National Clinical Guideline Centre

Cost-effectiveness analysis

Pneumonia (community- and hospital-acquired)

Cost-effectiveness analysis for microbiological tests in patients with moderate- and high-severity community-acquired pneumonia

Clinical Guideline 191

Appendix L

3 December 2014

Final version

Commissioned by the National Institute for Health and Care Excellence

Funding
National Institute for Health and Care Excellence

Contents

1 Cost-effectiveness analysis or microbiological tests in patients with moderate- and high-severity community acquired pneumonia

1.1 Methods

1.1.1 Model overview

Patients with moderate- and high-severity community-acquired pneumonia (CAP) commonly receive a suite of microbiological tests on admission to hospital with the hope of isolating a causative pathogen. Due to high mortality rates, correct antibiotic treatment is essential. However, it is unknown if the additional cost of tests to identify the pathogen/s provide additional benefits in terms of patient outcomes.

The GDG identified this area as a high priority for original economic analysis for patients with moderate- and high-severity CAP. Low-severity CAP is associated with a low mortality rate and an economic analysis on this population was not prioritised as the benefit of conducting microbiological tests in this population is likely to be negligible and the GDG did not recommend them in people with low-severity CAP.

This economic analysis addresses the question:

In adults with moderate- and high- severity CAP in a hospital setting, what microbiological test or combination of tests at presentation is the most cost effective?

1.1.1.1 Comparators

There are multiple microbiological testing strategies for those admitted to hospital with moderate- and high-severity CAP. The most relevant strategies chosen by the GDG due to their common usage in the UK were analysed in this model:

- no testing (clinical judgement)
- blood culture
- sputum culture
- urinary pneumococcal antigen
- urinary legionella antigen
- a combination of a blood culture and a sputum culture
- a combination of a blood culture, a urinary pneumococcal antigen and a urinary legionella antigen
- all tests in combination.

1.1.1.2 Population

The population used in this analysis were patents of an equal male to female ratio, with an average age of 72 years admitted to hospital with moderate- or high-severity CAP.

After estimating the average probability of death in the model (see 1.1.3.5), we concluded that the base case analysis was more applicable to the moderate-severity CAP group while an additional analysis was undertaken to obtain the results for the high-severity CAP group. Apart from the baseline mortality and the empiric treatment, all the other parameters were assumed to be equal in

the two subgroups and therefore we will refer to moderate- and high-severity CAP throughout the parameters explanation.

1.1.1.3 Time horizon and perspectives used

The time horizon chosen for the model was a lifetime time horizon, with a single in-hospital episode including diagnosis and treatment over a short time period with a lifetime extrapolation. The analysis took the perspective of the NHS and personal and social services, in line with the NICE reference case.

1.1.1.4 Deviations from NICE reference case

As explained in section 1.1.3.6 no applicable quality-of-life data were identified for moderate- and high-severity CAP. An additional systematic search for quality-of-life data for severe sepsis was undertaken which provided better-quality evidence, and which was deemed applicable to this population by the GDG. As such, the quality-of-life data used in the model may not directly represent patients with moderate- and high-severity CAP. These data, obtained from a Dutch population, were reported as SF-36 scores and mapped onto EQ-5D scores in the study.[7]

Adverse events from antibiotic therapy were not considered in this analysis. However, the impact from adverse events was not expected to be significant, especially when compared with the impact of mortality which was established by the GDG to be the critical outcome. Also not considered in this analysis were benefits of antimicrobial stewardship, because benefits from reducing antibiotic resistance accrue to both the individual and society as a whole, and mechanisms to estimate these benefits in a decision model have not yet been established. Therefore, the QALY gain associated with a strategy that increases targeted treatment may be understated.

It was not appropriate to discount costs in this analysis as all costs were incurred within the first 30 days. However, given that health outcomes were extrapolated to a lifetime time horizon, QALYs were discounted by 3.5% per annum as in the NICE reference case.

1.1.2 Modelling approach

1.1.2.1 Model structure

A full model structure is provided in section 1.4.

The population, as detailed in section 1.1.1.2, was tested using a microbiological strategy as detailed in section 1.1.1.1. Dependent on the pathogen present, patients were then given either targeted treatment or empirical treatment was continued. This change or continuation of treatment could have been based on the correct or incorrect identification of the pathogen which in turn depended on the sensitivity and specificity of the test/ tests used. As a result, the probability of patients being alive or dead at 30 days was determined by whether patients received the appropriate antibiotic treatment or not, as well as on pathogen-specific mortality probabilities. Costs and QALYs were determined by the initial strategy adopted and the probability of incorrect (falsely positive and falsely negative) test results and their outcomes, namely the increase in mortality. After 30 days, the model assumes there is no impact of pneumonia on mortality and standard UK life expectancies were used to generate lifetime QALYs. This model is unable to quantify some benefits of targeted treatment such as a reduction in adverse events or the reduction in antimicrobial resistance.

1.1.2.2 Assumptions

Due to lack of certain data, and pragmatic constraints relating to model complexity, a number of assumptions were made to facilitate the development of this model. These assumptions were agreed in discussion with the GDG and are detailed below.

Pathogens and tests

- In order to make the model feasible, it was assumed that patients have only a single causative pathogen, so that the overall pathogens prevalence adds up to 1. However, in real clinical practice more than 1 pathogen can be present and this was acknowledged in the treatment management assumptions of the model, where in some circumstances (for example when 2 tests performed in combination showed positive results to 2 different pathogens) treatment could cover more than 1 organism.
- Based on the prevalence in the UK, the only pathogens considered were:
  - *Streptococcus pneumoniae*
  - *Haemophilus influenzae*
  - *Staphylococcus aureus*
    - *'Staphylococcus species' (initial result showing Staphylococci awaiting species typing)*
  - *L. pneumophila*
  - 'Atypical' pathogens
  - Gram-negative pathogens
- Different tests in routine use detect different pathogens as described in Table 1**.**
  - Blood culture could detect:
    - *S. pneumoniae*, *H. influenzae*, *S. aureus* and Gram negative organisms
  - Routine sputum culture could detect:
    - *S. pneumoniae*, *H. influenzae*, *S. aureus* and Gram negative organisms
  - Urinary pneumococcal antigen could detect:
    - *S. pneumoniae*
  - Urinary legionella antigen could detect:
    - *Legionella pneumophila* serogroup 1
  - No routine test could reliably detect atypical pathogens

Table 1: Detection of pathogens by single test

	S. pneumonia	***H. influenza***	***S. aureus***	***L. pneumophila***	**Atypical pathogens**	**Gram-negative pathogens**
Blood culture	Yes	Yes	Yes	No	No	Yes
Routine sputum culture	Yes	Yes	Yes	No	No	Yes
Urinary pneumococcal antigen	Yes	No	No	No	No	No
Urinary legionella antigen	No	No	No	Yes	No	No

- Tests were assumed to produce the following false positive results in certain circumstances:
  - Blood culture could only be false positive to:

- *'Staphylococcus* species' – *see section 1.1.3.3*
  - Sputum culture could only be false positive to:
    - *S. pneumonia*, *H. influenza* and Gram negative organisms.
  - In combination strategies, urinary antigen tests do not produce false positive results. This assumption was used for model simplification as the specificity of these tests is around 100% when given alone.
- Sensitivities and specificities for some tests had to be based on expert opinion, as shown in Table 4.
- Sensitivity for all *Staphylococcus* species was assumed to be the same as the contamination rate for positive blood cultures with 'coagulase negative staphylococci' at 5%.

Combinations

- In the combination of blood culture and sputum culture:
  - the result of the blood culture was trusted over sputum culture, unless the blood culture reported 'Staphylococcus species' in which case treatment for both organisms would be required.
- In the combination of blood culture and urinary antigen tests:
  - the result of the urinary *Legionella* antigen test over "all tests" was trusted, unless the blood culture reported 'Staphylococcus species' in which case treatment for both organisms would be required
  - the result of the urinary pneumococcal antigen test was trusted over blood culture even if 'Staphylococcus species' was reported.
- When all tests were done in combination:
  - generally the results of blood culture were trusted over other tests
  - the result of the urinary *Legionella* antigen over "all tests" was trusted, unless the blood culture reported 'Staphylococcus species' in which case treatment for both organisms was required
  - the result of the urinary pneumococcal antigen test was trusted over blood culture even if 'Staphylococcus species' was reported
  - blood culture results were trusted over sputum culture.

Treatment pathway

Treatments were defined as 'incorrect' if the pathogen was resistant to the antibiotic treatment as defined by Table 2.

- All patients were treated empirically with a narrow-spectrum beta-lactam and a macrolide for moderate-severity CAP or a broad-spectrum beta-lactam and a macrolide (the cost used in the model was based on patients receiving co-amoxiclav) for high-severity CAP. All patients were started on intravenous (IV) antibiotics with switch to oral antibiotics after two days. The proportion of those admitted to and time spent in an intensive care unit (ICU) was assumed to be similar across all pathogens, as this parameter is most influenced by severity of pneumonia rather than pathogen. As such, the cost of ICU was not included in the model.
- The model did not allow for recurrence or relapse of pneumonia.
- All patients had a hospital stay of at least seven days.
- Patients treated 'incorrectly' had an additional three days' length of stay (LOS) over those treated 'appropriately'.

Table 2: Antibiotic susceptibility

	S. pneumonia	*H. influenza*	*S. aureus*	*L. pneumophila*	Atypical pathogens	Gram-negative pathogens
Empirical	S	S	S	S	S	S/R[(b)]
Broad-spectrum beta-lactam	S	T	S	R	R	S/R[(b)]
Narrow-spectrum beta-lactam	T	S[(a)]	R	R	R	R
Flucloxacillin	S	R	T	R	R	R
Macrolide	S	S/R[(C)]	S	S	S	R
Fluoroquinolone	S	S	S	T	S	S
Piperacillin with tazobactam	S	S	S	R	R	T

Note: S = susceptible, R = resistant, T = targeted treatment
(a) Susceptible but not to benzylpenicillin
(b) Some susceptible, some resistant
(c) H. influenza could be either resistant or have intermediate susceptibility to macrolides

Changes in management

- A change in management was defined as a change in antibiotic prescription only.
- The pathogen detected dictated the change in antibiotics as per targeted treatment reported on Table 2:
  - If *S. pneumonia* was detected, it was assumed treatment would consist only of narrow-spectrum beta-lactam
    - Patients who deteriorated, or did not respond to (incorrectly treated) narrow-spectrum beta-lactam would be switched to a broad-spectrum beta-lactam after 48 hours.
  - If *H. influenza* was detected, it was assumed treatment would consist only of broad-spectrum beta-lactam
  - If *S. aureus* was detected, it was assumed that antibiotic treatment would be changed to flucloxacillin
    - If *'Staphylococcus species'* was detected, 24 hours of flucloxacillin was added to empirical treatment (to allow for further typing of the staphylococcus species). Initial false positives would be treated with flucloxacillin in addition to empirical treatment only for 24 hours.
  - If *Legionella pneumophila* was detected, it was assumed that treatment would be changed to a fluoroquinolone
  - If a Gram-negative pathogen was detected, it was assumed that a switch to piperacillin with tazobactam would be made.
    - Patients correctly treated with piperacillin with tazobactam would remain on IV antibiotic for seven days, due to the nature of Gram-negative pathogens.
    - Patients who deteriorated, or did not respond to (incorrectly treated) IV piperacillin with tazobactam would be switched to another broad-spectrum beta-lactam after 48 hours.
- If a patient tested negative after all the tests envisaged in the strategy, empirical treatment would be continued without further tests.

Quality of life

- It was assumed that patients with moderate- and high-severity CAP would only ever return to 95% of their pre-pneumonia quality-of-life, which would occur after six months.

- Severe sepsis was used a proxy for moderate- and high-severity pneumonia due to quality-of-life data limitations. See section 1.1.3.6 for details.

Mortality

- Due to mortality data limitations, the GDG refined mortality estimates, using published and unpublished data and clinical experience. This was done through discussion and a consensus was agreed. See section 1.1.3.5 for details.
- Mortality was assumed to happen within 30 days. After 30 days, the model assumes mortality is not affected by pneumonia.

1.1.2.2.1 Uncertainty

Probabilistic sensitivity analysis

Where possible, the model was built probabilistically to take account of the uncertainty around input parameter point estimates. A probability distribution was defined for each model input parameter which was to be modelled in this way. When the model was run, a value for each input was randomly selected simultaneously from its respective probability distribution; mean costs and mean incremental QALYs were calculated using these values. The model was run repeatedly – 20,000 times for the base case and sensitivity analyses, where appropriate, and the results summarised.

In addition, various deterministic sensitivity analyses were undertaken to test the robustness of model assumptions. In these, one or more inputs were changed and the analysis rerun to evaluate the impact on results and whether conclusions on which intervention should be recommended would change. Threshold analyses were also conducted which allowed determination of the threshold at which the value of a particular parameter is likely to change the conclusion.

The way in which distributions are defined reflects the nature of the data, so for example, costs were given a gamma distribution, which is bounded by zero but positively skewed, reflecting the true nature of costs. All of the variables that were probabilistic in the model and their distributional parameters are detailed in Table 3 and in the relevant input summary tables in section 1.1.3.1. Probability distributions in the analysis were parameterised using error estimates from data sources.

Table 3: Description of the type and properties of distributions used in the probabilistic sensitivity analysis

Parameter	Type of distribution	Properties of distribution
Prevalence of pathogens; sensitivity and specificity	Beta	Bounded between 0 and 1. As the sample size and the number of events were specified alpha and Beta values were calculated as follows: Alpha = (number of patients hospitalised) Beta = (Number of patients)-(number of patients hospitalised)
Costs, quality of life decrement	Gamma	Bounded at 0, positively skewed. Derived from mean and its standard error. Alpha and Beta values were calculated as follows: Alpha = $(mean/SE)^2$ Beta = $SE^2/mean$ Where costs were based on GDG opinion Alpha and Lambda values were calculated as follows: Alpha = $(mean/SE)^2$ Lambda = $mean/SE^2$

Parameter	Type of distribution	Properties of distribution
NHS Reference Costs (diagnostic and treatment)	Lognormal	Where appropriate, the lognormal distribution may provide a better fit than the gamma distribution for costs. The natural log of the mean was calculated as follows: Natural log of the mean = [Ln(mean) – (lnSE)^2]/2 Where the natural log of the standard error (lnSE) was calculated by: $\sqrt{\ln \frac{SE^2 + mean^2}{mean^2}}$

The following variables, were left deterministic (were not varied in the probabilistic analysis): the cost-effectiveness threshold (which was deemed to be fixed by NICE) and the resource, including time and staff costs, required to implement each strategy (assumed to be fixed according to national pay scales and programme content), length of hospital stay, cost of antibiotic treatment, or life expectancy.

1.1.3 Model inputs

1.1.3.1 Summary table of model inputs

Model inputs were based on clinical evidence identified in the systematic review undertaken for the guideline, supplemented by additional data sources as required. Model inputs were validated with clinical members of the GDG. A summary of the model inputs used in the base case (primary) analysis is provided in Table 4 below. More details about sources, calculations and rationale for selection can be found in the sections following this summary table.

Table 4: Summary of base case model inputs and parameter distributions used in the model

Parameter description	Point estimate	Probability distribution	Distribution parameters	Source
Patient characteristics				
Age when starting model	72	-	-	Hospital Episode Statistics[6]
Discounted life expectancy at start of model	10.819 years	-	-	Interim life tables[13] – see 1.1.3.5
Prevalence of pathogens				
S. pneumonia	0.6341	NA	NA	Calculated using prevalence reported in BTS Guidelines 2009[10] – see 1.1.3.2
H. influenza	0.0846	Beta	α = 59, β = 640	Calculated using prevalence reported in BTS Guidelines 2009[10] – see 1.1.3.2
S. aureus	0.0309	Beta	α = 22, β = 678	Calculated using prevalence reported in BTS Guidelines 2009[10] – see 1.1.3.2

Parameter description	Point estimate	Probability distribution	Distribution parameters	Source
L. pneumophila	0.0585	Beta	α = 41, β = 658	Calculated using prevalence reported in BTS Guidelines 2009[10] – see 1.1.3.2
Atypical pathogens	0.1756	Beta	α = 123, β = 576	Calculated using prevalence reported in BTS Guidelines 2009[10] – see 1.1.3.2
Gram-negative pathogens	0.0163	Beta	α = 11, β = 688	Calculated using prevalence reported in BTS Guidelines 2009[10] – see 1.1.3.2
Cost of tests (£)				
Blood culture	£23.71	Gamma	α = 96.12, λ = 4.054	GDG expert opinion – see 1.1.3.4
Sputum culture	£19.37	Gamma	α = 96.12, λ = 4.054	GDG expert opinion – see 1.1.3.4
Urinary legionella antigen	£40	Gamma	α = 7.11, λ =0.1778	GDG expert opinion – see 1.1.3.4
Urinary pneumococcal antigen	£40	Gamma	α = 7.11, λ =0.1778	GDG expert opinion – see 1.1.3.4
Cost of hospital treatment (£ per day)				
Bed day - within standard LOS	£324	Lognormal	μ = 5.75, σ = 0.26	NHS reference costs 2011-12 – DZ11A, 'Non-elective long stay - Lobar, Atypical or Viral Pneumonia, with Major CC[3]
Bed day - excess LOS	£228	Lognormal	μ = 5.40, σ = 0.25	NHS reference costs 2011-12 – DZ11A, Non-elective long stay excess bed days - Lobar, Atypical or Viral Pneumonia, with Major CC[3]
Cost of antibiotics (£ per day)				
Broad-spectrum beta-lactam (IV)	£3.18	-	-	MIMS online (Augmentin)[1]
Narrow-spectrum beta-lactam (IV)	£4.38	-	-	MIMS online (Amoxil)[1]
Macrolide (IV)	£18.90	-	-	MIMS online (Klaricid)[1]
Flucloxacillin (IV)	£5.33	-	-	MIMS online (Magnapen)[1]
Piperacillin with tazobactam (IV)	£45.51	-	-	MIMS online (Tazocin)[1]
Fluoroquinolone (IV)	£52.80	-	-	MIMS online (Tavonic)[1]
Broad-spectrum beta-lactam (PO)	£1.02	-	-	MIMS online (Co-amoxiclav)[1]
Narrow-spectrum beta-lactam (PO)	£0.42	-	-	MIMS online (Amoxicillin)[1]

Parameter description	Point estimate	Probability distribution	Distribution parameters	Source
Macrolide (PO)	£0.47	-	-	MIMS online (Clarithromycin)[1]
Flucloxacillin (PO)	£0.36	-	-	MIMS online (Flucloxacillin)[1]
Fluoroquinolone (PO)	£4.03	-	-	MIMS online (Levofloxacin)[1]
Blood culture and sensitivities - sensitivity and specificity				
Sensitivity to *S. pneumonia*	0.25	Beta	α = 18.5, β = 55.5	GDG expert opinion
Sensitivity to *H. influenza*	0.25	Beta	α = 18.5, β = 55.5	GDG expert opinion
Sensitivity to *S. aureus*	0.25	Beta	α = 18.5, β = 55.5	GDG expert opinion
Sensitivity to Gram-negative pathogens	0.25	Beta	α = 18.5, β = 55.5	GDG expert opinion
Specificity to *Staphylococcus* spp	0.95	Beta	α = 0.3, β = 0.016	GDG expert opinion
Sputum culture – sensitivity and specificity				
Sensitivity to *S. pneumoniae*	0.55	Beta	α = 0.10.7, β = 8.755	GDG expert opinion and Barrett-Connor (1971)[2]
Sensitivity to *H. influenza*	0.55	Beta	α = 0.10.7, β = 8.755	GDG expert opinion and Barrett-Connor (1971)[2]
Sensitivity to *S. aureus*	0.80	Beta	α = 4.2, β = 1.05	GDG expert opinion
Sensitivity to Gram-negative pathogens	0.80	Beta	α = 4.2, β = 1.05	GDG expert opinion
Specificity to *S. pneumonia*	0.71	Beta	α = 6.54, β = 2.67	GDG expert opinion and Guckian et al (1978)[5]
Specificity to *H. influenza*	0.71	Beta	α = 6.54, β = 2.67	GDG expert opinion and Guckian et al (1978)[5]
Specificity to Gram - negative pathogens	0.74	Beta	α = 5.76, β = 2.024	GDG expert opinion and Guckian et al (1978)[5]
Urinary pneumococcal antigen - sensitivity and specificity				
Sensitivity to *S. pneumonia*	0.74	Beta	α =88.2, β = 31.0	Sinclair et al (2013)[17]
Specificity to *S. pneumonia*	0.97	Beta	α =65.2, β = 1.88	Sinclair et al (2013)[17]
Urinary legionella antigen - sensitivity and specificity				
Sensitivity to *L. pneumophila*	0.74	Beta	α = 122.4, β = 43.0	Shimada et l (2009)[15]
Specificity to *L. pneumophila*	0.99	Beta	α = 352.6, β = 3.2	Shimada et al (2009)[15]
Length of stay (days)				
Average LOS	7	-	-	BTS Audit – Personal communication
Additional LOS for incorrect treatment	3	-	-	GDG expert opinion
Average LOS of those who die	14	-	-	Mortensen et al (2002)[12]

Parameter description	Point estimate	Probability distribution	Distribution parameters	Source
Quality of life				
UK population average EQ-5D score	0.825			Kind et al (1998)[9]
Long-term quality of life proportion applied to the UK general population	95%	-	-	GDG assumption
Disutility over six months – correct treatment	0.033429	Gamma	α =1.94, λ =58.04	Calculated using Hofhuis et al (2008)[7] – see 1.1.3.6
Disutility over six months – incorrect/empirical treatment	0.035131	Gamma	α = 2.143, λ =60.99	Calculated using Hofhuis2008[7] – see 1.1.3.6
Probability of mortality – non-targeted treatment				
S. pneumonia	0.14	Beta	α = 18, β = 111	Lim et al (2001)[11]
H. influenza	0.05	Beta	α = 1, β = 19	Lim et al (2001)[11]
S. aureus	0.50	Beta	α = 2, β = 2	Lim et al (2001)[11]
L. pneumophila	0.11	Beta	α = 1, β = 8	Lim et al (2001)[11]
Atypical	0.05	Beta	α = 3, β = 57	Lim et al (2001)[11]
Gram-negative pathogens s	0.4	Beta	α = 14.6, β = 21.9	Lim et al (2001)[11] and GDG expert opinion – see 1.1.3.5
Probability of mortality – targeted treatment				
S. pneumonia	0.14	Beta	α = 18, β = 111	Lim et al (2001)[11]
H. influenza	0.05	Beta	α = 1, β = 19	Lim et al (2001)[11]
S. aureus	0.30	Beta	α = 17.2, β = 40.13	Lim et al (2001)[11] and GDG expert opinion
L. pneumophila	0.11	Beta	α = 1, β = 8	Lim et al (2001)[11]
Atypical	0.05	Beta	α = 3, β = 57	Lim et al (2001)[11]
Gram-negative pathogens	0.25	Beta	α = 1, β = 3	Lim et al (2001)[11]

CC = complications and comorbidities; IV = intravenous; PO = per os (orally)

1.1.3.2 Prevalence

The prevalence of pathogens in the UK was taken from the BTS CAP Guidelines.[10] In reality, many pathogens can cause pneumonia, 12.8% are viruses whilst in 30.8% of cases no pathogen is identified and 15% of those with an identified aetiology have multiple pathogens.[10] As discussed in the assumptions above, for simplification only one of six pathogens could cause pneumonia and multiple pathogens were not considered as the prevalence of pathogens in these patients is unknown. We used the prevalence of pathogens from hospital, as these were likely to be more closely aligned with moderate- and high-severity CAP than the prevalence found in the community. The prevalence of these six pathogens, as reported in the BTS CAP guidelines was 61.5%. These pathogens were chosen by the GDG as these were considered to be the most common pathogens. In agreement with the GDG, all pathogens were equally scaled up to 61.50 using a factor of 1.626 (100/61.50) so the sum of the prevalence of these pathogens in our model equalled 100%.

Table 5: Prevalence of pathogens

Pathogen	Prevalence from BTS[10] (%)	Prevalence for model (%)[a]
S. pneumonia	39.00	63.41
H. influenza	5.20	8.46
L. pneumophila	3.60	5.85
S. aureus	1.90	3.09
Atypical pathogens	10.80	17.56
Gram negative pathogens	1.00	1.63

(a) Scaled up using a factor of 100/61.50 = 1.626

When performing tests, blood culture can return a result of '*Staphylococcus species*' within 24 hours. However, at this stage there is uncertainty as the test is unable to determine if this is *S. aureus* or contamination *with S. epidermidis* and this uncertainty has been built into the model. At 48 hours, it is possible to accurately identify the staphylococcal species and adjust treatment accordingly. In order to run the model probabilistically, *S. pneumoniae* was not entered as a numerical value in the model. This pathogens prevalence is the residual of all the other pathogens so that the prevalence still sums to 100%.

1.1.3.3 Test accuracy

As there was a lack of data relating tests directly to clinical outcomes (e.g. mortality) in the clinical review that could be used for the model, sensitivity and specificity data were used. The clinical review did not collect information on each test for particular pathogens given the different antibiotics used. Some systematic reviews were available, but when studies on test accuracy for specific pathogens were not available, the GDG was asked to provide sensitivity and specificity rates for the various tests; however even when studies were available, some data was modified by the GDG due to data limitations, such as the age of the data, and uncertainty about whether the reported sensitivity and specificity applied to patients with moderate- and high-severity CAP. In addition, some data were not available for this population and had to be assumed, as is demonstrated in Table 4. Due to the scaling up of prevalence, the GDG was concerned that the test accuracy could have been overestimated. A sensitivity analysis was undertaken to analyse this.

The GDG was unaware of data on sensitivity or specificity of blood culture for *Staphylococcus* (any species). The GDG felt that as the false positive rate of coagulase-negative staphylococcus is 5%, a specificity of 95% could be assumed. Two recent meta-analyses were used to inform the accuracy of the urinary antigen tests.[15,17] As the specificity is very close to 100%, in the combination arms it was assumed that no false positives were possible after a urinary antigen test.

When tests are used in combination, they are not necessarily independent. For example, a positive test for S. *pneumoniae* is more likely with a pneumococcal urinary antigen test given a positive test pneumococcal blood culture. We explored this issue by reducing all sensitivities after the initial test by 100%, and increased all sensitivities after the initial test to 130% of the original sensitivity. Beyond this the model did not run as some probabilities summed to more than 1.

1.1.3.4 Resource use and costs

1.1.3.4.1 Antibiotics

Every patient receives empirical antibiotics at diagnosis for the first 24 hours. Dependent on the results of the microbiological test, the relevant targeted treatment is given as explained in section 1.1.2.2, which can involve IV only, oral only or IV and oral treatment.

No difference in length of antibiotic treatment was made for patients with correctly targeted treatment, incorrectly targeted treatment or those on empirical therapy. This is due to the low costs of oral antibiotics, the similarity between them across the regimens and to the fact that a difference in the cost of care is already covered by the extended length of stay in those cases where the treatment chosen based on the test results does not cover the pathogen present. Doses for moderate- and high-severity CAP were based on doses from the BNF.[8]

1.1.3.4.2 Microbiological test costs

Microbiological tests are listed in HRG codes within NHS reference costs, but costs for specific tests are not detailed. However, because this model evaluated the cost effectiveness of individual tests or specific combinations thereof, we had to estimate the cost of specific tests rather than the cost of the broad category. Costs were based on those from standard UK laboratories; these figures reported in Table 4 include not only the cost of the test itself but also staff time to take, prepare and interpret the tests. It was not possible to provide a breakdown of these costs as only a bundled cost was provided. The institutions that provided the costs asked to remain anonymous. While the cost of most tests was similar across laboratories, varying prices for the urinary antigen tests were reported by GDG members. The manufacturer of these kits confirmed that locally-negotiated discounts were not responsible for the difference in costs. An average price was therefore calculated and confirmed by the GDG.

The cost of the "all tests" strategy comprised the sum of all the individual tests as it was assumed that they would be performed simultaneously.

1.1.3.4.3 Hospital costs

All individuals in the model were admitted to hospital as this was the patient population prioritised by the GDG. The cost of the average LOS is not dependent on the strategy and every individual with moderate- and high-severity CAP would be admitted for a standard number of days, around 7 according to the BTS Audit (personal communication). The difference between strategies was additional bed days resulting from the incorrect treatment given because of the false positives and false negatives produced by each test strategy. For this reason, we assigned an additional cost of three days to patients in the model who were incorrectly treated (pathogen was resistant to the assigned treatment) and 7 days for patients who died (Mortensen et al, 2002).[12] The cost per excess bed day (£228) was taken from the NHS Reference cost, (code DZ11A, 'Non-elective long stay excess bed days - Lobar, Atypical or Viral Pneumonia, with Major CC'). The 'Major comorbidities and complications' code was chosen as this was likely to be a better proxy for moderate- and high-severity CAP.

1.1.3.5 Life expectancy and mortality

The average life expectancy for this population, aged 72, was calculated using interim life tables.[13] The result is detailed below in Table 6. In the model we assumed that the ratio male:female was 50:50.

Table 6: Life expectancy

	Male	Female
Life expectancy	13.054	15.190
Life expectancy (discounted)[a]	10.157	11.481
Average life expectancy (discounted)[b]	10.819	

(a) Discounted at 3.5% per annum
(b) Assuming equal male to female ratio

No systematic search was conducted for data on mortality in patients with CAP. The GDG advised of good studies on mortality. The study by Lim et al (2001)[11] with 237 patients, was selected as it was the most recent UK study on the mortality of patients with CAP, and was also used in the BTS CAP guidelines.[10] However these mortality data were collected across all severities of patients admitted to hospital with CAP (not specifically moderate- and high-severity CAP). The figures reported do not distinguish between those patients on targeted or empirical treatment. To adjust for this, the GDG used their expert knowledge to modify the mortality estimates for these two subgroups dependent on the treatment strategy. For the majority of pathogens, the mortality was considered likely to be to be the same in patients who were treated correctly or incorrectly because of the susceptibility of pathogens to empirical treatment. The GDG increased the mortality probability associated with Gram-negative pneumonia on non-targeted treatment due to the resistance of some Gram-negative pathogens to this strategy. The GDG reported clinical experience of improved survival with targeted treatment for *S. aureus* pneumonia and therefore decreased the mortality probability related to *S. aureus* for pneumonia treated with targeted antibiotic therapy. These variations made by the GDG were tested in a sensitivity analysis (see 1.2.2.5). Additionally, as the long-term impact on mortality from CAP is unknown, the assumption was made that there was no impact after the initial 30 days.

The expected probability of death due to pneumonia for the average patient in the model was 12.6%, which is given by the pathogen-specific mortality adjusted by the prevalence of the pathogen. This value is in line with the mortality in the moderate-severity group although is too low for the high-severity group. In fact, according to the definition of moderate- and high-risk severity as stratified by CURB65 score, see Table 7**,** the base case results should be considered applicable to the moderate-severity group only. A sensitivity analysis, detailed in section 1.1.5.1, will be undertaken to increase the mortality probabilities for the high-severity group.

Table 7: ITU Prevalence

CURB65 risk	Mortality
Low-risk group	2.1%
Moderate-risk group	10.3%
High-risk group	22.1%

1.1.3.6 Utilities

A systematic search was undertaken to identify quality-of-life data for CAP. Few data were available and where they were, these were for low-severity CAP or ventilator-associated pneumonia (VAP). The GDG felt that these data could not be extrapolated to patients with moderate- and high-severity CAP. The GDG advised that other medical conditions, such as severe sepsis and meningitis, were likely to have similar quality-of-life effects to moderate- and high-severity CAP and as such, these conditions were used as proxies. An additional systematic search was conducted to identify studies evaluating quality-of-life in severe sepsis. A systematic search for quality-of-life data in meningitis would have been undertaken had no applicable data for severe sepsis been identified. Two papers provided good short- and long-term quality-of-life data. The GDG decided that the utilities used in a recent HTA[18] were not appropriate as patients did not return to a quality-of-life close to that with which they started and the assumption made within this model was that patients return to 95% quality-of-life within six months. As such, these data was not applicable to patients with moderate- and high-severity CAP. Instead, a utility measure from Hofhuis et al (2008)[7] was used. This study reported SF-36 scores at ICU discharge, hospital discharge, three months after discharge and six months after discharge in people with sepsis. These scores were converted into EQ-5D scores using a mapping function and linearly adjusted by a factor of 0.948 to match the UK baseline as the baseline score used in Hofhuis was considerably higher than the UK average[9] as shown in Table 8.

Table 8: Utility scores

Time point	EQ-5D scores from Hofhuis et al (2008)[7]	Adjustment factor	EQ-5D scores used in the mode (EQ-5D scores from study X adjustment factor)
Baseline	0.8681	0.95	0.8250
Hospital discharge (EQ-5D score between hospital discharge and 30 days)	0.6978		0.6631
3 months after discharge (QoL between 30 days and 90 days)	0.8178		0.7772
6 months after discharge (QoL between 90 days and 180 days)	0.8371		0.7955

The QALY loss from CAP was calculated by working out the difference in EQ-5D score between baseline and the given the time period above which is then applied to the relevant number of days (e.g. the difference between EQ-5D between admission and discharge is given by the difference in EQ-5D divided by 365 days and multiplied by 10 days). These differences were then summed to calculate the total QALY loss associated with a correct and incorrect treatment as shown in **Table 9**.

Table 9: QALY loss due to correctly treated and incorrectly treated community-acquired pneumonia

QALY loss	Correct treatment	Incorrect treatment
Between admission to hospital and discharge(a)	0.0071	0.0102
Between hospital discharge and 30 days(b)	0.0102	0.0089
Between 30 days and 90 days	0.0079	0.0079
Between 90 days and 180 days	0.0073	0.0073
Total	0.0324	0.0342
Incremental (incorrect – correct treatment)		0.0017

(a) 10 days of disutility for incorrect treatment, 7 days of disutility for correct treatment
(b) 20 days for incorrect treatment, 23 days for correct treatment

1.1.4 Computations

The mean cost and effectiveness and the incremental cost effectiveness of the microbiological testing strategies were calculated using TreeAge Pro 2009.

1.1.4.1 Calculating costs

For each strategy, the expected cost is calculated as follows:

I **Expected cost = C_{test} + C_{ant} + (pChange*C_{ant2}) + (pIncorrectDiagnosis * C_{LOS}) + (pDeath * $C_{LOSDeath}$)**

where

C_{test} = cost of the initial strategy (tests conducted)

C_{ant} = cost of initial antibiotic strategy

pChange = probability of changing treatment strategy due to test result

C_{ant2} = cost of second antibiotic strategy

pIncorrectDiagnosis = probability of an incorrect diagnosis

C_{LOS} = cost of additional LOS

pDeath = probability of death within 30 days

$C_{LOSDeath}$ = cost of additional LOS for patients who eventually died

Costs are accrued only during the first 30 days and no discounting on costs was applied.

The incremental cost associated with a strategy is calculated as the difference between the expected cost with that strategy and the expected cost with the comparators.

1.1.4.2 Calculating QALYs

For each strategy, the expected QALYs are calculated as follows:

II Total discounted expected QALYs = $QALY_{lifetime}$ – (pCorrecttreatment * $QALY_{correcttreatment}$) – (pIncorrecttreatment * $QALY_{incorrecttreatment}$)

Where

$QALY_{lifetime}$ = QALYs accrued over lifetime of a patient (discounted)

pCorrecttreatment = probability that the treatment given is correct

$QALY_{correcttreatment}$ = QALY loss when treated correctly

pIncorrecttreatment = probability that the treatment given in incorrect

$QALY_{incorrecttreatment}$ = QALY loss when treated incorrectly

The incremental QALYs gained associated with a strategy are calculated as the difference between the expected QALYs with that strategy and the expected QALYs with the comparators.

1.1.4.3 Estimation of cost effectiveness

The widely used cost-effectiveness metric is the incremental cost-effectiveness ratio (ICER). This is calculated by dividing the difference in costs associated with two alternatives by the difference in QALYs. The decision rule then applied is that if the ICER falls below a given cost-per-QALY threshold the result is considered to be cost effective. If both costs are lower and QALYs are higher the option is said to dominate and an ICER is not calculated.

$$ICER = \frac{Costs(B) - Costs(A)}{QALYs(B) - QALYs(A)}$$ Where: Costs/QALYs(X) = total costs/QALYs for option X	• Cost-effective if: ICER < Threshold

When there are more than two comparators, as in this analysis, options must be ranked in order of increasing cost then options ruled out by dominance or extended dominance before calculating ICERs excluding these options. An option is said to be dominated, and ruled out, if another intervention is less costly and more effective. An option is said to be extendedly dominated if a combination of two other options would prove to be less costly and more effective.

It is also possible, for a particular cost-effectiveness threshold, to re-express cost-effectiveness results in term of net monetary benefit (NMB). This is calculated by multiplying the total QALYs for a comparator by the threshold cost-per-QALY value (for example, £20,000) and then subtracting the

total costs (formula below). The decision rule then applied is that the comparator with the highest NMB is the most cost-effective option at the specified threshold. That is the option that provides the highest number of QALYs at an acceptable cost.

$Net\ Benefit(X) = (QALYs\ (X) \times \lambda) - Costs\ (X)$ *Where: Costs/QALYs(X) = total costs/QALYs for option X; λ = threshold*	• Cost-effective if: highest net benefit

Both methods of determining cost effectiveness will identify exactly the same optimal strategy. For ease of computation NMB is used in this analysis.

Results are also presented graphically where total costs and total QALYs for each microbiological testing strategy are shown. Comparisons not ruled out by dominance or extended dominance are joined by a line on the graph where the slope represents the incremental cost-effectiveness ratio.

1.1.5 Sensitivity analyses

The GDG wished to explore whether any modification of important inputs and assumptions in the base case analysis would have an effect on the results. The following sensitivity analyses were conducted:

1.1.5.1 SA1: high-severity mortality

When constructing the model we focused on moderate- and high-severity CAP. However, the mortality estimates in the base case lead to a total mortality of between 13.0% and 12.1% in the empirical and targeted treatment groups respectively. As such, the base case analysis lends better to the moderate-severity than the high-severity group, given the mortality stratified by risk group in Table 7 from the severity assessment clinical review.

As such, a threshold analysis will be undertaken to increase the mortality probabilities linearly across all the pathogens to assess if the optimal strategy changes.

1.1.5.2 SA2: availability of sputum culture

A high proportion of patients with moderate- and high-severity CAP are unable to produce sputum when admitted to hospital. The GDG wanted to explore how the model results would change if this option was removed from the model. This was done by using a switch in the model to remove the cost of sputum culture and set all sensitivities of sputum culture to 0, and all specificities to 1. This then forced the model to ignore targeted treatment guided by sputum culture results.

1.1.5.3 SA3: prevalence of pathogens

The GDG wished to know how the prevalence of pathogens would impact on the model results, because there was a high probability of either having an atypical pathogen or *S. pneumoniae* in the base case. As such, the prevalence from the ICU reported in the BTS CAP Guidelines[10] was used instead for very high severity.

Table 10: ITU Prevalence

Parameter description	Point estimate	Probability distribution	Distribution parameters	Source
S. pneumonia[a]	NA	NA	NA	NA
H. influenza	0.06762	Beta	α = 7, β = 97	Calculated using prevalence reported in

Parameter description	Point estimate	Probability distribution	Distribution parameters	Source
				BTS Guidelines 2009[10]
S. aureus	0.15480	Beta	α = 16, β = 88	Calculated using prevalence reported in BTS Guidelines 2009[10]
L. pneumophila	0.31673	Beta	α = 33, β = 71	Calculated using prevalence reported in BTS Guidelines 2009[10]
Atypical pathogens	0.04804	Beta	α = 5, β = 99	Calculated using prevalence reported in BTS Guidelines 2009[10]
Gram-negative pathogens	0.02847	Beta	α = 3, β = 101	Calculated using prevalence reported in BTS Guidelines 2009[10]

(a) Within the model the point estimate for S. pneumoniae is calculated as the residual of the other pathogens probabilities.

1.1.5.4 SA4: quality of life

The GDG wished to know how returning to full quality-of-life after pneumonia would affect the results. As such, a sensitivity analysis was conducted. To do this, we used the full average UK quality-of-life, instead of weighting it by 0.95 as we did in the base case.

1.1.5.5 SA5: mortality probability of pathogens

Due to the GDG modifying the mortality estimates, a one-way sensitivity analysis was conducted on the mortality probabilities of Gram-negative pathogens treated with empirical treatment and *S. aureus* with targeted treatment. This was to minimise the uncertainty around these estimates.

1.1.5.6 SA6: reduction in test sensitivities

The GDG were concerned that through the assumption of increasing the prevalence, the efficacy of the tests could be over-estimated as the sensitivities would have been obtained given a lower prevalence. To ensure the rigour of the base case results, we undertook a threshold analysis to assess the reduction needed in sensitivities that would change the results. The reduction in sensitivity was assumed to be the same across all pathogens, with reduced test sensitivity being calculated as follows:

$$sensitivity\ of\ test\ to\ pathogen\ (X) \times factor\ (Y) = reduced\ test\ sensitivity$$

It was decided to assess this impact across both the base case prevalence the ICU prevalence used in SA2.

1.1.5.7 SA7: quality-of-life gain from targeted treatment

The model was not designed to include the benefits of targeted treatment, such as lower rates of antimicrobial resistance and minimised adverse events. The GDG were interested in investigating the amount of QALY gain that a targeted treatment would require in order to alter results and as such a threshold analysis was undertaken. The GDG felt that this would be useful to guide their decision, as extra tests may be warranted if the QALY gain required to make them cost effective was not improbable.

1.1.6 Model validation

The model was developed in consultation with the GDG; model structure, inputs and results were presented to and discussed with the GDG for clinical validation and interpretation.

The model was systematically checked by the health economist undertaking the analysis; this included inputting null and extreme values and checking that results were plausible given inputs. The model was peer reviewed by a second experienced health economist from the NCGC; this included systematic checking of the model calculations.

1.1.7 Interpreting Results

NICE's report 'Social value judgements: principles for the development of NICE guidance' sets out the principles that GDGs should consider when judging whether an intervention offers good value for money. In general, an intervention was considered to be cost effective if either of the following criteria applied (given that the estimate was considered plausible):

- The intervention dominated other relevant strategies (that is, it was both less costly in terms of resource use and more clinically effective compared with all the other relevant alternative strategies), or
- The intervention costs less than £20,000 per quality-adjusted life-year (QALY) gained compared with the next best strategy.

In this analysis, as there were several interventions, the NMB was used to rank the strategies on the basis of their relative cost effectiveness. The highest NMB identified the optimal strategy at a willingness to pay of £20,000 per QALY gained.

The strategy with the highest net benefit is the one that should be recommended. However, since we were unable to capture the incidence, cost or disutilities of treatment-specific adverse events, and other issues such as antibiotic resistance, caution should be exercised in recommending microbiological testing strategies which lead to more inappropriate treatments. It should also be noted that this economic analysis applied to patients with moderate- and high-severity CAP only.

1.2 Results

1.2.1 Base case

In the base case, model inputs were set as shown in Table 4 and the model was run both deterministically and probabilistically.

Table 11: Base case (moderate-severity CAP) – probabilistic results

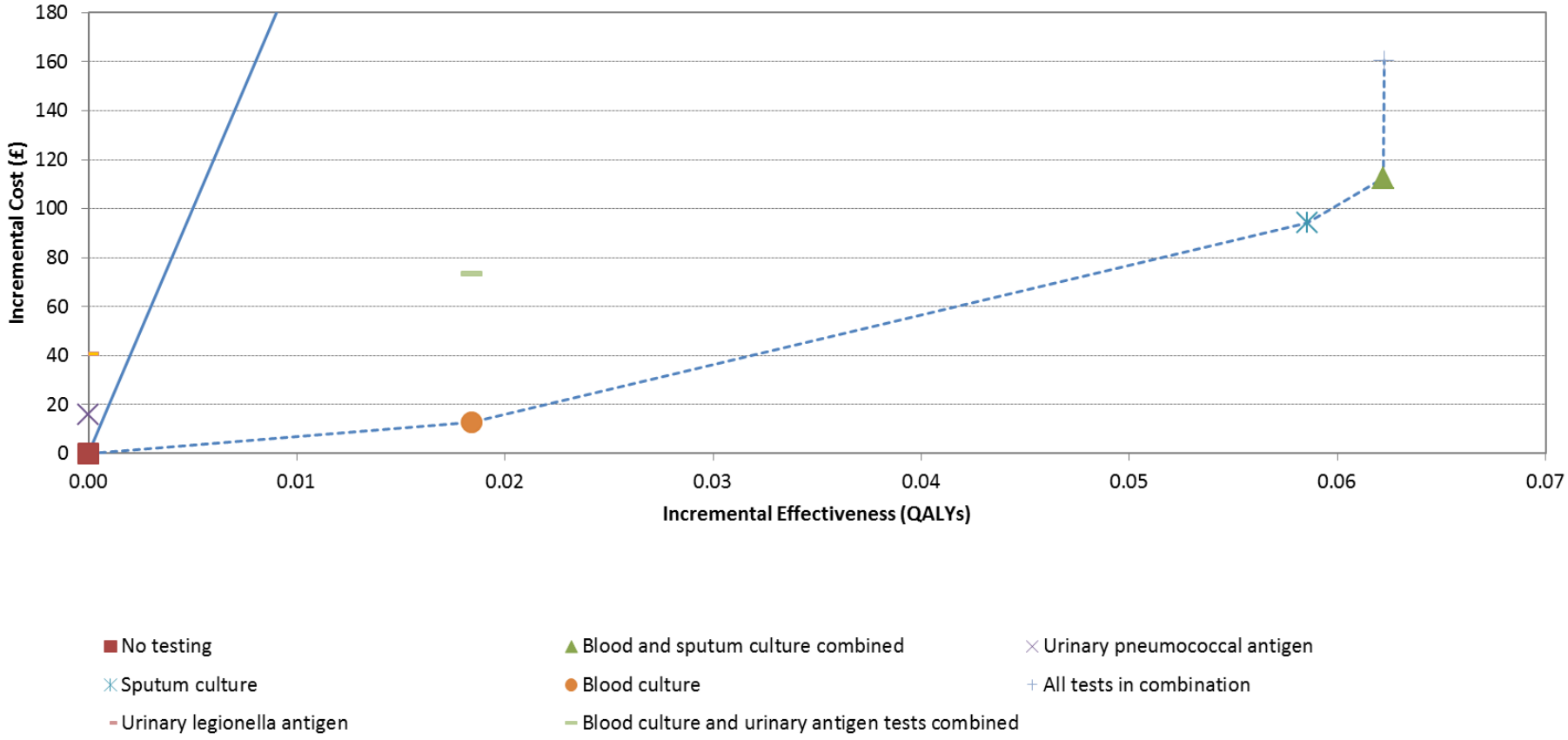
Strategy	Cost (£)	QALYs	NMB at £20k/ QALY	Rank(a)	Probability optimal strategy(b)
Blood culture and sputum culture	£2,683	7.4103	145,524	1	58%
All tests	£2,731	7.4103	145,475	2	5%
Sputum culture	£2,664	7.4066	145,468	3	18%
Blood culture	£2,582	7.3670	144,758	4	3%
Blood culture and urinary antigen tests	£2,642	7.3670	144,698	5	0%
No testing	£2,570	7.3488	144,406	6	15%
Urinary pneumococcal antigen	£2,589	7.3488	144,387	7	2%
Urinary legionella antigen	£2,610	7.3488	144,366	8	0%

(a) Ranked by average NMB (£20,000 per QALY threshold)
(b) Percentage of simulations in which the microbiological testing strategy was the optimal strategy

Table 11 shows that a blood culture in combination with a sputum culture was the optimal microbiological testing strategy, as it was associated with the highest average net monetary benefit in 58% of the simulations.

The cost-effectiveness plane in Figure 1 provides a visual demonstration of the cost-effectiveness of the compared strategies. The strategies to the right of the £20,000 per QALY threshold (the blue solid line) were the strategies with positive incremental NMB compared to no testing and were therefore more cost effective. Those strategies to the left of the £20,000 per QALY threshold were not cost effective compared to no testing (urinary pneumococcal antigen and urinary legionella antigen) and have a negative incremental NMB. However, given the main benefits of these tests for targeting treatment were not included in the base case, this was to be expected. To establish which of the microbiological testing strategies with positive incremental NMB is optimal, we can look at Figure 1. The line depicting the ICER between blood culture and no testing is less steep than the cost-effectiveness threshold. Blood culture and urinary antigen tests combined is dominated (more costly and no more effective) by blood culture and the all tests in combination strategy is dominated (more costly and equally effective) by blood and sputum culture combined. The line depicting the ICER between sputum culture and blood culture is also less steep than the cost-effectiveness threshold, indicating sputum culture is cost effective compared to blood culture. We can then consider a blood culture in combination with a sputum culture against sputum culture alone, which is less steep than the cost-effectiveness threshold. As such, a blood culture in combination with a sputum culture is the most cost-effective microbiological testing strategy in the base case analysis. The results were similar when the model was run deterministically.

Figure 1: Cost-effectiveness plane (base case analysis)



1.2.2 Sensitivity analyses

1.2.2.1 SA1: high-severity mortality

As described in section 1.1.5, a threshold analysis was undertaken to assess the impact of an increase in all pathogen mortality probabilities up to double. This factor was applied linearly across all pathogens simultaneously. We were unable to linearly increase mortality beyond double as mortality becomes a certainty with *S. aureus* treated empirically. Table 12 shows the actual values used in the upper limit of the threshold analysis.

Table 12: Pathogen mortality probabilities doubled

Pathogen	Empirical mortality	Targeted mortality
S. pneumonia	0.28	0.28
H. influenza	0.10	0.10
S. aureus	1.00	0.60
L. pneumophila	0.22	0.22
Atypical pathogens	0.10	0.10
Gram-negative pathogens	0.80	0.50

An increase in mortality probabilities by up to double leads to no change in the optimal strategy, a blood culture in combination with a sputum culture. This results in a total mortality of between 26.0% and 24.3% in the empirical and targeted treatment groups respectively.

As we realised the model may have underestimated any possible benefit from targeted treatment for pathogens other than S. aureus and Gram-negative, we decreased the mortality with targeted treatment for every pathogen. When the current targeted treatment mortality estimates were multiplied by a 0.99 factor (a decrease of 1% of their current values), all tests in combination was the most cost effective strategy.

1.2.2.2 SA2: availability of sputum culture

As described in section 1.1.5 sputum culture was removed from the model. The model was run both deterministically and probabilistically.

Table 13: Sputum not available – probabilistic results

Strategy	Cost (£)	QALYs	NMB	Rank[a]	Probability optimal strategy[b]
Blood culture	£2,587	7.367	£144,753	1	83%
Blood culture and urinary antigen tests	£2,648	7.367	£144,692	2	0%
All tests	£2,650	7.367	£144,690	3	0%
No testing	£2,574	7.349	£144,406	4	14%
Urinary pneumococcal antigen	£2,590	7.349	£144,390	5	2%
Urinary legionella antigen	£2,615	7.349	£144,365	6	0%

(a) Ranked by average NMB (£20k per QALY)
(b) Percentage of simulations in which the microbiological testing strategy was the optimal strategy

Table 13 demonstrates that in this scenario, the most cost-effective microbiological testing strategy is a blood culture alone in 83% of the simulations. In the probabilistic analysis, blood culture alone had an ICER of £705 per QALY gained when compared with the next best alternative, no testing. The results were similar when the model was run deterministically.

1.2.2.3 SA3: prevalence of pathogens

As described in section 1.1.5 the prevalence of pathogens in the ICU was used instead of pathogens in the general hospital setting. The model was run both deterministically and probabilistically.

Table 14: ITU prevalence – probabilistic results

Strategy	Cost (£)	QALYs	NMB	Rank[a]	Probability optimal strategy[b]
All tests	£2,768	7.147	£140,172	1	46%
Blood culture and sputum culture	£2,789	7.147	£140,151	2	31%
Sputum culture	£2,772	7.132	£139,868	3	3%
Blood culture	£2,673	6.972	£136,767	4	1%
Blood culture and urinary antigen tests	£2,745	6.972	£136,695	5	0%
No testing	£2,673	6.898	£135,287	6	18%
Urinary pneumococcal antigen	£2,699	6.898	£135,261	7	0%
Urinary legionella antigen	£2,717	6.898	£135,243	8	0%

(a) Ranked by average NMB (£20K per QALY threshold)
(b) Percentage of simulations in which the microbiological testing strategy was the optimal strategy

Table 14 demonstrates that in this scenario, the most cost-effective microbiological testing strategy is all tests in combination in 46% of the simulations. In the probabilistic analysis, the all tests in combination strategy had an ICER of £538 per QALY gained when compared with the next best alternative, a blood culture in combination with a sputum culture. The results were similar when the model was run deterministically. However, there is some uncertainty associated with this effect as a blood culture in combination with a sputum culture was the most cost-effective testing strategy in 31% of the simulations.

1.2.2.4 SA4: quality-of-life

As described in section 1.1.5, the average UK quality-of-life figure was used, instead of 95% of this. The model was run both deterministically and probabilistically.

Table 15: Lifetime QoL – probabilistic results

Strategy	Cost (£)	QALYs	NMB	Rank[a]	Probability optimal strategy[b]
Blood culture and sputum culture	£2,691	7.803	£153,369	1	58%
All tests	£2,739	7.803	£153,321	2	6%
Sputum culture	£2,673	7.799	£153,307	3	17%
Blood culture	£2,591	7.757	£152,549	4	3%
Blood culture and urinary antigen tests	£2,652	7.757	£152,488	5	0%
No testing	£2,578	7.738	£152,182	6	14%

Strategy	Cost (£)	QALYs	NMB	Rank[a]	Probability optimal strategy[b]
Urinary pneumococcal antigen	£2,594	7.738	£152,166	7	3%
Urinary legionella antigen	£2,619	7.738	£152,141	8	0%

(a) Ranked by average NMB (£20K per QALY threshold)
(b) Percentage of simulations in which the microbiological testing strategy was the optimal strategy

Table 15 demonstrates that in this scenario, the most cost-effective microbiological testing strategy was a blood culture in combination with a sputum culture in 58% of the simulations. In the probabilistic analysis, the all tests in combination strategy was the most clinically-effective strategy, but was not cost effective (ICER: £1,384,075 per QALY gained) when compared with a blood culture in combination with a sputum culture. A blood culture in combination with a sputum culture had an ICER of £4723 per QALY gained when compared with the next best alternative, a sputum culture alone. The results were similar when the model was run deterministically.

1.2.2.5 SA5: mortality probability of pathogens

As described in section 1.1.5 mortality probabilities were varied in a one-way sensitivity analysis to see if changing the mortality probability would alter the model results. The range of mortality probabilities to explore in this sensitivity analysis was suggested by the GDG.

Table 16: Mortality probabilities

Mortality probabilities	Optimal strategy[(a)]
0.25	Blood culture and sputum culture
0.30	Blood culture and sputum culture
0.35	Blood culture and sputum culture
0.40	Blood culture and sputum culture
0.45	Blood culture and sputum culture
0.50	Blood culture and sputum culture

(a) Ranked by average NMB (£20K per QALY threshold)

Table 16 demonstrates that at all mortality probabilities explored, a blood culture in combination with a sputum culture was the optimal microbiological testing strategy. With this range of mortality the ICER for a blood culture in combination with a sputum culture compared with sputum culture alone ranged from £4,190 to £18,163 per QALY gained.

1.2.2.6 SA6: reduction in test sensitivities

As described in section 1.1.5, a threshold analysis was undertaken to assess the impact of a reduction on test sensitivity rates. This factor was applied to all sensitivities simultaneously.

Table 17: Reduction in test sensitivities using base case prevalence

Reduction	Optimal strategy[(a)]
Up to 88%	Blood culture and sputum culture
Between 88% and 94%	Blood culture
Above 94%	No testing

(a) Ranked by average NMB (£20K per QALY threshold)

Table 17 demonstrates that until sensitivities of tests are reduced by 88% of their base case sensitivity, a blood culture and a sputum culture was the most cost-effective strategy. As such, it is highly unlikely that this result would change.

Table 18: Reduction in test sensitivities using ITU prevalence

Reduction	Optimal strategy[(a)]
Up to 25%	All tests in combination
Between 25% and 96%	Blood culture and sputum culture
Between 96% and 98%	Blood culture
Above 98%	No testing

(a) Ranked by average NMB (£20K per QALY threshold)

Table 18 demonstrates that until sensitivities of tests are reduced by 25% of their base case sensitivity, all tests in combination were the most cost-effective strategy. Between this reduction and a reduction of 96% of their original sensitivity, a blood culture and sputum culture was the most cost-effective strategy. If they were reduced below this, blood culture was the most cost-effective strategy until the sensitivities were reduced by 98 % of their base case sensitivity when no testing became the most cost-effective strategy. It can be argued that there is more uncertainty as to whether all tests in combination would be an optimal strategy.

1.2.2.7 SA7: quality-of-life gain from targeted treatment

As described in section 1.1.5, a quality-of-life gain was added for those strategies that led to targeted treatment being given. This threshold analysis demonstrated when the benefits of targeted treatment would change the results.

Table 19: Targeted treatment QALY gain

QALY gain	Optimal strategy[a]
0 to 0.0134	Blood culture and sputum culture
above 0.0134	All tests in combination

(a) Ranked by average NMB (£20K per QALY threshold)

Table 19 demonstrates that if there was a QALY gain from targeted treatment of less than 0.0134 over the lifetime of a patient, blood culture and sputum culture remained the most cost-effective strategy. If targeted treatment was able to provide a QALY gain of more than 0.0134, all tests in combination would have been the most cost-effective strategy.

1.3 Discussion

1.3.1 Summary of results

In the base case (moderate-severity CAP), the most cost-effective microbiological testing strategy was to perform a blood culture and a sputum culture. This remained the same when all mortality probabilities were doubled (to account for the high-severity CAP), quality-of-life returned to pre-pneumonia levels and a range of specific pathogen mortality probabilities were used.

However, in those patients where sputum was not available, the most cost-effective strategy was blood culture alone and when ICU prevalence was used, the most cost-effective strategy was all tests in combination.

When base case test sensitivities were reduced by more than 88%, a blood culture replaced the combination of a blood culture and a sputum culture as the most optimal strategy. When ICU prevalence of pathogens was used, sensitivities of tests needed to be only reduced by 25% in order for the combination of a blood culture and a sputum culture to replace all tests in combination as the optimal strategy.

If there would be a QALY gain from targeted treatment it needed to be above 0.0134 QALYs before all tests in combination would become the cost-effective strategy compared to a blood culture and a sputum culture.

1.3.2 Limitations and interpretation

As has already been mentioned, due to the lack of evidence, and pragmatic constraints relating to model complexity, a number of assumptions were made to facilitate this model, with both the data inputs and the model structure. A considerable number of inputs within this model used data that is either an assumption by the GDG, indirect evidence, or with little evidence from good randomised controlled trials. This data limitation does cause uncertainty around the model results, yet the probabilistic nature of the model and the sensitivity analyses undertaken ensures that this risk is minimised.

A key assumption that may not translate to clinical practice is that this model assumed that patients only had a single causative pathogen. Moderate- and high-severity CAP can be caused by multiple pathogens and it is possible that it may be more acceptable to undertake additional tests to identify the rarer pathogens in this scenario. Further to this, with 30% of cases having unidentified aetiology, the true prevalence of these pathogens may be different to that within the model.

This model also assumed that there was no treatment failure and that there were no adverse events, which would be likely to impact both the cost of some strategies and their QALYs gained. However, it was considered that estimating the incidence of treatment failure and adverse events would have introduced too many unnecessary complications given the relatively limited impact of these effects compared to mortality.

In addition, there is no accepted method of estimating a cost for the advantages of antimicrobial stewardship. Reducing the need for inappropriate antibiotics may lead to long-term economic benefits, on both an individual and societal level, through the use of lower cost antibiotics and the continued ability to use basic antibiotics for common conditions. With the development of new antibiotics slowing, this is a key issue, both in terms of costs and quality of life.

The evidence on quality-of-life reductions from severe CAP is extremely limited. Using severe sepsis as a proxy does have limitations. This may either under- or over-value the true quality-of-life reductions associated with moderate- and high-severity CAP and ineffective treatment.

Further to this, the model was unable to capture the fact that Legionnaires' disease became a notifiable disease in early 2010 in England. For those with high-severity CAP, Legionella urinary antigen tests should still be considered for surveillance reasons.

The model may have not fully captured the benefits of urinary pneumococcal and legionella antigen tests as these pathogens are susceptible to empirical treatment and no decrease in mortality was assumed with targeted treatment for these two pathogens. The health benefit of all tests in combination is therefore likely to be underestimated by the model, as the paper by Uematsu et al (2014)[19] included in the clinical review shows - a lower mortality is evident in the all tests strategy,

while in our model there is no QALY gain in conducting urinary antigen tests in addition to blood and sputum culture tests.

Overall, this model is likely to provide an acceptable assessment of this area and until better data exist, it is unlikely that further uncertainty can be reduced.

1.3.3 Generalisability to other populations/settings

All of these findings relate to an adult population with confirmed moderate- and high-severity CAP in hospital. These results should not be used to inform decisions for patients with unconfirmed pneumonia or low-severity CAP.

1.3.4 Comparisons with published studies

Three published studies were identified in the literature review. Oosterheert[14] assessed the cost savings of targeting antimicrobial therapy in patients with severe pneumonia. The authors concluded that Gram staining and urinary antigen tests to detect *S. pneumoniae* provided no cost savings. Falguera et al[4] described the cost of targeted and empirical treatment arms, when testing for *S. pneumoniae* and *L. pneumophila* using urinary antigen tests, concluding that targeted treatment was more expensive with no additional benefits. Sinclair et al[16] also concluded that urinary pneumococcal antigen test was more expensive, without any benefits.

However, the studies above did not take into account quality-of-life in their analyses and did not adopt a lifetime time horizon.

1.3.5 Conclusion/evidence statement

A blood culture in combination with a sputum culture is the optimal microbiological testing strategy for patients with confirmed moderate- and high-severity CAP, managed in a hospital setting.

When patients are unable to produce sputum, blood culture alone is the optimal strategy.

If the prevalence of pathogens is closer to those observed in the ICU, all tests in combination is the optimal strategy.

Our analysis advocates that there needs to be a relatively modest QALY gain from targeted treatment in order for all tests in combination to be the optimal strategy.

1.3.6 Implications for future research

Within this model there are a number of limitations with the data and assumptions had to be made to fill the evidence gaps as has been explained above. It would be useful to rerun this model with up-to-date evidence including UK based moderate- and high-severity CAP quality-of-life data, recent UK pathogen-prevalence data for this population and mortality of patients with targeted and non-targeted treatments within the UK. A further extension to this piece of work may be to include the possibility of multiple pathogens causing moderate- and high-severity CAP or to include treatment failure and adverse events as outcomes to assess how this impacts the cost effectiveness of these microbiological tests.

1.4 Model Structure

The following section provides the model structure. Due to its size, it is broken down into separate arms of the model.

Figure 2: Model population and microbiological testing strategies

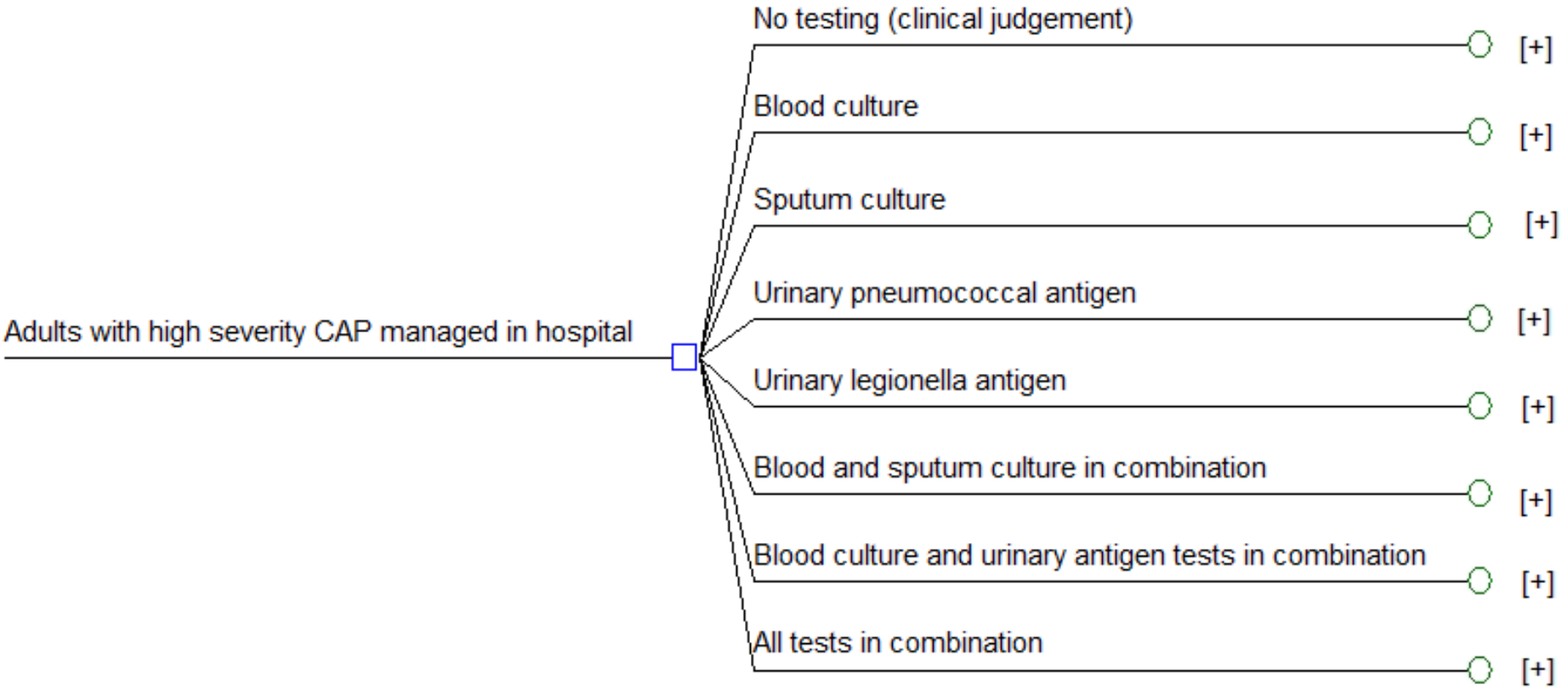


Figure 3: No testing (clinical judgement) arm

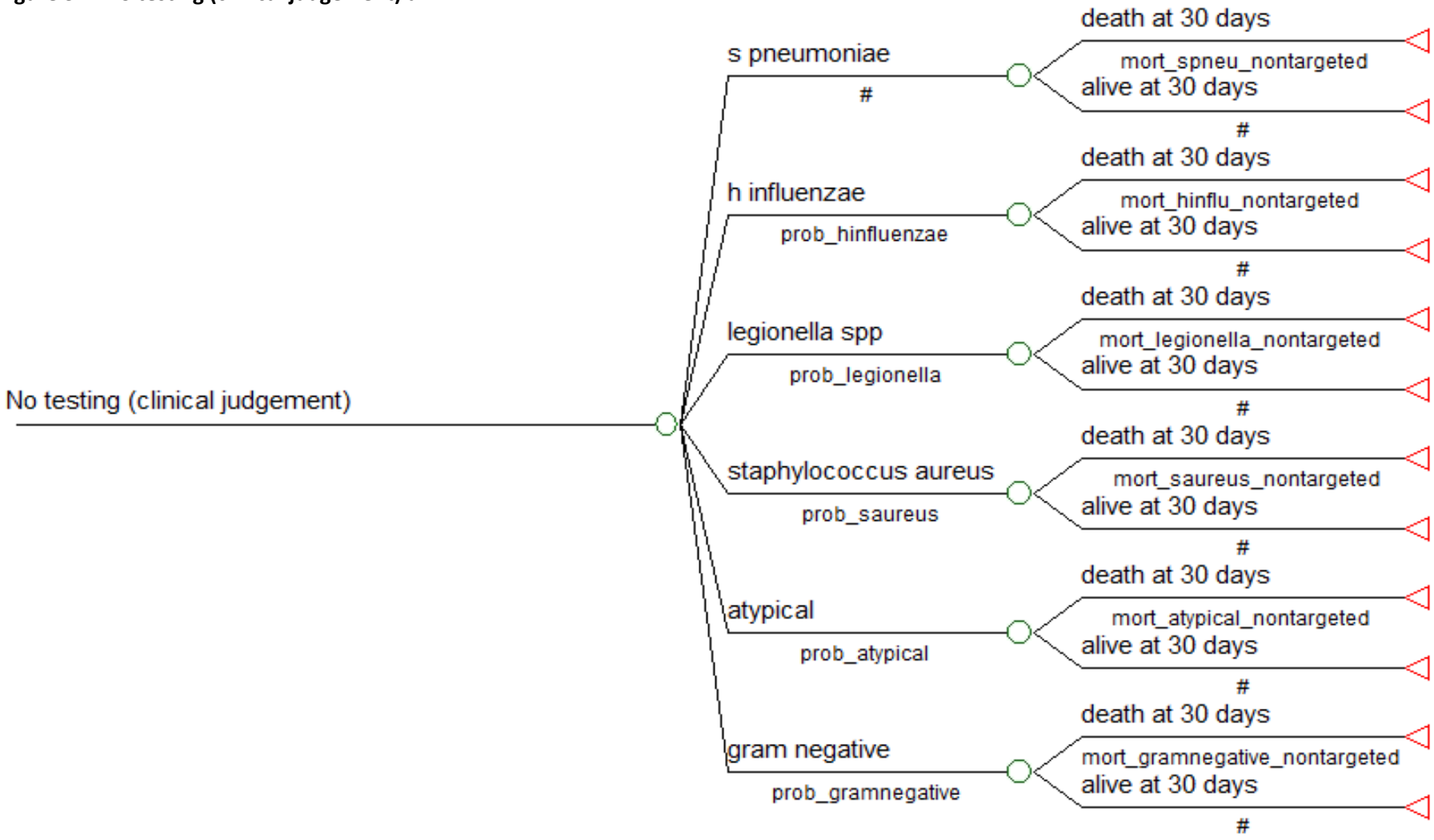


Figure 4: Blood culture arm

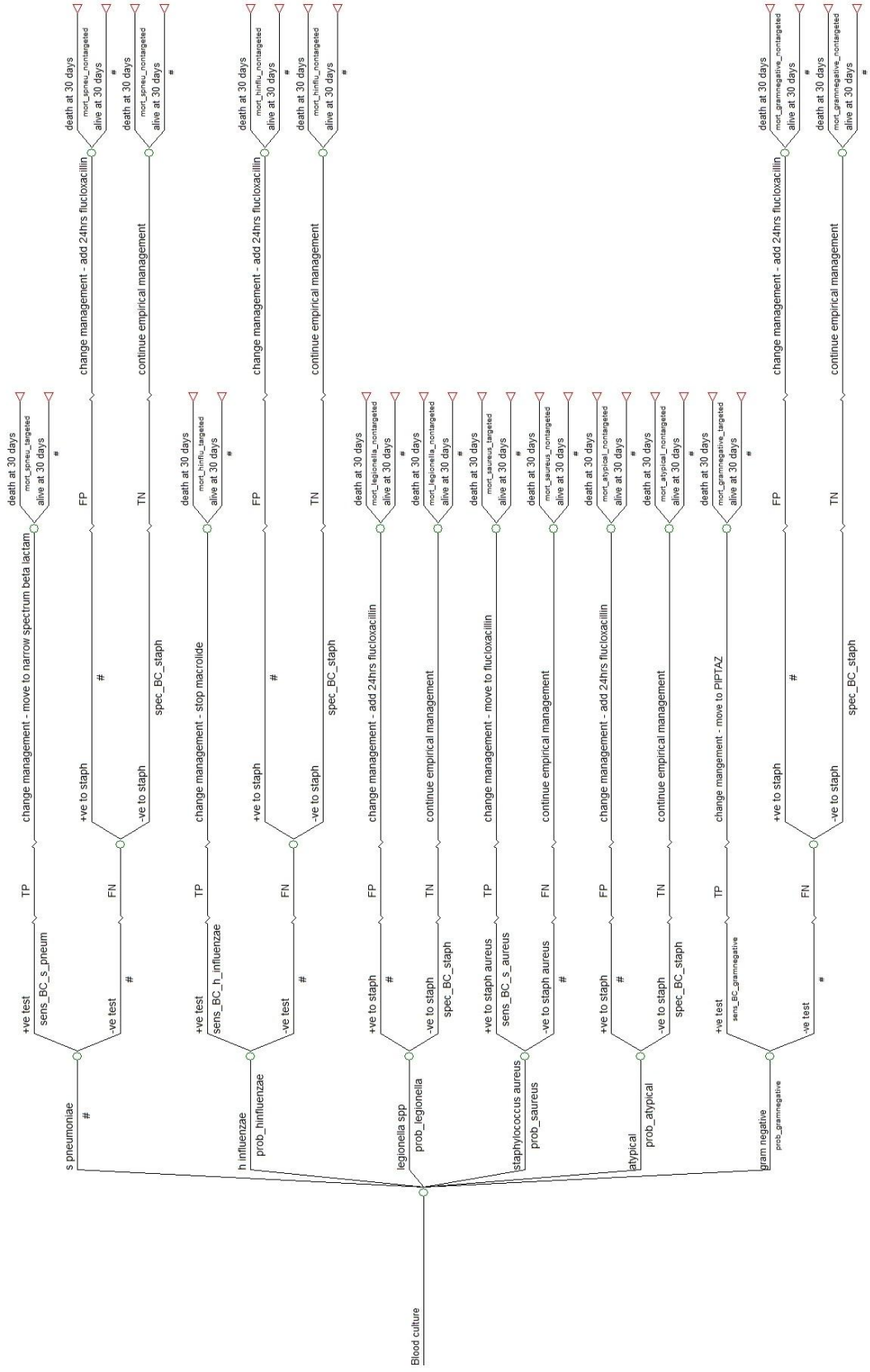


Figure 5: Routine sputum culture arm (part 1)

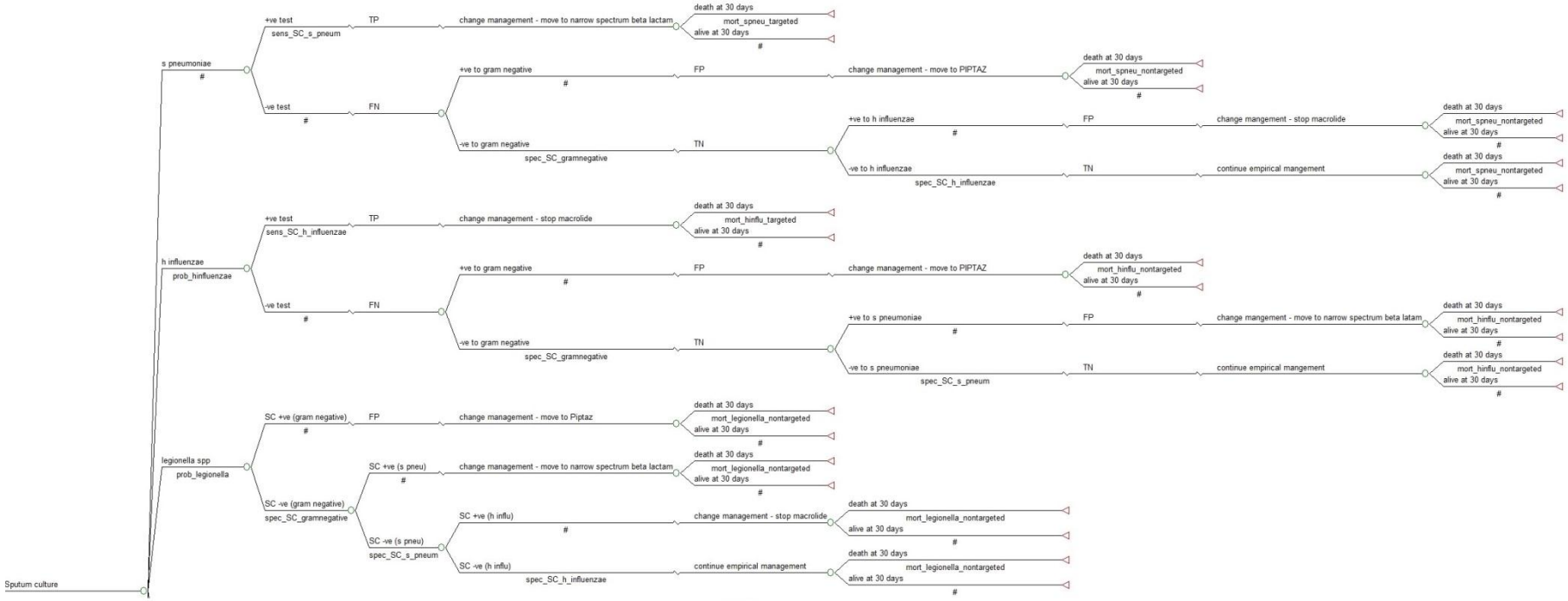


Figure 6: Routine sputum culture arm (part 2)

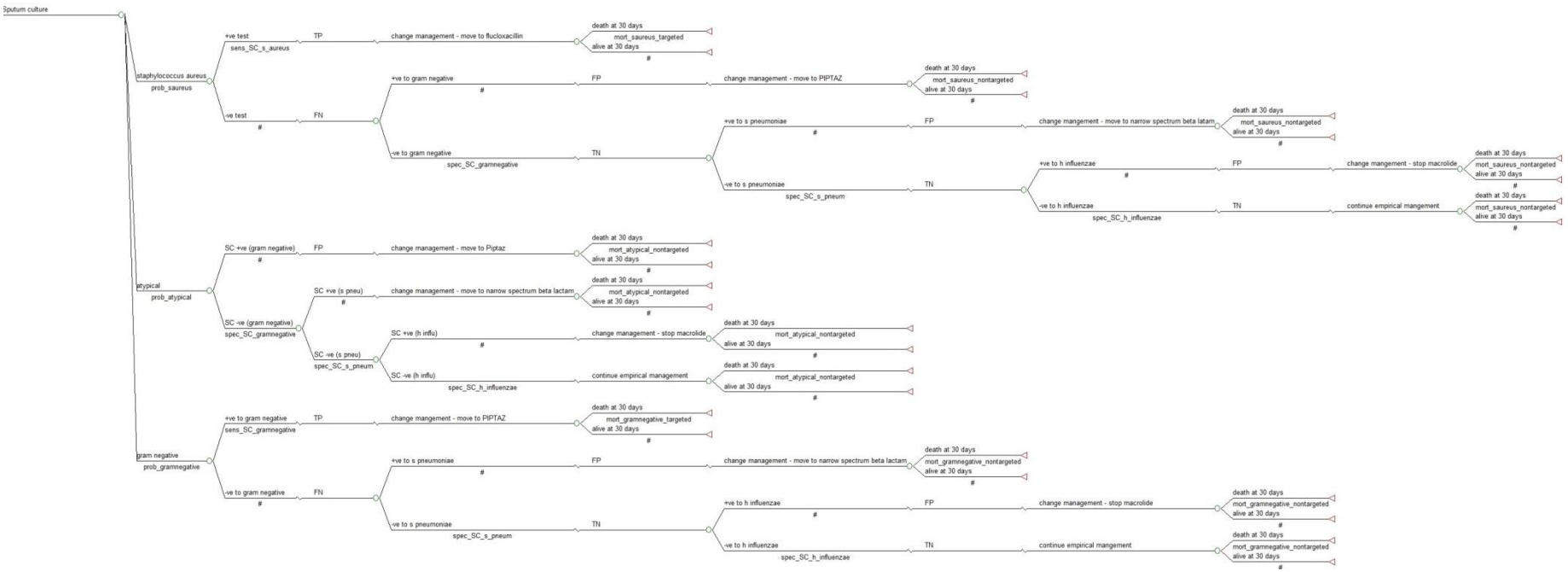


Figure 7: Urinary pneumococcal antigen arm

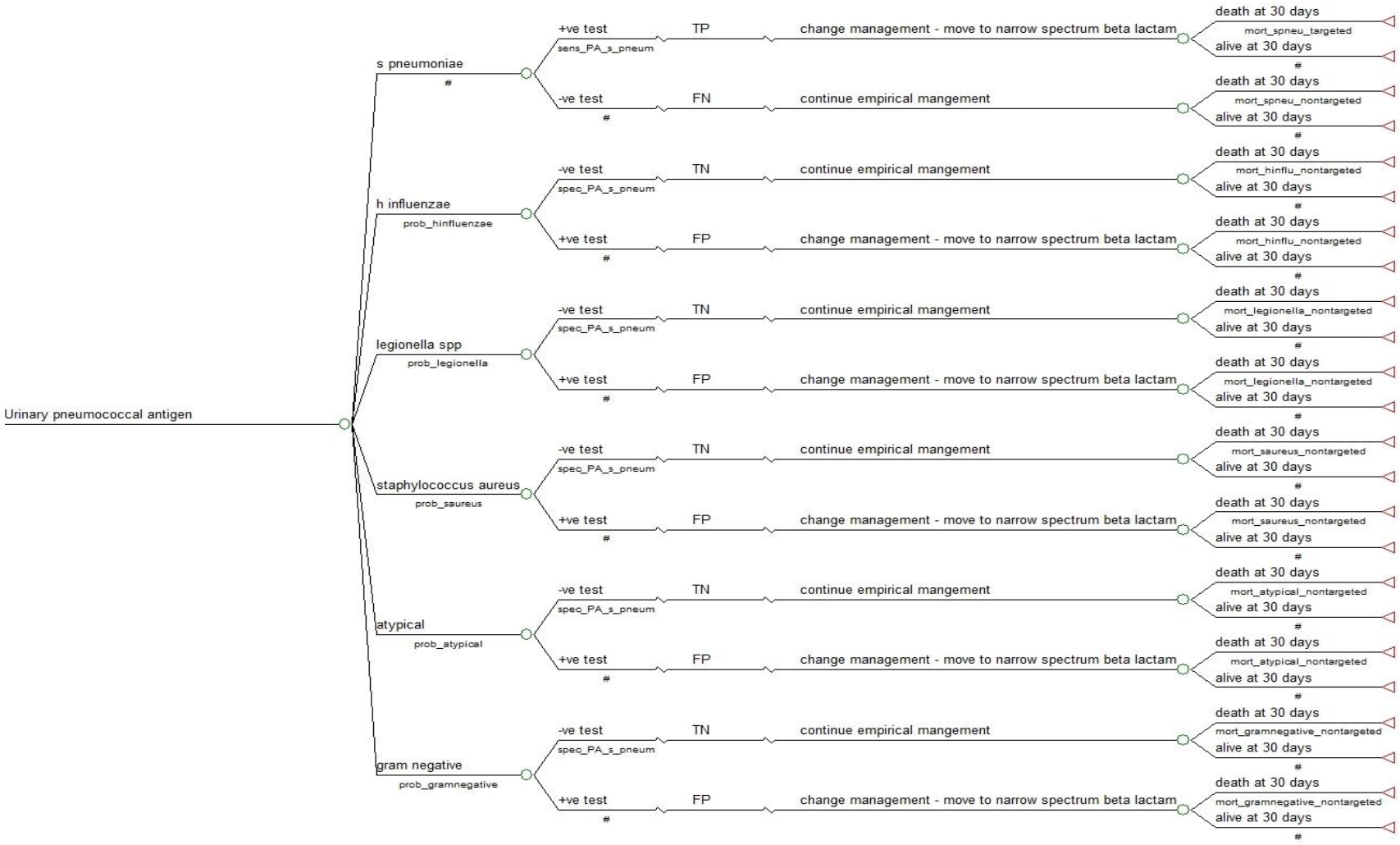


Figure 8: Urinary legionella antigen arm

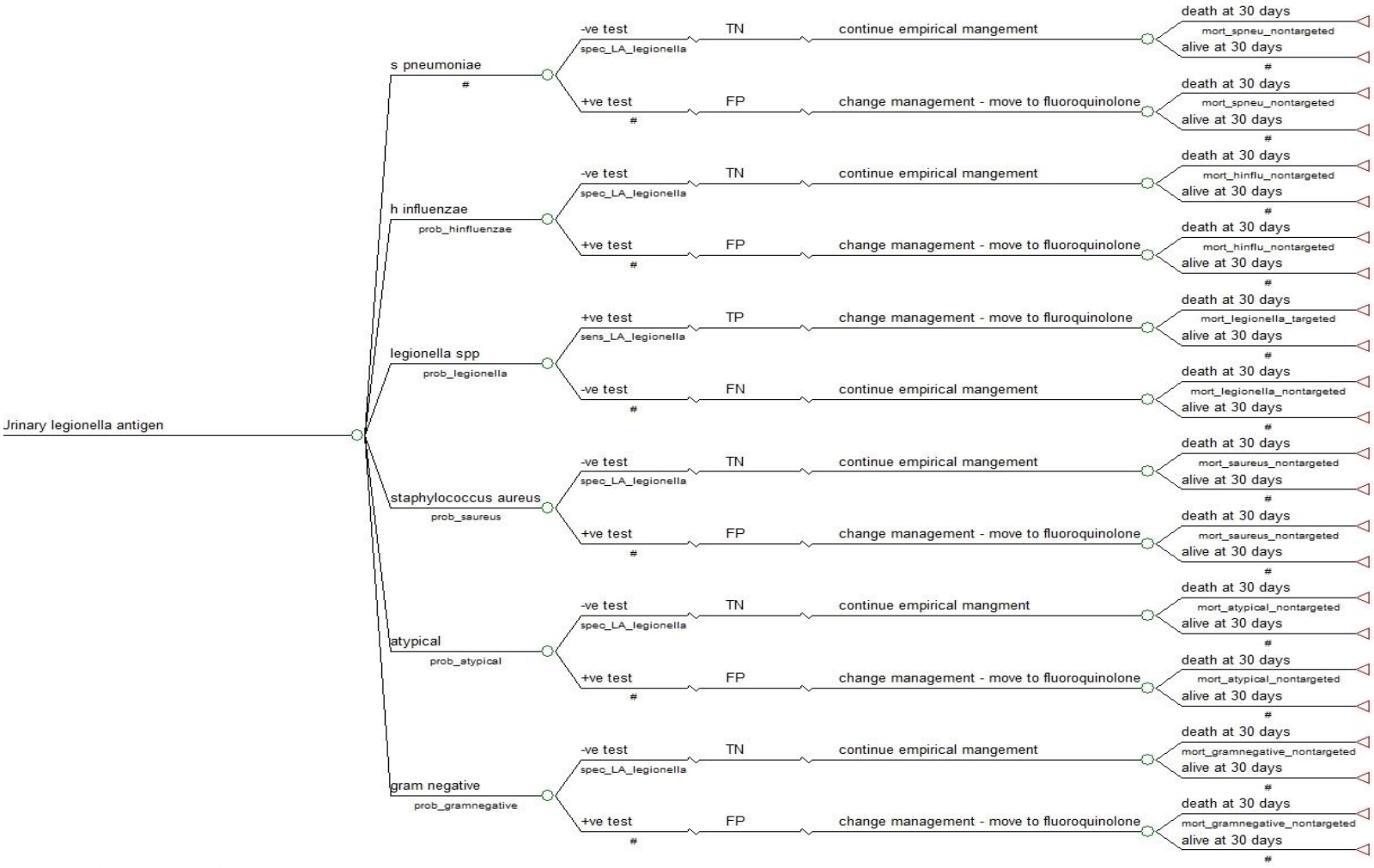
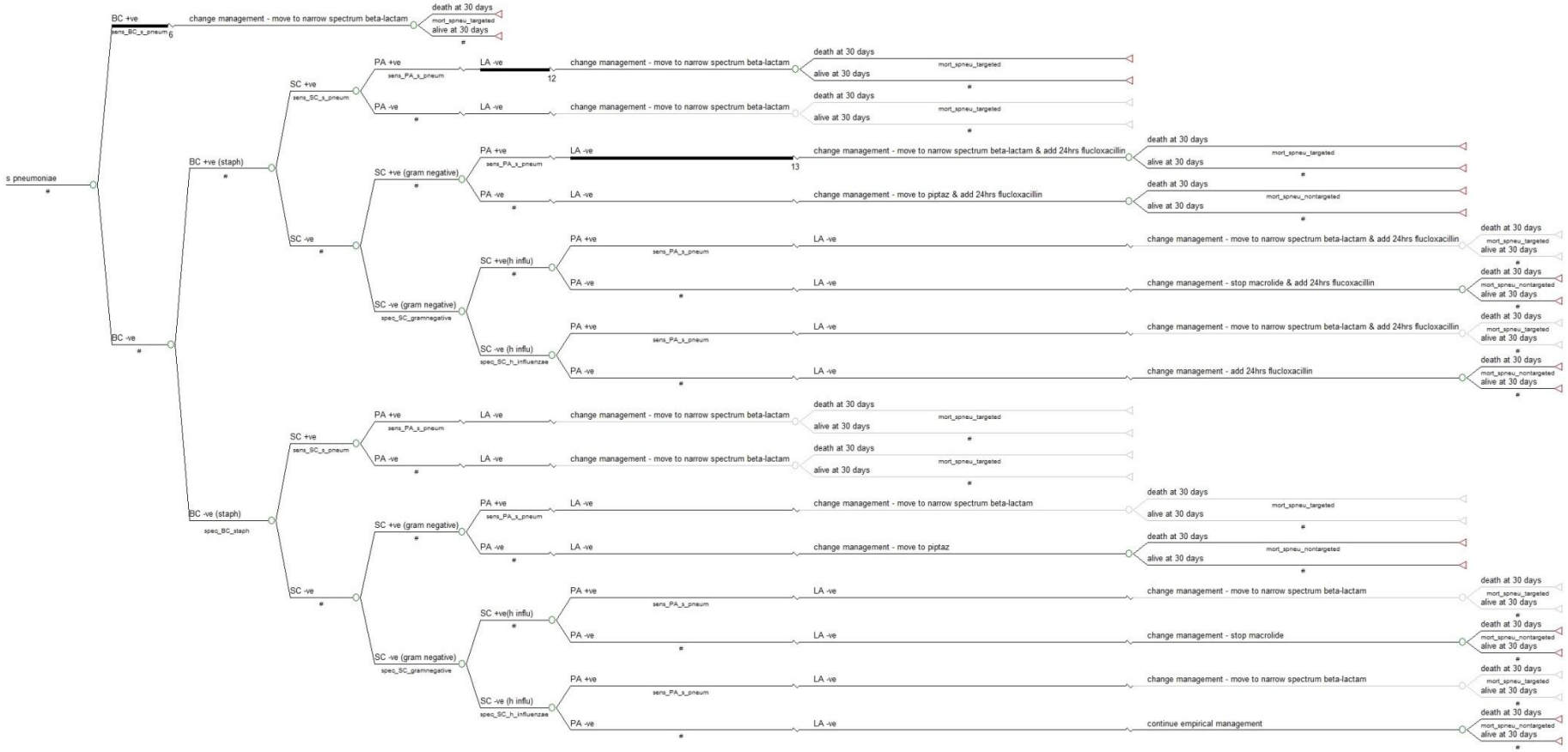


Figure 9: All tests in combination – *S. pneumoniae* arm



Note: The H. influenzae and Gram-negative arms follow the same structure

Figure 10: All tests in combination – *L. pneumophila* arm

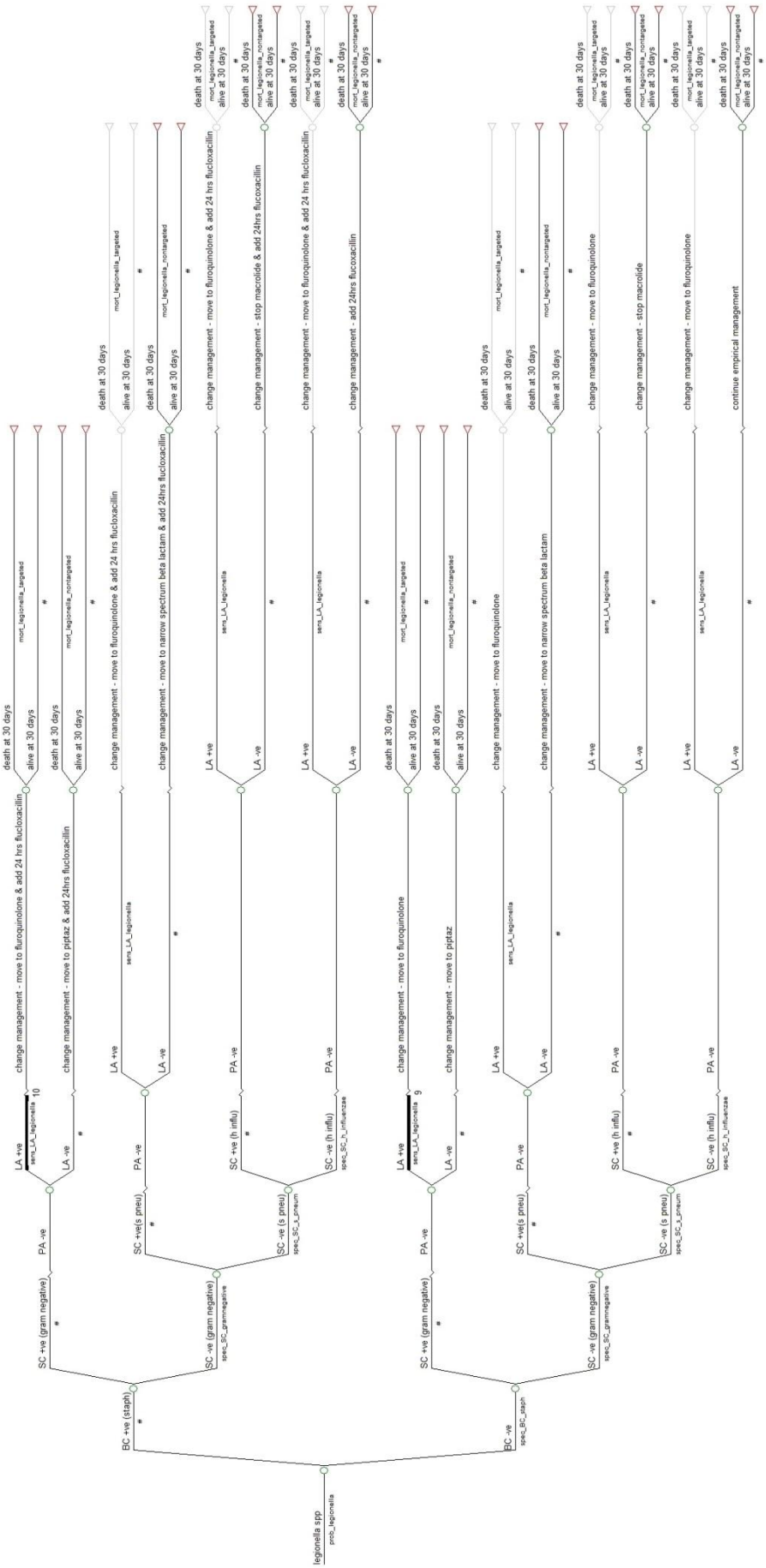


Figure 11: All tests in combination – *S. aureus* arm

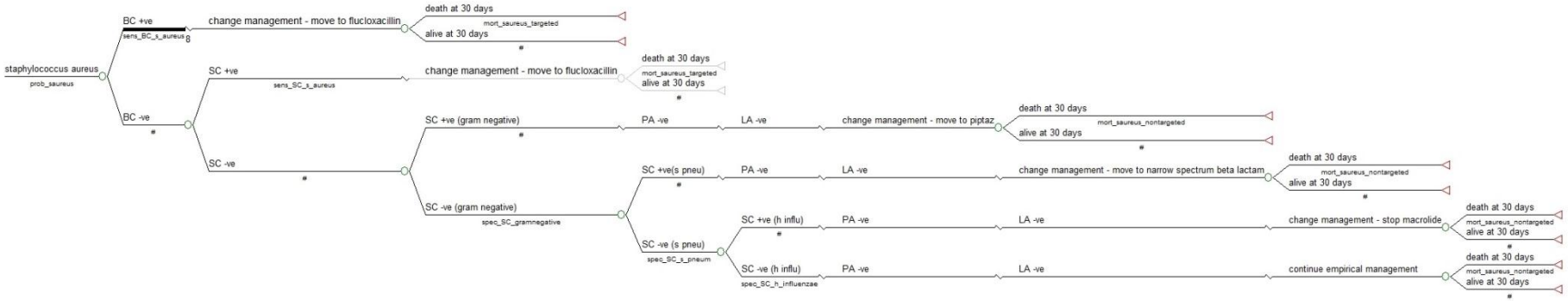


Figure 12: All tests in combination – atypical pathogen arm

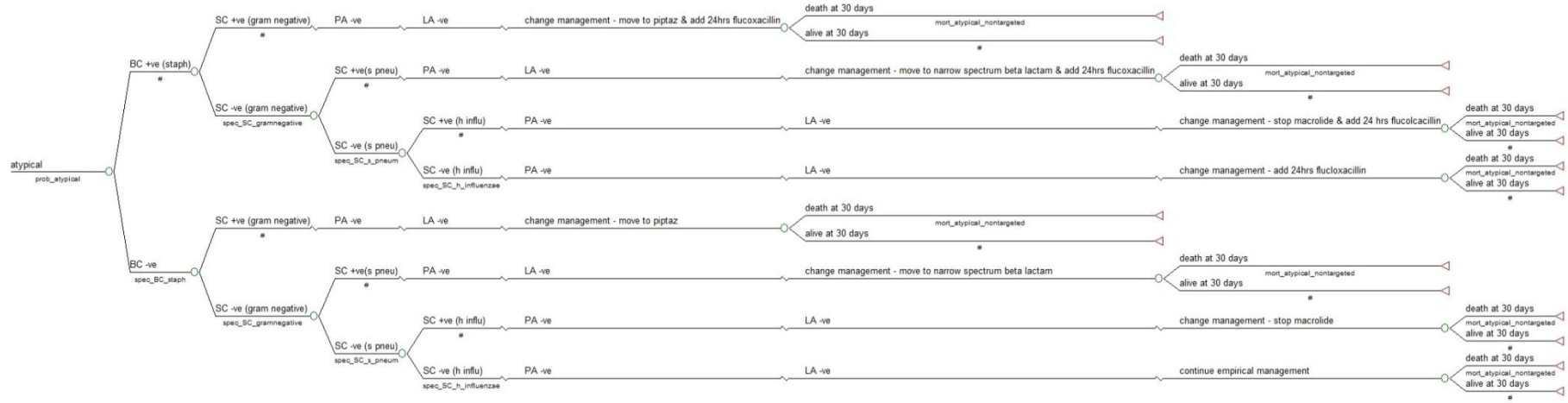
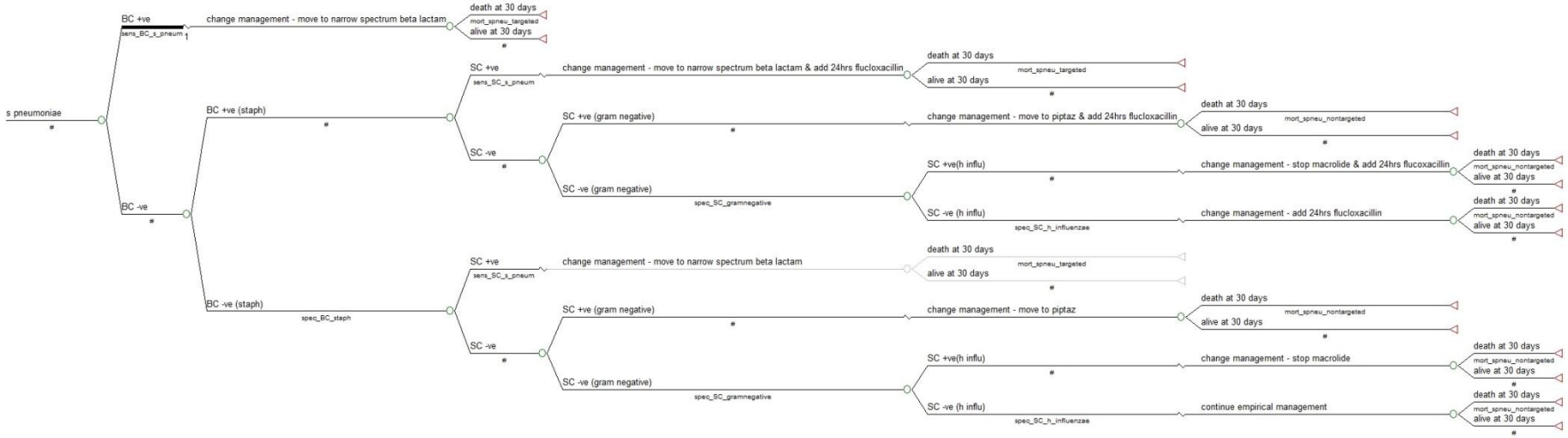
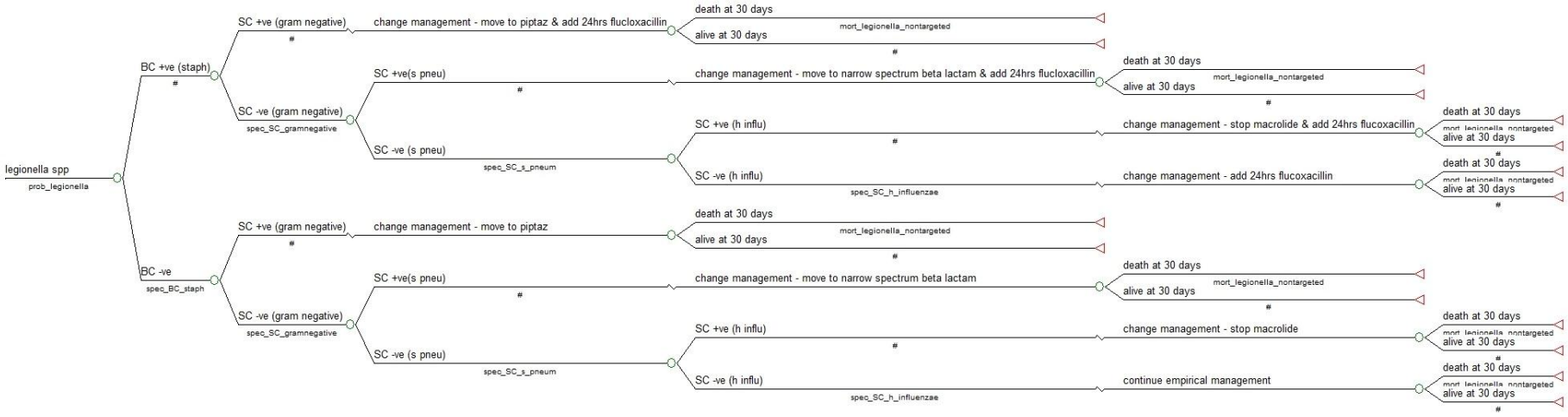


Figure 13: Blood culture and routine sputum culture – *S. pneumoniae* arm



Note: The H. influenzae and Gram-negative arms follow the same structure

Figure 14: Blood culture and routine sputum culture – *L. pneumophila* arm



Note: The atypical pathogen arm follows the same structure

Figure 15: Blood culture and routine sputum culture – *S. aureus* arm

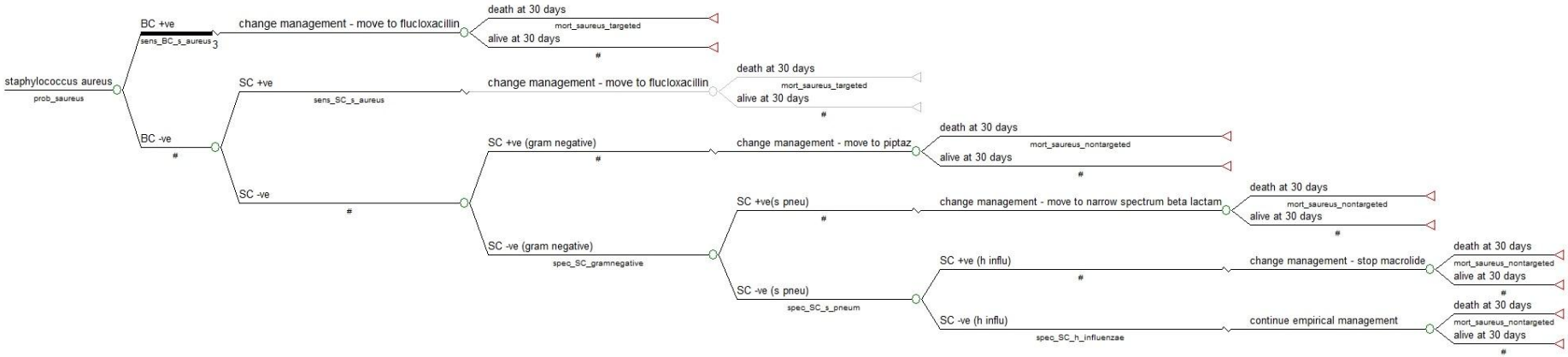
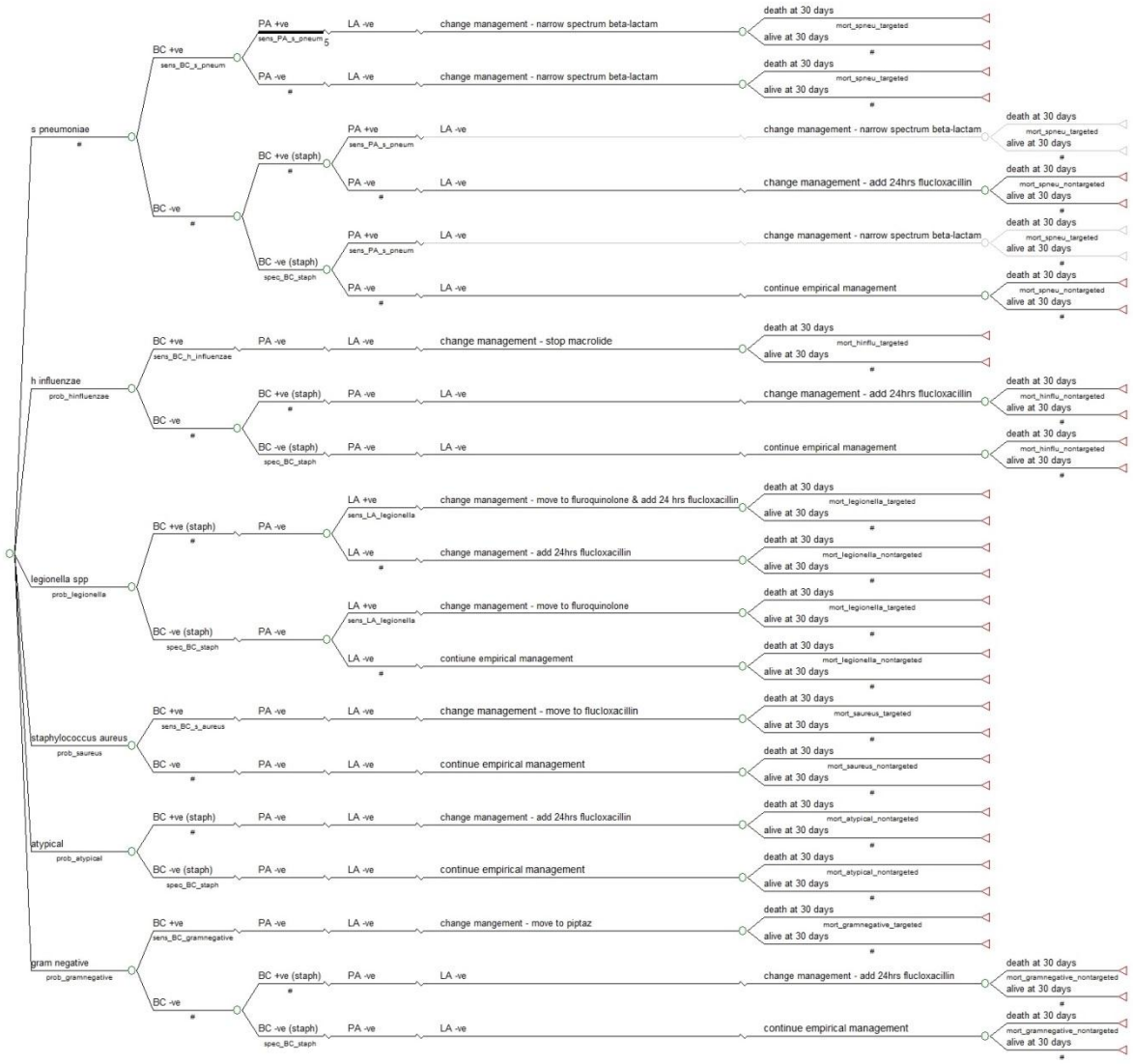


Figure 16: Blood culture and urinary antigen arm



1.5 References

1 MIMS Online. 2013. Available from: http://www.mims.co.uk/ [Last accessed: 15 July 2013]

2 Barrett-Connor E. The nonvalue of sputum culture in the diagnosis of pneumococcal pneumonia. American Review of Respiratory Disease. 1971; 103(6):845-848

3 Department of Health. NHS reference costs 2011-12. 2012. Available from: https://www.gov.uk/government/publications/nhs-reference-costs-financial-year-2011-to-2012 [Last accessed: 15 July 2013]

4 Falguera M, Ruiz-Gonzalez A, Schoenenberger JA, Touzon C, Gazquez I, Galindo C et al. Prospective, randomised study to compare empirical treatment versus targeted treatment on the basis of the urine antigen results in hospitalised patients with community-acquired pneumonia. Thorax. 2010; 65(2):101-106

5 Guckian JC, Christensen WD. Quantitative culture and gram stain of sputum in pneumonia. American Review of Respiratory Disease. 1978; 118(6):997-1005

6 Health and Social Care Information Centre. Hospital episode statistics. 2013. Available from: http://www.hscic.gov.uk/hes [Last accessed: 15 July 2013]

7 Hofhuis JGM, Spronk PE, van Stel HF, Schrijvers AJP, Rommes JH, Bakker J. The impact of severe sepsis on health-related quality of life: a long-term follow-up study. Anesthesia and Analgesia. 2008; 107(6):1957-1964

8 Joint Formulary Committee. British National Formulary (BNF). 65th edition. London: British Medical Association and The Royal Pharmaceutical Society of Great Britain; 2013. Available from: http://www.bnf.org.uk

9 Kind P, Dolan P, Gudex C, Williams A. Variations in population health status: results from a United Kingdom national questionnaire survey. BMJ. 1998; 316(7133):736-741

10 Lim WS, Baudouin SV, George RC, Hill AT, Jamieson C, Le J, I et al. BTS guidelines for the management of community acquired pneumonia in adults: update 2009. Thorax. 2009; 64(Suppl 3):iii1-55

11 Lim WS, Macfarlane JT, Boswell TC, Harrison TG, Rose D, Leinonen M et al. Study of community acquired pneumonia aetiology (SCAPA) in adults admitted to hospital: implications for management guidelines. Thorax. 2001; 56(4):296-301

12 Mortensen EM, Coley CM, Singer DE, Marrie TJ, Obrosky DS, Kapoor WN et al. Causes of death for patients with community-acquired pneumonia: results from the Pneumonia Patient Outcomes Research Team cohort study. Archives of Internal Medicine. 2002; 162(9):1059-1064

13 Office for National Statistics. Interim life tables, 2009-2011. 2013. Available from: http://www.ons.gov.uk/ons/rel/lifetables/interim-life-tables/2009-2011/index.html [Last accessed: 15 July 2013]

14 Oosterheert JJ, Bonten MJM, Buskens E, Schneider MME, Hoepelman IM. Algorithm to determine cost savings of targeting antimicrobial therapy based on results of rapid diagnostic testing. Journal of Clinical Microbiology. 2003; 41(10):4708-4713

15 Shimada T, Noguchi Y, Jackson JL, Miyashita J, Hayashino Y, Kamiya T et al. Systematic review and metaanalysis: Urinary antigen tests for legionellosis. Chest. 2009; 136(6):1576-1585

16 Sinclair A, Xie X, and Dendukuri N. The clinical effectiveness and cost of a pneumococcal urine antigen immunochromatographic test (BinaxNOW Streptococcus pneumoniae) in the diagnosis of community acquire Streptococcus pneumoniae pneumonia in patients admitted to hospital. Technology Assessment Unit of the McGill University Health Centre (MUHC), 2012. Available from: https://secureweb.mcgill.ca/tau/sites/mcgill.ca.tau/files/muhc_tau_2011_57_binaxnow.pdf

17 Sinclair A, Xie X, Teltscher M, Dendukuri N. Systematic review and meta-analysis of a urine-based pneumococcal antigen test for diagnosis of community-acquired pneumonia caused by Streptococcus pneumoniae. Journal of Clinical Microbiology. 2013; 51(7):2303-2310

18 Soares MO, Welton NJ, Harrison DA, Peura P, Hari M, Harvey SE et al. An evaluation of the feasibility, cost and value of information of a multicentre randomised controlled trial of intravenous immunoglobulin for sepsis (severe sepsis and septic shock): incorporating a systematic review, meta-analysis and value of information analysis. Health Technology Assessment. 2012; 16(7):1-186

19 Uematsu H, Hashimoto H, Iwamoto T, Horiguchi H, Yasunaga H. Impact of guideline-concordant microbiological testing on outcomes of pneumonia. International Journal for Quality in Health Care. 2014; 26(1):100-107