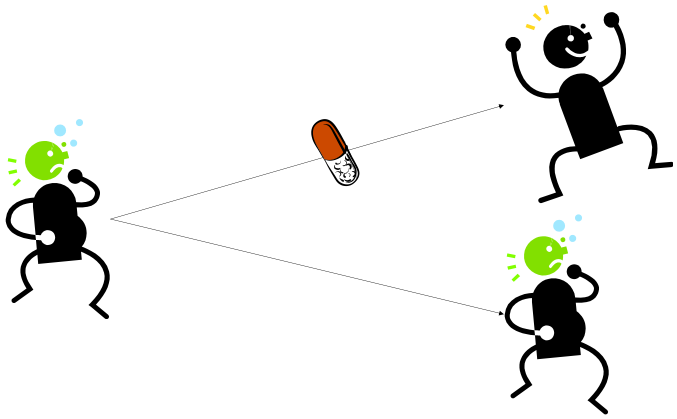
Appendices

Appendix 1. Summary of concepts and motivation for randomisation-based efficacy estimators

1. The importance of randomisation when inferring causal treatment effects

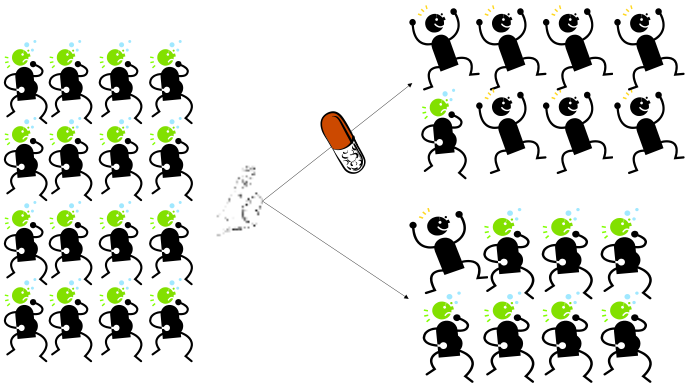
- One of the key reasons we perform experiments is to determine the effect that a treatment has on some outcome of interest – the causal effect.
- In general, we would like to infer these causal effects to the level of individuals. However, without simultaneously observing the effect of both giving and not giving treatment, we will never be able to calculate a true individual-level treatment effect.

Figure 1: Illustration of an individual-level treatment effect



- Instead, we calculate population-level (or average) treatment effects, where the average outcomes of individuals in the treated group are compared to those in the untreated group and we use this calculation as an estimate for the individual-level effects (that we only ever partially observe).
- For this estimate to be valid, the choice to be in the treated / untreated group must be made at random.
    - If the choice is not made at random, the estimate is likely to be biased unless the decision to choose one group over the other (i.e. the selection mechanism) is fully measured and adjusted for. However, this is very unlikely to be the case in practice, where typically some variables that contribute to the selection mechanism are likely to remain unmeasured.

Figure 2: Illustration of a population-level (average) treatment effect from a randomised experiment



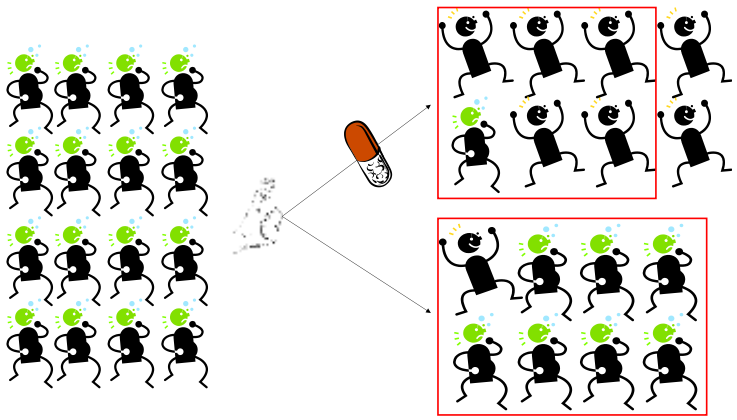
2. What does an Intention to Treat (ITT) analysis allow and what does it not allow?

- ITT analysis allows for a comparison of groups as randomised, independent of both observed and (most importantly) *unobserved* confounders. It reflects the design of the trial and uses randomisation to avoid selection bias. To preserve randomisation, deviations following randomisation (such as lack of adherence to allocated treatment) are not adjusted for.
- When all participants receive their allocated medication as intended, an ITT analysis provides an unbiased estimate of the effect of both prescribing and taking treatment.
- When some participants do not receive their treatment as intended, an ITT analysis can only be guaranteed to provide an unbiased estimate of the effect of prescribing treatment.

3. What is a per-protocol analysis and why is it usually inappropriate to perform in an RCT?

- A per-protocol analysis generally only includes participants who followed study protocol as intended. Examples of protocol deviations could be:
    - Participant was incorrectly randomised
    - Being in the treatment arm and not taking treatment
    - Being in the control arm and taking treatment
    - Not providing follow-up data
- A per-protocol analysis makes the assumption that analysed participants are equivalent to excluded participants (i.e. that the choice to deviate from protocol is made completely at random, or, if there is a selection mechanism, that it has been fully measured and adjusted for).
- However, these exclusions occur post-randomisation, and as illustrated in Point 1, selection mechanisms that are not based on randomisation are likely to yield biased estimates of treatment effects. Therefore using a per-protocol population to estimate treatment effects in RCTs should usually be avoided.

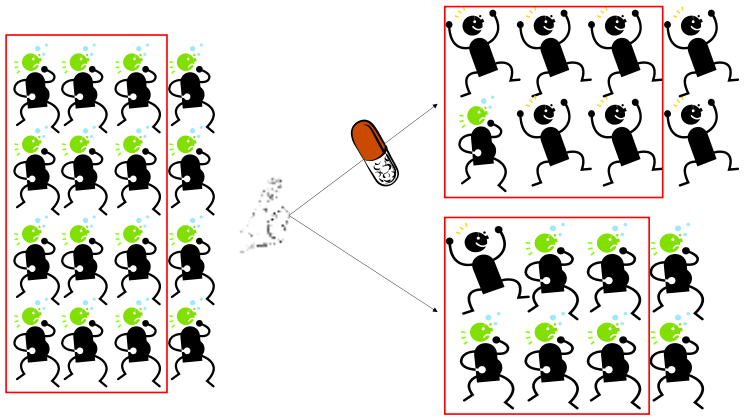
Figure 3: Illustration of per protocol analysis



4. What are randomisation-based efficacy estimators and why are they generally a better approach?

- Randomisation-based efficacy estimators (RBEE) compare the effect of treatment in those who were allocated to and adhered to treatment with those allocated to control who would have adhered to treatment (if allocated to the treatment arm).
- Dependent on data type, there are many ways of calculating a RBEE, but most methods rely on the following core assumptions:
    1. Participants' adherence/compliance-type is a latent trait, a baseline characteristic that is independent of randomisation. One way to think of RBEEs is as the ITT effect in the sub-group of participants who would always adhere to treatment.
    2. Due to randomisation, the proportion of participants classed as non-adherers will be the same in each group.
    3. In the absence of treatment, randomisation in and of itself has no effect on outcome.
- By making these assumptions, observed adherence data can be used to classify individuals and obtain estimates of the effect of receiving treatment on outcome that are not prone to the selection bias commonly seen in traditional efficacy analyses.
- While a binary definition of adherence is often used, this can either make the third core assumption implausible (by including participants in the non-adherent group that may have received some treatment and may therefore benefit from it) or involve a restrictive definition of adherence (e.g. took at least one tablet).
- A continuous definition of adherence makes this third assumption plausible, as zero can represent those who received no treatment. However, the use of a continuous definition implies the additional assumption of a linear relationship between adherence and treatment effect, which is likely to have varying degrees of plausibility depending on setting.

Figure 4: Illustration of randomisation-based efficacy estimator



Appendix 2: Stata syntax for the structural mean models

Structural mean model for "mean clinician-rated symptom severity between days two and four after initial presentation" outcome using two-stage least squares instrumental variables regression

```
ivregress 2sls y c (x=z)
```

In the syntax above, y = outcome, c = covariate, x = exposure, and z = randomisation indicator

Generalised linear (double logistic) structural mean model for "development of new or worsening symptoms" and "presence of any non-respiratory symptoms" outcomes using generalised method of moments

```
logit y x z

matrix from = e(b)

predict xblog, xb

gmm (invlogit(xblog - x*{psi})-ey0), instruments(z)

matrix from = (from, e(b))

gmm (y - invlogit({xb: x z} + {b0})) (invlogit({xb:} + {b0} - x*{psi}) - ey0), instruments(1:x z)
instruments(2:z) winitial(unadjusted, independent) from(from)

lincom[psi]_cons, eform

estat overid
```

In the syntax above, y = outcome, x = exposure, z = randomisation indicator, ey0 = mean exposure-free potential outcome (to stabilise the model, this has been fixed as the proportion of people with positive outcomes in the control group. It can however be directly estimated from the model). This model requires an additional stage (an associational model) because collapsing the logistic SMM over observed exposure (z) depends on the distribution of z. It is therefore not possible to derive causal odds ratios in a single stage. The stages are first run individually to obtain initial values for the joint estimation. The stages are then run jointly to produce standard errors that correctly incorporate the error from the first stage of the model.

Appendix 3: Additional sensitivity analysis with missing adherence data imputed

The aim of this paper was to demonstrate how randomisation-based efficacy estimators can be used to produce unbiased adherence-adjusted estimates of benefits and harms from treatment with amoxicillin for patients consulting with an LRTI. The main effectiveness findings (reference 7 in the main manuscript) were used as the reference results. However, two participants did not have adherence data available for the symptom severity between days 2 and 4 post-randomisation and non-respiratory symptoms/side effects in the 4 weeks post-randomisation outcomes. A total of 104 participants did not have adherence data available for the new or worsening symptoms in the 4 weeks post-randomisation outcome. While the two former outcomes were collected via symptom diaries, the latter was collected from patient notes, and was consequently available for more participants. Table 2 in the manuscript suggests that the level of adherence in participants without self-reported diary or tablet count data was considerably lower (self-reported telephone data was primarily collected in those who did not return diaries). In the presence of missing adherence data, there may remain some residual bias. To understand how severe this bias could be (particularly, how low the odds ratio for new or worsening symptoms could be), Table 9 provides the findings of additional sensitivity analyses where participants with missing adherence data are assumed to have not taken any study medication (i.e. their adherence level is 0%). The findings demonstrate that making this most extreme assumption about missing adherence data did not alter the clinical conclusions that were drawn from the analyses.

Table 9: Efficacy analysis with missing adherence data imputed as 0%

Outcome	**Effectiveness***	**Effectiveness for whom adherence data were also available[†]**	**Efficacy per 10% increase in adherence[†]**	**Maximum efficacy (100% adherence)[†]**	**Efficacy per 10% increase in adherence*[§]**	**Maximum efficacy (100% adherence)*[§]**
Adjusted between-group mean difference in symptom severity between days 2 and 4 post-randomisation	-0.07 (-0.15 to 0.01)	-0.07 (-0.15 to 0.01)	-0.008 (-0.017 to 0.001)	-0.08 (-0.17 to 0.01)	-0.008 (-0.017 to 0.001)	-0.08 (-0.17 to 0.01)
Odds ratio for developing new or worsening symptoms in the 4 weeks post-randomisation	0.79 (0.63 to 0.99)	0.81 (0.64 to 1.03)	0.978 (0.960 to 0.998)	0.81 (0.66 to 0.98)	0.973 (0.954 to 0.994)	0.76 (0.62 to 0.94)
Odds ratio for reporting non-respiratory symptoms/side effects in the 4 weeks post-randomisation	1.28 (1.03 to 1.59)	1.28 (1.04 to 1.59)	1.028 (1.011 to 1.046)	1.32 (1.12 to 1.57)	1.028 (1.011 to 1.046)	1.32 (1.11 to 1.56)

* Analysis based on 1789, 2027 and 1727 participants for the symptom severity, new symptoms and side effect outcomes respectively. † Analysis based on 1787, 1923 and 1725 participants for the symptom severity, new symptoms and side effect outcomes respectively. § Assuming those participants with missing adherence data did not take any medication (i.e. their adherence level is 0%).