resistance	x0.16	100	75	25	0						
		2F Reduced impact on microbiota	x0.14	100	60	0							
Health System Benefit	x0.3	3A Adverse Events	x0.18	100	80	30	0						
		3B Drug -drug interactions	x0.13	100	65	45	0						
		3C Mode of administration	x0.15	100	95	85	45	30	0				
		3D Dose Frequency	x0.15	100	95	65	50	0					
		3E Product stability and storage	x0.09	100	45	0							
		3F Monitoring requirements	x0.12	100	40	0							
		3G Reduced hospital admissions or length of stay	x0.18	100	60	30	0						

TABLE A3: SCORES FOR DUMMY PRODUCT 2

Category & Weight		Criteria & Weight		Levels & Points									
				L1	L2	L3	L4	L5	L6	L7	L8	L9	L10
Relative Effectiveness and Unmet Clinical need	x0.45	1A Activity against WHO priority pathogens	x0.27	36	32	32	28	26	22	21	20	18	14
		1B Activity against clinically relevant resistance mechanisms	x0.23	29	26	24	21	19	17	17	15	14	
		1C Activity against UK unmet needs	x0.25	100	45	0							
		1D Clinical effectiveness compared to best standard of care	x0.25	100	80	70	60	50	0				
Pharmacological Benefit	x0.25	2A Chemical entity novelty	x0.17	100	75	45	0						
		2B Target site novelty	x0.16	100	80	0							
		2C Penetration to relevant anatomical site	x0.18	100	85	0							
		2D Absence of cross resistance	x0.18	100	70	30	0						
		2E Absence of rapidly emerging resistance	x0.16	100	75	25	0						
		2F Reduced impact on microbiota	x0.14	100	60	0							
Health System benefit	x0.3	3A Adverse Events	x0.18	100	80	30	0						
		3B Drug -drug interactions	x0.13	100	65	45	0						
		3C Mode of administration	x0.15	100	95	85	45	30	0				
		3D Dose Frequency	x0.15	100	95	65	50	0					
		3E Product stability and storage	x0.09	100	45	0							
		3F Monitoring requirements	x0.12	100	40	0							
		3G Reduced hospital admissions or length of stay	x0.18	100	60	30	0						

TABLE A4: SCORES FOR DUMMY PRODUCT 3

Category & Weight		Criteria & Weight		Levels & Points									
				L1	L2	L3	L4	L5	L6	L7	L8	L9	L10
Relative Effectiveness and Unmet Clinical need	x0.45	1A Activity against WHO priority pathogens	x0.27	36	32	32	28	26	22	21	20	18	14
		1B Activity against clinically relevant resistance mechanisms	x0.23	29	26	24	21	19	17	17	15	14	
		1C Activity against UK unmet needs	x0.25	100	45	0							
		1D Clinical effectiveness compared to best standard of care	x0.25	100	80	70	60	50	0				
Pharmacological Benefit	x0.25	2A Chemical entity novelty	x0.17	100	75	45	0						
		2B Target site novelty	x0.16	100	80	0							
		2C Penetration to relevant anatomical site	x0.18	100	85	0							
		2D Absence of cross resistance	x0.18	100	70	30	0						
		2E Absence of rapidly emerging resistance	x0.16	100	75	25	0						
		2F Reduced impact on microbiota	x0.14	100	60	0							
Health System benefit	x0.3	3A Adverse Events	x0.18	100	80	30	0						
		3B Drug -drug interactions	x0.13	100	65	45	0						
		3C Mode of administration	x0.15	100	95	85	45	30	0				
		3D Dose Frequency	x0.15	100	95	65	50	0					
		3E Product stability and storage	x0.09	100	45	0							
		3F Monitoring requirements	x0.12	100	40	0							
		3G Reduced hospital admissions or length of stay	x0.18	100	60	30	0						

TABLE A6: SCORES FOR DUMMY PRODUCT 4

Category & Weight		Criteria & Weight		Levels & Points									
				L1	L2	L3	L4	L5	L6	L7	L8	L9	L10
Relative Effectiveness and Unmet Clinical need	x0.45	1A Activity against WHO priority pathogens	x0.27	36	32	32	28	26	22	21	20	18	14
		1B Activity against clinically relevant resistance mechanisms	x0.23	29	26	24	21	19	17	17	15	14	
		1C Activity against UK unmet needs	x0.25	100	45	0							
		1D Clinical effectiveness compared to best standard of care	x0.25	100	80	70	60	50	0				
Pharmacological Benefit	x0.25	2A Chemical entity novelty	x0.17	100	75	45	0						
		2B Target site novelty	x0.16	100	80	0							
		2C Penetration to relevant anatomical site	x0.18	100	85	0							
		2D Absence of cross resistance	x0.18	100	70	30	0						
		2E Absence of rapidly emerging resistance	x0.16	100	75	25	0						
		2F Reduced impact on microbiota	x0.14	100	60	0							
Health System benefit	x0.3	3A Adverse Events	x0.18	100	80	30	0						
		3B Drug-drug interactions	x0.13	100	65	45	0						
		3C Mode of administration	x0.15	100	95	85	45	30	0				
		3D Dose Frequency	x0.15	100	95	65	50	0					
		3E Product stability and storage	x0.09	100	45	0							
		3F Monitoring requirements	x0.12	100	40	0							
		3G Reduced hospital admissions or length of stay	x0.18	100	60	30	0						

TABLE A7: SCORES FOR DUMMY PRODUCT 5

Category & Weight		Criteria & Weight		Levels & Points									
				L1	L2	L3	L4	L5	L6	L7	L8	L9	L10
Relative Effectiveness and Unmet Clinical need	x0.45	1A Activity against WHO priority pathogens	x0.27	36	32	32	28	26	22	21	20	18	14
		1B Activity against clinically relevant resistance mechanisms	x0.23	29	26	24	21	19	17	17	15	14	
		1C Activity against UK unmet needs	x0.25	100	45	0							
		1D Clinical effectiveness compared to best standard of care	x0.25	100	80	70	60	50	0				
Pharmacological Benefit	x0.25	2A Chemical entity novelty	x0.17	100	75	45	0						
		2B Target site novelty	x0.16	100	80	0							
		2C Penetration to relevant anatomical site	x0.18	100	85	0							
		2D Absence of cross resistance	x0.18	100	70	30	0						
		2E Absence of rapidly emerging resistance	x0.16	100	75	25	0						
		2F Reduced impact on microbiota	x0.14	100	60	0							
Health System benefit	x0.3	3A Adverse Events	x0.18	100	80	30	0						
		3B Drug-drug interactions	x0.13	100	65	45	0						
		3C Mode of administration	x0.15	100	95	85	45	30	0				
		3D Dose Frequency	x0.15	100	95	65	50	0					
		3E Product stability and storage	x0.09	100	45	0							
		3F Monitoring requirements	x0.12	100	40	0							
		3G Reduced hospital admissions or length of stay	x0.18	100	60	30	0						

A3. Sensitivity analyses

TABLE A8: SENSITIVITY ANALYSIS FOR CRITERION 1C (BASELINE HIGHLIGHTED IN GREY)

Level	Dummy 1	Dummy 2	Dummy 3	Dummy 4	Dummy 5
100	80.5	85.4	68.0	58.6	52.2
45	74.3	79.2	61.8	52.4	46.1
0	69.2	74.2	56.8	47.3	41.0

TABLE A9: SENSITIVITY ANALYSIS FOR CRITERION 1D (BASELINE HIGHLIGHTED IN GREY)

Level	Dummy 1	Dummy 2	Dummy 3	Dummy 4	Dummy 5
100	79.9	84.8	67.5	63.6	57.3
80	77.7	82.6	65.2	61.4	55.1
70	76.5	81.5	64.1	60.3	53.9
60	75.4	80.3	63.0	59.1	52.8
50	74.3	79.2	61.8	58.0	51.7
0	68.7	73.6	56.2	52.4	46.1

TABLE A10: SENSITIVITY ANALYSIS FOR CRITERION 2E (BASELINE HIGHLIGHTED IN GREY)

Level	Dummy 1	Dummy 2	Dummy 3	Dummy 4	Dummy 5
100	74.3	80.2	62.8	53.4	47.1
75	73.3	79.2	61.8	52.4	46.1
25	71.3	77.2	59.8	50.4	44.1
0	70.3	76.2	58.8	49.4	43.1

TABLE A11: SENSITIVITY ANALYSIS FOR CRITERION 3A (BASELINE HIGHLIGHTED IN GREY)

Level	Dummy 1	Dummy 2	Dummy 3	Dummy 4	Dummy 5
100	75.4	79.2	61.8	52.4	47.1
80	74.3	78.1	60.7	51.3	46.1
30	71.6	75.4	58.0	48.6	43.4
0	70.0	73.8	56.4	47.0	41.7

TABLE A12: SENSITIVITY ANALYSIS FOR CRITERION 3G (BASELINE HIGHLIGHTED IN GREY)

Level	Dummy 1	Dummy 2	Dummy 3	Dummy 4	Dummy 5
100	76.4	81.4	64.0	54.6	48.2
60	74.3	79.2	61.8	52.4	46.1
30	72.7	77.6	60.2	50.8	44.4
0	71.0	76.0	58.6	49.2	42.8

6. References

Årdal, C., Balasegaram, M., Laxminarayan, R., McAdams, D., Outterson, K., Rex, J.H. and Sumpradit, N., 2020. Antibiotic development - economic, regulatory and societal challenges. *Nature Reviews. Microbiology*, 18(5), pp.267–274. 10.1038/s41579-019-0293-3.

Brassel, S., Al Taie, A. and Steuten, L., 2023. Value assessment of antimicrobials using the STEDI framework – How steady is the outcome? *Health Policy*, 136, p.104892. 10.1016/j.healthpol.2023.104892.

Butler, M.S., Gigante, V., Sati, H., Paulin, S., Al-Sulaiman, L., Rex, J.H., Fernandes, P., Arias, C.A., Paul, M., Thwaites, G.E., Czaplewski, L., Alm, R.A., Lienhardt, C., Spigelman, M., Silver, L.L., Ohmagari, N., Kozlov, R., Harbarth, S. and Beyer, P., 2022. Analysis of the Clinical Pipeline of Treatments for Drug-Resistant Bacterial Infections: Despite Progress, More Action Is Needed. *Antimicrobial Agents and Chemotherapy*, 66(3), pp.e01991-21. 10.1128/aac.01991-21.

Dane, A., Rex, J.H., Newell, P. and Stallard, N., 2022. The Value of the Information That Can Be Generated: Optimizing Study Design to Enable the Study of Treatments Addressing an Unmet Need for Rare Pathogens. *Open Forum Infectious Diseases*, 9(7), p.ofac266. 10.1093/ofid/ofac266.

Dutescu, I.A. and Hillier, S.A., 2021. Encouraging the Development of New Antibiotics: Are Financial Incentives the Right Way Forward? A Systematic Review and Case Study. *Infection and Drug Resistance*, 14, pp.415–434. 10.2147/IDR.S287792.

European Centre for Disease Prevention and Control and WHO Regional Office for Europe, 2022. *Antimicrobial resistance surveillance in Europe 2022 - 2020*. [online] Available at: https://www.ecdc.europa.eu/en/publications-data/antimicrobial-resistance-surveillance-europe-2022-2020-data [Accessed 23 Aug. 2023].

European Medicines Agency, 2022. *Guideline on the evaluation of medicinal products indicated for treatment of bacterial infections [CPMP/EWP/558/95 Rev 3]*.

Hofer, M.P. and Hampson, G., 2023. *Proposals for a Novel UK Antimicrobial Subscription Model: The Investor Perspective*. [online] OHE Contract Research Report, London: Office of Health Economics. Available at: www.ohe.org/publications/novel-uk-antimicrobial-subscription-model-investor-perspective/ .

Karlsberg Schaffer, S., West, P., Towse, A., Henshall, C., Mestre-Ferrandiz, J. and Fischer, A., 2017. Assessing the Value of New Antibiotics: Additional Elements of Value for Health Technology Assessment Decisions.

Lanini, S., Ioannidis, J.P.A., Vairo, F., Pletschette, M., Portella, G., Di Bari, V., Mammone, A., Pisapia, R., Merler, S., Nguhuni, B., Langer, M., Di Caro, A., Edwards, S.J.L., Petrosillo, N., Zumla, A. and Ippolito, G., 2019. Non-inferiority versus superiority trial design for new antibiotics in an era of high antimicrobial resistance: the case for post-marketing, adaptive randomised controlled trials. *The Lancet Infectious Diseases*, 19(12), pp.e444–e451. 10.1016/S1473-3099(19)30284-1.

Leonard, C., Crabb, N., Glover, D., Cooper, S., Bouvy, J., Wobbe, M. and Perkins, M., 2023. Can the UK 'Netflix' Payment Model Boost the Antibacterial Pipeline? *Applied Health Economics and Health Policy*, 21(3), pp.365–372. 10.1007/s40258-022-00786-1.

NHS England, 2023. *The Antimicrobial Products Subscription Model: consultation on proposals*. [online] Available at: https://www.engage.england.nhs.uk/survey/the-antimicrobial-products-subscription-model/ [Accessed 12 Jul. 2023].

NICE, 2023. *Models for the evaluation and purchase of antimicrobials*. [CorporatePage] NICE. Available at: https://www.nice.org.uk/about/what-we-do/life-sciences/scientific-advice/models-for-the-evaluation-and-purchase-of-antimicrobials [Accessed 12 Jul. 2023].

NICE Decision Support Unit, 2022. *Homepage*. [online] Available at: https://www.sheffield.ac.uk/nice-dsu [Accessed 10 Aug. 2023].

Outterson, K. and Rex, J.H., 2020. Evaluating for-profit public benefit corporations as an additional structure for antibiotic development and commercialization. *Translational Research*, 220, pp.182–190. 10.1016/j.trsl.2020.02.006.

Prasad, N.K., Seiple, I.B., Cirz, R.T. and Rosenberg, O.S., 2022. Leaks in the Pipeline: a Failure Analysis of Gram-Negative Antibiotic Development from 2010 to 2020. *Antimicrobial Agents and Chemotherapy*, 66(5), pp.e00054-22. 10.1128/aac.00054-22.

Rex, J.H. and Outterson, K., 2016. Antibiotic reimbursement in a model delinked from sales: a benchmark-based worldwide approach. *The Lancet. Infectious Diseases*, 16(4), pp.500–505. 10.1016/S1473-3099(15)00500-9.

Rezaei, J., Arab, A. and Mehregan, M., 2022. Equalizing bias in eliciting attribute weights in multiattribute decision-making: experimental research. *Journal of Behavioral Decision Making*, 35(2), p.e2262. 10.1002/bdm.2262.

Rothery, C., Woods, B., Schmitt, L., Claxton, K., Palmer, S. and Sculpher, M., 2018. FRAMEWORK FOR VALUE ASSESSMENT OF NEW ANTIMICROBIALS.

Tervonen, T., Gelhorn, H., Sri Bhashyam, S., Poon, J.-L., Gries, K.S., Rentz, A. and Marsh, K., 2017. MCDA swing weighting and discrete choice experiments for elicitation of patient benefit-risk preferences: a critical assessment. *Pharmacoepidemiology and Drug Safety*, 26(12), pp.1483–1491. 10.1002/pds.4255.

Wellcome Trust, 2020. *The growing crisis in antibiotic R&D*.

World Health Organization, 2022. *Antibacterial products in clinical development for priority pathogens*. [online] Available at: https://www.who.int/observatories/global-observatory-on-health-research-and-development/monitoring/antibacterial-products-in-clinical-development-for-priority-pathogens [Accessed 13 Jul. 2023].

About us

With over 60 years of expertise, the Office of Health Economics (OHE) is the world's oldest independent health economics research organisation. Every day we work to improve health care through pioneering and innovative research, analysis, and education.

As a global thought leader and publisher in the economics of health, health care, and life sciences, we partner with Universities, Government, health systems and the pharmaceutical industry to research and respond to global health challenges.

As a government-recognised Independent Research Organisation and not-for-profit, our international reputation for the quality and independence of our research is at the forefront of all we do. OHE provides independent and pioneering resources, research and analyses in health economics, health policy and health statistics. Our work informs decision-making about health care and pharmaceutical issues at a global level.

All of our work is available for free online at www.ohe.org.

Areas of expertise

- Evaluation of health policy
- The economics of health care systems
- Health technology assessment (HTA) methodology and approaches
- HTA's impact on decision making, health care spending and the delivery of care
- Pricing and reimbursement for biologics and pharmaceuticals, including value-based pricing, risk sharing and biosimilars market competition
- The costs of treating, or failing to treat, specific diseases and conditions
- Drivers of, and incentives for, the uptake of pharmaceuticals and prescription medicines
- Competition and incentives for improving the quality and efficiency of health care
- Incentives, disincentives, regulation and the costs of R&D for pharmaceuticals and innovation in medicine
- Capturing preferences using patient-reported outcomes measures (PROMs) and time trade-off (TTO) methodology
- Roles of the private and charity sectors in health care and research
- Health and health care statistics

CONTRACT RESEARCH REPORT
NOVEMBER 2023

ohe.org

The Office of Health Economics A Company Limited by Guarantee of Registered No.09848965
OHE Consulting Ltd Registered Company No.09853113
OHE is a Charity Registration No.1170829
Registered Office 2nd Floor Goldings House, Hay's Galleria, 2 Hay's Lane, London, SE1 2HB