Antibiotic review kit for hospitals (ARK-Hospital): a stepped-wedge cluster-randomised controlled trial


Summary
Background Strategies to reduce antibiotic overuse in hospitals depend on prescribers taking decisions to stop unnecessary antibiotic use. There is scarce evidence for how to support these decisions. We evaluated a multifaceted behaviour change intervention (i.e., the antibiotic review kit) designed to reduce antibiotic use among adult acute general medical inpatients by increasing appropriate decisions to stop antibiotics at clinical review.

Methods We performed a stepped-wedge, cluster (hospital)-randomised controlled trial using computer-generated sequence randomisation of eligible hospitals in seven calendar-time blocks in the UK. Hospitals were eligible for inclusion if they admitted adult non-elective general or medical inpatients, had a local representative to champion the intervention, and could provide the required study data. Hospital clusters were randomised to an implementation date occurring at 1–2 week intervals, and the date was concealed until 12 weeks before implementation, when local preparations were designed to start. The intervention effect was assessed using data from pseudonymised routine electronic health records, ward-level antibiotic dispensing, and monthly hospital antibiotic use. Co-primary outcomes were monthly antibiotic defined daily doses per adult acute general medical admission (hospital-level, superiority) and all-cause mortality within 30 days of admission (patient level, non-inferiority margin of 5%). Outcomes were assessed in the modified intention-to-treat population (i.e., excluding sites that withdrew before implementation). Intervention effects were assessed by use of interrupted time series analyses within each site, estimating overall effects through random-effects meta-analysis, with heterogeneity across prespecified potential modifiers assessed by use of meta-regression. This trial is completed and is registered with ISRCTN, ISRCTN12674243.

Findings 58 hospital organisations expressed an interest in participating. Three pilot sites implemented the intervention between Sept 25 and Nov 20, 2017. 43 further sites were randomised to implement the intervention between Feb 12, 2018, and July 1, 2019, and seven sites withdrew before implementation. 39 sites were followed up for at least 14 months. Adjusted estimates showed reductions in total antibiotic defined daily doses per acute general medical admission (–4.8% per year, 95% CI –9.9 to –0.8) following the intervention. Among 7160421 acute general medical admissions, the ARK intervention was associated with an immediate change of –2.7% (95% CI –5.7 to 0.3) and sustained change of 3.0% (–0.1 to 6.2) in adjusted 30-day mortality.

Interpretation The antibiotic review kit intervention resulted in sustained reductions in antibiotic use among adult acute general medical inpatients. The weak, inconsistent intervention effects on mortality are probably explained by the onset of the COVID-19 pandemic. Hospitals should use the antibiotic review kit to reduce antibiotic overuse.

Funding UK National Institute for Health and Care Research.

Copyright © 2022 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.
Research in context

Evidence before this study

Patients who are acutely ill often need antibiotics before full diagnostic information is available. Consequently, reducing overuse of antibiotics in hospitals requires prescribers to review and, when appropriate, stop unnecessary antibiotic prescriptions. Evidence-based tools to support prescribers to stop unnecessary antibiotics do not exist.

We searched PubMed, with no language or date restrictions, on Jan 31, 2022, for clinical studies that focused on improving antibiotic use for adults who were admitted to hospital using the terms "anti-bacterial agents therapeutic use" and "antibiotic stewardship." Among the 427 studies found, most were uncontrolled evaluations of different approaches to education, decision support, and feedback. These studies included one before-and-after study, which identified no effect of unsupported clinician-led prescription review on antibiotic use. Three small, hospital-level, cluster-randomised trials were identified. One trial evaluated different approaches to feedback, one compared different hospital specialities, and one reported that intense feedback was effective in reducing antibiotic use. All three trials were small and none considered clinical outcomes or sustainability. Research is needed to deliver effective interventions that are ready for implementation into clinical practice. This weak evidence base explains the differences that exist in national policy recommendations around clinician-led antibiotic prescription review for hospital antibiotic stewardship between, for example, the UK and the USA.

Added value of this study

We evaluated a multifaceted intervention to support clinician-led antibiotic prescription review (ie, the antibiotic review kit [ARK] intervention) and showed that ARK was effective in achieving safe sustained reductions in organisation-level antibiotic use among acute, general medical hospital admissions. Our findings deal with the uncertainty about whether clinician-led prescription review is an effective approach to antibiotic stewardship in hospital practice by being highly pragmatic, evaluating sustainability, and robustly exploring potential patient-level harms of this approach to reducing antibiotic use. Furthermore, the ARK-Hospital Programme delivers resources to support effective clinician-led prescription review ready for adoption into clinical practice.

Implications of all available evidence

The ARK intervention is safe and effective in reducing antibiotic use among adult acute, general medical hospital admissions. The tools used are now freely available for adoption into practice. Available evidence comes from research using paper-based prescribing and future research should establish how antibiotic prescription reviews should be built into electronic prescribing systems.

Plymouth, UK (K Fox MBBS); Airedale NHS Foundation Trust, Keighley, UK (F Foot MSc); Hampshire Hospitals NHS Foundation Trust, Basingstoke, UK (V Garcia-Arias MSc); William Harvey Hospital, East Kent Hospitals University NHS Foundation Trust, Ashford, UK (S Glass MBChB); Pharmacy Department, Altnagelvin Area Hospital, Western Health and Social Care Trust, Londonderry, UK (C Gormley MSc); Whiston Hospital, St Helens and Knowsley Teaching Hospitals NHS Trust, Prescot, UK (R Gray MBChB); Cumberland Infirmary Carlisle, North Cumbria Integrated Care NHS Foundation Trust, Carlisle, UK (C Hamilton MBChB); Wirral University Teaching Hospital NHS Foundation Trust, Wirral, UK (D Harvey MBChB); Nottingham University Hospitals NHS Trust, Nottingham, UK (T Hills BPharm); Royal Berkshire NHS Foundation Trust, Reading, UK (S Iyer FRCP); Country Hospital, Wye Valley NHS Trust, Hereford, UK (A Johnson FRCP); John Radcliffe Hospital, Oxford, UK (Prof T E A Peto, N Jones PhD, Prof D W Cockx); The Royal Wolverhampton NHS Trust, Wolverhampton, UK (P Kang MPharm); Great Western Hospitals NHS Foundation Trust, Swindon, UK (G Kipl MSc); Royal Free London NHS Foundation Trust, London, UK (D Mack FRCPath, S Hopkins FRCP); Betsi Cadwaladr University Health Board, Ysbyty Gwynedd, Bangor, UK (C Makanga MSc); York and Scarborough Teaching Hospitals NHS Foundation Trust, York, UK (D Mawer MBChB); South Eastern Health and Social Care Trust, Ulster Hospital, Belfast, UK (C McCullagh MSc); North Middlesex University Hospital NHS Trust, London, UK (M Mirfenderesky FRCP); Countess of Chester Hospital NHS Foundation Trust, Chester, UK (R McEwen MBChB); James Cook University Hospital, South Tees Hospitals NHS Foundation Trust, Middlesbrough, UK (S Nag FRCPath); Antrim Area Hospital, Northern Health and Social Care Trust, Antrim, UK are used whenever appropriate.6 However, translating research on antimicrobial stewardship into practice is hampered by the poor quality of evidence, particularly weaknesses in the intervention design process and the study designs used, which are usually underpowered, not experimental, and do not consider clinical outcomes.7,8

In primary care, restrictive strategies for antimicrobial stewardship, such as avoiding or delaying antibiotics in respiratory tract infection, can be safe and effective.9,10 By contrast, hospital antimicrobial stewardship interventions that enable improved prescribing are more acceptable than restrictive strategies and can reduce overuse and length of stay without compromising mortality.11 The need to ensure that patients with serious bacterial infections are treated promptly before a diagnosis is confirmed means that ongoing review and revision of hospital antibiotic prescriptions is required to safely minimise unnecessary use. In England, the Department of Health’s guidance Start Smart—Then Focus requires prescribers to review and revise antibiotic prescriptions every 48–72 h.12 In the USA, the analogous term antibiotic timeouts is used, but revised Centers for Disease Control and Prevention guidance in 2019 prioritised pharmacist-led audit and feedback to prescribers,13 in light of a non-interventional guidance in 2019 prioritised pharmacist-led audit and feedback to prescribers,13 but revised Centers for Disease Control and Prevention guidance in 2019 prioritised pharmacist-led audit and feedback to prescribers,13 in light of a non-interventional before-and-after study, which identified no effect of unsupported clinician-led prescription review on antibiotic use. All three trials were small and none considered clinical outcomes or sustainability. Research is needed to deliver effective interventions that are ready for implementation into clinical practice. This weak evidence base explains the differences that exist in national policy recommendations around clinician-led antibiotic prescription review for hospital antibiotic stewardship between, for example, the UK and the USA.

Antibiotic use was reduced after introducing Start Smart—Then Focus in 2011, antibiotic consumption in English hospitals continued to rise year on year until the COVID-19 pandemic in 2020.14 This increase was despite financial incentives to reduce hospital prescribing, first through a Commissioning for Quality and Innovation framework in 2016–18 and then its incorporation into the National Health Service (NHS) standard contract for acute hospitals.15 Although high rates of prescription review were achieved,16 most review decisions were to adjust rather than stop antibiotics. The antibiotic review kit (ARK) programme for hospitals aimed to develop and evaluate a multifaceted behaviour change intervention to safely reduce antibiotic use in acute general medical inpatients. The ARK programme created a four-component intervention to help prescribers to take appropriate decisions to stop or continue antibiotics at prescription review, comprising a novel prescribing decision aid, an online training tool supporting use of the decision aid, guidance for implementing audit and feedback, and a patient leaflet.17,18 In this study we report the immediate and sustained effect of the ARK intervention on antibiotic consumption at the hospital level and clinical outcomes at the patient level.

Methods

Study design

Following a feasibility evaluation at Brighton and Sussex University Hospitals NHS Trust, an acute hospital,19 the ARK intervention was evaluated at eligible participating hospitals across all UK nations by use of a stepped-wedge
cluster-randomised controlled trial.27 A cluster design was essential to avoid contamination from health-care professionals moving between teams within a hospital. A stepped-wedge design was essential given the few UK secondary care organisations that could be randomised (appendix p 46).

Ethical approval was from the South Central Oxford C Research Ethics Committee (17/SC/0034) and the Confidentiality Advisory Group (17/CAG/0015) without individual patient consent because electronic health records were pseudonymised and no personal identifiable data was collected other than date of death.

Clusters and participants

The unit of observation was a hospital organisation offering services for non-elective medical admissions (appendix p 3). Sites were approached through professional networks and the Society for Acute Medicine. Eligible sites needed to admit adult (ie, aged >15 years) general or medical inpatients, have a local representative (known as a champion) who was willing to lead intervention implementation, and be able to provide the required study data. Since the intervention targeted prescribers on acute general medical wards and used electronic health records to ascertain patient-level outcomes, the study population was defined using the consultant specialty codes that were most often used to admit adult general medicine inpatients (appendix p 45).27 Sites were asked to exclude patients who opted out of having their health records used for research purposes (appendix pp 3–4). The protocol is included in the appendix (p 156).

Randomisation and masking

Eligible sites were randomised by use of a computer-generated list by the trial statistician (ASW), including the pilot sites (one block of three sites) and main trial sites (six blocks of six sites and one block of seven sites), to an intervention implementation date. Implementation was staggered across sites in 1–2 week intervals, with breaks over the Christmas period and in August given the high rate of staff holidays or when the funder requested a delay over the Christmas period and in August given the high risk rate of staff holidays. Implementation was scheduled in pilot sites (one block of three sites) and main trial sites (six blocks of six sites and one block of seven sites) to a randomised implementation date. Implementation was staggered across sites in 1–2 week intervals, with breaks over the Christmas period and in August given the high rate of staff holidays or when the funder requested a delay.

Procedures

The intervention comprised a decision aid that was intended to be embedded in the hospital prescription process, prompting prescribers to clarify the level of diagnostic uncertainty at antibiotic initiation by classifying infection risk as possible or probable, and then either stopping the prescription if a clear indication for ongoing antibiotic treatment could not be established at 48–72 h review or finalising the prescription if a clear indication could be established; online training to motivate and support use of the decision aid; implementation guidance, including audit and feedback tools; and a patient leaflet.28,29 The ARK tools are freely available through the British Society for Antimicrobial Chemotherapy. By supporting decisions to stop antibiotics at clinical review, the intervention aimed to safely reduce antibiotic use through reducing treatment duration, rather than by targeting the appropriateness of initial prescriptions. Fidelity of intervention implementation was assessed with eight predefined criteria by the study team at each individual site up to 16 weeks after implementation (table).

Study data were collected for 24 months before and after implementation at the first main trial site until at least 14 months following implementation of the final site, to facilitate outcome assessment before and after implementation. Time periods for the co-primary outcomes are shown in the appendix (appendix p 46).

All outcomes were assessed using pseudonymised electronic health records from adult (age ≥16 years) acute general medical admissions (further details, including data cleaning, are shown in the appendix pp 6–9, 48), bulk antibiotic dispensing, on the wards that implemented ARK, and C difficile test results. Date of death within 90 days of admission (in or out of hospital) was obtained by sites through linkage with national registries. Patient-level antibiotic data and laboratory results (ie, microbiology, haematology, biochemistry, and imaging tests) were provided by few sites, preventing further analysis of these data (appendix p 5). Uptake of the intervention was assessed through a process evaluation and prescription review audits.

Outcomes

The trial had two co-primary outcomes: antibiotic defined daily doses (DDD)s per adult acute general medical admission (superiority) and all-cause mortality within 30 days of admission (in or out of hospital); non-inferiority, relative margin 5% for an immediate step change associated with implementation, assuming a constant rate before and after implementation). Both outcomes were assessed by estimating the immediate effect and the sustained year-on-year effect.

Secondary antibiotic (superiority) outcomes were total antibiotic DDDs per acute general medical bed-day and DDDs per admission for specific antibiotic groups, including carbapenems, parenteral and oral administration, broad-spectrum and narrow-spectrum antibiotics, and the UK Health Security Agency’s interpretations of Access, Watch, and Reserve from WHO’s Essential Medicines List (appendix pp II–12).30 Admissions, rather than bed-days, were used as the denominator in the primary analysis because bed-days can be influenced by non-medical reasons for prolonged hospital stays (eg, awaiting discharge to another place of care). Although...
the protocol specified patient-level antibiotic outcomes (ie, days on antibiotics, antibiotic-days per admission or bed-day, and antibiotic restart after discontinuation for >48 h).\(^1\) only four sites had both the electronic prescribing systems and information technology resources required to provide these data, so these outcomes could not be analysed (figure 1). Secondary non-inferiority outcomes were 90-day mortality, admission to an intensive care unit, length of stay, emergency hospital readmission (to any specialty) within 30 days of discharge, and *Clostridioides difficile* infection or colonisation within 90 days of admission.

Ddds per admission for piperacillin–tazobactam and quinolones and length of stay of 48 h or more were considered in exploratory analyses.

### Statistical analysis

An interrupted time series analysis estimated the immediate effect of the intervention (ie, step change) and sustained effect on year-on-year trends after implementation versus before implementation within each site for the co-primary outcomes of antibiotic DDDS per adult acute general medical admission and all-cause mortality within 30 days of admission, by use of the randomised implementation date. All analyses used an intention-to-treat approach that was modified to exclude seven sites that withdrew after randomisation but before implementation, from which no data were collected. Overall intervention effects were then estimated by use of random effects meta-analysis, using meta-regression to assess heterogeneity in effects across prespecified potential effect modifiers. As the intervention did not change after the pilot, following the approved protocol, the primary analysis included pilot and main trial sites to maximise power. With a minimum of 36 sites, the stepped-wedge cluster-randomised design had more than 85% power to exclude an immediate 5% relative increase in 30-day mortality and to detect a 15% relative reduction in antibiotic use associated with intervention implementation.\(^2\) The 5% relative non-inferiority margin for the 30-day mortality outcome means that to declare non-inferiority, the upper bound of 95% CI for the relative change in mortality had to not exceed a 5% increase. Further details are given in the appendix (appendix pp 9–10).

- **Randomisation date**
  - Median (range) Aug 6, 2018 (May 25, 2017–April 8, 2019)

- **Randomised implementation date**
  - Median (range) Nov 5, 2018 (Sept 25, 2017–July 1, 2019)

- **Any implementation delay**
  - Yes 9 (23%)

- **Region**
  - South of England 13 (33%)
  - North of England 10 (26%)
  - Midlands and east of England 6 (15%)
  - Northern Ireland 3 (8%)
  - Wales 2 (5%)
  - Scotland 1 (3%)

- **Size (ie, acute beds available)**
  - Median (IQR; range) 690 (540–939; 268–1484)
  - Large (>850 beds) 13 (33%)
  - Medium (551–850 beds) 14 (35%)
  - Small (<550 beds) 12 (31%)

- **Specialty of the principal investigator**
  - Microbiologist 19 (49%)
  - Pharmacist 10 (26%)
  - Acute medicine 7 (18%)
  - Microbiology and infectious diseases 2 (5%)
  - Microbiology and acute medicine 1 (3%)

- **Prescribing system at implementation**
  - Paper 25 (64%)
  - E-prescribing (Cerner) 6 (15%)
  - E-prescribing (JAC) 3 (8%)
  - E-prescribing (other) 5 (13%)

- **How the decision aid was implemented**
  - Hard stop 21 (54%)
  - Soft stop 9 (23%)
  - Neither 9 (23%)

- **Baseline acute or general medicine admissions**
  - Median (IQR; range) 38 045 (26 008–56 294; 11 416–77 946)

- **Baseline DDDS per admission in the 12 months before implementation**
  - Total 3.05 (2.19–5.35; 0.43–16.16)
  - Parenteral administration 0.91 (0.65–1.43; 0.15–5.00)
  - Oral administration 2.16 (1.56–3.85; 0.26–13.02)
  - Broad spectrum 1.02 (0.72–1.53; 0.03–4.39)
  - Narrow spectrum 2.23 (1.40–3.44; 0.40–12.61)
  - Access 1.52 (0.83–2.39; 0.29–9.68)
  - Access or watch† 0.77 (0.51–1.31; 0.02–3.97)
  - Watch 0.88 (0.43–1.19; 0.02–3.96)
  - Reserve 0.09 (0.05–0.19; <0.01–0.55)
  - Carbapenems 0.06 (0.03–0.08; 0.00–0.31)
  - Piperacillin-tazobactam 0.12 (0.05–0.22; <0.01–0.80)
  - Quinolones 0.26 (0.11–0.42; <0.01–1.43)

\(^1\) Articles

### Articles

- 11 (Sept 25, 2017–July 1, 2019)
  - 4 (May 25, 2017–April 8, 2019)

**Clusters (n=39)**

<table>
<thead>
<tr>
<th>Articles</th>
<th>Clusters (n=39)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomisation date</td>
<td>Median (range) Aug 6, 2018 (May 25, 2017–April 8, 2019)</td>
</tr>
<tr>
<td>Randomised implementation date</td>
<td>Median (range) Nov 5, 2018 (Sept 25, 2017–July 1, 2019)</td>
</tr>
<tr>
<td>Any implementation delay</td>
<td>Yes 9 (23%)</td>
</tr>
<tr>
<td>Region</td>
<td>South of England 13 (33%)</td>
</tr>
<tr>
<td>Size (ie, acute beds available)</td>
<td>Median (IQR; range) 690 (540–939; 268–1484)</td>
</tr>
<tr>
<td>Specialty of the principal investigator</td>
<td>Microbiologist 19 (49%)</td>
</tr>
<tr>
<td>Prescribing system at implementation</td>
<td>Paper 25 (64%)</td>
</tr>
<tr>
<td>How the decision aid was implemented</td>
<td>Hard stop 21 (54%)</td>
</tr>
<tr>
<td>Baseline acute or general medicine admissions</td>
<td>Median (IQR; range) 38 045 (26 008–56 294; 11 416–77 946)</td>
</tr>
<tr>
<td>Baseline DDDS per admission in the 12 months before implementation</td>
<td>Total 3.05 (2.19–5.35; 0.43–16.16)</td>
</tr>
<tr>
<td>Parenteral administration</td>
<td>0.91 (0.65–1.43; 0.15–5.00)</td>
</tr>
<tr>
<td>Oral administration</td>
<td>2.16 (1.56–3.85; 0.26–13.02)</td>
</tr>
<tr>
<td>Broad spectrum</td>
<td>1.02 (0.72–1.53; 0.03–4.39)</td>
</tr>
<tr>
<td>Narrow spectrum</td>
<td>2.23 (1.40–3.44; 0.40–12.61)</td>
</tr>
<tr>
<td>Access</td>
<td>1.52 (0.83–2.39; 0.29–9.68)</td>
</tr>
<tr>
<td>Access or watch†</td>
<td>0.77 (0.51–1.31; 0.02–3.97)</td>
</tr>
<tr>
<td>Watch</td>
<td>0.88 (0.43–1.19; 0.02–3.96)</td>
</tr>
<tr>
<td>Reserve</td>
<td>0.09 (0.05–0.19; &lt;0.01–0.55)</td>
</tr>
<tr>
<td>Carbapenems</td>
<td>0.06 (0.03–0.08; 0.00–0.31)</td>
</tr>
<tr>
<td>Piperacillin-tazobactam</td>
<td>0.12 (0.05–0.22; &lt;0.01–0.80)</td>
</tr>
<tr>
<td>Quinolones</td>
<td>0.26 (0.11–0.42; &lt;0.01–1.43)</td>
</tr>
</tbody>
</table>

**Table continues on next page**
exploratory analysis by use of logit models and exploratory analyses for DDDs per admission for piperacillin–tazobactam and quinolones were modelled by use of negative binomial regression. All clinical outcomes were modelled by use of a robust variance adjustment by patient.

Since the COVID-19 pandemic profoundly affected both primary and secondary outcomes, models included a binary indicator for March–June, 2020, for antibiotic outcomes (measured monthly) and March 1–June 30, 2020, for patient-level clinical outcomes, unless otherwise noted. Sensitivity analyses excluded admissions after March 1, 2020, including 12 sites with less than 12 months of data after implementation as a result. Antibiotic models additionally adjusted for seasonal effects by including month of year as a sin()+cos() function to ensure smooth risk changes year to year. Non-antibiotic models also adjusted for individual admission-level covariates, regardless of statistical significance (based on the findings of Walker and colleagues): sex, age, immunosuppression, deprivation percentile, Charlson comorbidity index and its interaction with age, admission method, admission source, admission specialty, patient classification, admission day of the week (ie, weekend vs weekday), admission day of year and time of day (both modelled as a sin()+cos() function, with an interaction between time of day and day of week), and number of overnight admissions and any previous overnight complex (ie, >1 consultant episode, excluding episodes in the emergency department and rehabilitation) admission in the past year. Ethnicity was missing for a median of 8–8% admissions (IQR 4.5–18.4) per site so was not adjusted for. Further details are shown in the appendix (pp 4–5, 21–28).

All analyses used Stata/MP version 17.0. The data monitoring committee reviewed outcome data three times during the trial, using a Haybittle-Peto statistical rule for early stopping. This trial is registered with ISRCTN, number ISRCTN12674243.

Role of the funding source
The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results
58 UK acute hospital organisations expressed an interest in participating, of which 46 sites agreed to join the pilot or main trial. Three pilot sites implemented the intervention between Sept 25–Nov 20, 2017; 43 further sites were randomised to implement between Feb 12, 2018, and July 1, 2019, of which seven withdrew before implementation and were excluded from analyses as no data were collected (figure 1). 39 sites were included in the analysis of 30-day mortality and 38 sites were included in the analysis of total antibiotic DDDs per adult acute general medical admission because antibiotic data were not available from one site (figure 1, appendix p 29).

13 sites were classed as large (ie, >850 beds available, median 991), 14 medium (ie, 551–850 beds, median 670), and 12 small (ie, ≤550 beds, median 487; table). Sites were distributed across the UK, with the largest number in the south of England. Most champions were microbiologists. At implementation, prescribing was paper based at most sites. 21 (54%) of 39 sites implemented the decision aid with a hard stop to the initial prescription unless revised by 72 h, nine sites (23%) implemented as a soft stop, emphasising the need to stop or finalise within 72 h, and nine (23%) sites did neither.

Antibiotic use in the 12 months before randomised implementation varied widely, both in total DDDs per admission and specific agents, classes, and Access, Watch, and Reserve categories (table; appendix p 48). Access antibiotics accounted for 30.7–85.2% of total DDDs, Watch for 4.8–44.1%, and Reserve for 0.3–5.7%.

Site champions named a median of 19 (IQR 14–34, range 5–72) people as essential for doing the online training; a median of 78% (63–90) completed the training by 12 weeks, with 16 (41%) of 39 sites below the target of 70% or higher (figure 2A). The total number of staff completing training also varied substantially, with a median of 24 (IQR 15–41) staff trained per 100 acute...
one site implemented the intervention in the feasibility phase of the research. These data were published flow diagram of participating hospital organisations in antibiotic use after implementation. Individual-level data were not analysed due to scarcity. immediate decline (>90%) in reported antibiotic defined daily doses, which prevented valid assessment of trends §one pilot site introduced an electronic prescribing system 6 months after implementation leading to an from sites that never implemented the intervention and hence would show no intervention effect on mortality. was a non-inferiority comparison, it was more important to replace these sites than use resources collecting data decision not to fund data collection in sites withdrawing before implementation was made because, as mortality study materials typically due to the resource implications of participation in the research, conflicts with local antimicrobial stewardship initiatives, or being unable to provide mandatory electronic health record data. ‡the decision not to fund data collection in sites withdrawing before implementation was made because, as mortality was a non-inferiority comparison, it was more important to replace these sites than use resources collecting data from sites that never implemented the intervention and hence would show no intervention effect on mortality. §one pilot site introduced an electronic prescribing system 6 months after implementation leading to an immediate decline (>90%) in reported antibiotic defined daily doses, which prevented valid assessment of trends in antibiotic use after implementation. (individual-level data were not analysed due to scarcity.)

39 sites provided de-identified patient-level electronic health record data for clinical outcomes 38 sites provided antibiotic data§ 2 provided hospital-level data due to limitations in local pharmacy information systems 36 provided ward-level data where the intervention was implemented 4 provided both ward-level and individual patient-level data|| 39 sites implemented the intervention and were included in analyses)

7 sites withdrew after randomisation but before implementation 1 due to changes in personnel (eg, nominated champion left) 5 due to resource implications of participation in the research 7 additional sites recruited to replace those that withdrew after randomisation 12 sites excluded 1 was included in the feasibility assessment* 11 sites chose not to participate*

39 sites agreed to join the pilot (n=3) or to be randomised in future in the main trial (n=36)

71 UK acute hospital organisations expressed interest in participating

51 due to resource implications of participation in the research, conflicts with local antimicrobial stewardship initiatives, or being unable to provide mandatory electronic health record data. the decision not to fund data collection in sites withdrawing before implementation was made because, as mortality was a non-inferiority comparison, it was more important to replace these sites than use resources collecting data from sites that never implemented the intervention and hence would show no intervention effect on mortality. §one pilot site introduced an electronic prescribing system 6 months after implementation leading to an immediate decline (>90%) in reported antibiotic defined daily doses, which prevented valid assessment of trends in antibiotic use after implementation. (individual-level data were not analysed due to scarcity.)

sites contributed a median 23 months (range 14–37) of antibiotic data after implementation (appendix p 46). adjusting for the effects of COVID-19 and interrupted time series trends (shown by site in appendix pp 51–127), the intervention was associated with a –1.0% (95% CI –4.0 to 2.1) immediate change in total antibiotic DDDs per admission and a sustained –4.8% (–9.1 to –0.2) change per year subsequently (figure 3, 4), with little association between the immediate and longer-term intervention effects across sites (r=–0.088, p=0.60; appendix p 128). there was substantial heterogeneity in trajectories of DDDs per admission before intervention and after intervention (appendix p 129). Intervention effects were similar unadjusted (figure 3) and excluding all follow-up from March, 2020 (appendix pp 13–15). there was no evidence that immediate effects on total DDDs per admission at implementation (–0.5%, 95% CI –2.7 to 1.7, per additional fidelity criteria achieved) and on year-on-year trends after versus before intervention (–1.4%, –4.9 to 2.3, per additional fidelity criteria achieved) were associated with overall implementation fidelity (figure 4). immediate reductions in total DDDs per admission were greater among sites with processes for ongoing audit and feedback in place by implementation (by –16.6%, 95% CI –28.5 to –2.8, relative to sites that did not have processes in place by implementation) and greater among sites that submitted postimplementation audit data within 4 weeks following implementation (by –8.3%, –15.1 to –1.0, relative to sites that did not submit audit data within 4 weeks). however, the relative reduction in immediate implementation effect among sites that submitted postimplementation audit data within 4 weeks was not observed after we adjusted for whether the site had a process in place for ongoing audit and feedback by the implementation date in a multivariate model (appendix pp 16–20). we found non-significantly greater sustained reductions in total DDDs per admission among sites that introduced ARK categories into the prescribing process by implementation than among sites that had not introduced ARK categories by implementation (by –11.5%, 95% CI –22.9 to 1.7, relative to the reference group; appendix p 16) and among sites with higher uptake of the online learning by implementation (ie, with ≥20 people per 100 acute beds completing the training) than among sites with lower uptake (by –9.9%, 95% CI –19.7 to 1.1, versus sites training <20 people per 100 acute beds). Medium-sized sites also had non-significantly greater reductions in DDDs at implementation (by –7.4%, 95% CI –14.6 to 0.5, relative to small sites), with evidence for sustained year-on-year increases (by 14.6%, 0.1 to 31.3, relative to small sites; appendix pp 16–20). sites contributed a median 24 months (IQR 19–30, range 16–38) of data for all-cause 30-day mortality (in or out of hospital) after implementation (appendix p 46).
Analysis included 7160421 admissions (appendix p 47); 314313 (4.4%) people died within 30 days (2.6–7.2% across sites, median 4.6%, IQR 4.0–5.0). Overall, the ARK intervention was associated with a –2.7% (95% CI –5.7 to 0.3) immediate change in 30-day mortality odds (consistent with non-inferiority) and a 3.0% (–0.1 to 6.2) sustained change after versus before implementation (figure 3, 5). Sites with larger immediate mortality reductions tended to have larger sustained increases in mortality after versus before implementation (r=–0.28, p=0.082; appendix p 145). This inverse association suggests that the weak overall effects could be an artefact from the individual interrupted time series (appendix pp 51–127), given the substantial heterogeneity in trajectories of 30-day mortality before and after intervention (appendix p 146), potentially related to the varying capture of electronic health record data at sites over time (appendix pp 29–34). There was no evidence that effects on 30-day mortality were associated with implementation fidelity (immediate effect was 1.3% per additional fidelity criteria achieved, 95% CI –6.8 to 3.3; change in year-on-year trend after vs before implementation was –0.1%.

Figure 2: Intervention adherence during the first 12 weeks of implementation
Panels show the proportion of essential people who completed ARK training (A), staff who completed ARK training (B), staff who completed training per 100 acute beds (C), proportion of antibiotic prescriptions categorised using the decision aid at the initial prescription (D), proportion of antibiotic prescriptions reviewed versus baseline (E), and proportion of antibiotic prescriptions stopped at review and revise versus baseline (F). Sites are identified numerically by the order in which they were randomised to implement. The targets for 70% of essential people and more than 20 staff per 100 acute beds to complete ARK training are arbitrary but were prespecified for the funder as part of trial agreements and as part of prespecified fidelity criteria. Audit data were unavailable at two of 39 hospitals (ie, sites 31 and 38) and these sites are excluded from panels D–F. Six hospitals (sites 13, 18, 21, 27, 28, and 30) were missing baseline audit data and are therefore excluded from panels E–F.
Figure 3: Effects of the ARK intervention

Immediate effect at implementation (A) and effect on sustained year-on-year trend after versus before implementation (B). The top part of each panel shows the antibiotic primary (bold) and secondary outcomes, and the bottom part shows the clinical primary (bold) and secondary outcomes. Effects that were adjusted only for the effects of COVID-19 are shown in green, and fully adjusted effects are shown in red when there was evidence of an association or otherwise in blue.

ARK=antibiotic review kit. C difficile=Clostridioides difficile. DDD=defined daily dose. ICU=intensive care unit. IRR=incidence rate ratio (negative binomial regression). OR=odds ratio (logistic regression). SHR=subhazard ratio (competing risks regression). *Antibiotics are measured as DDDs per admission unless indicated otherwise. †Access or Watch depending on indication (analysed as a mutually exclusive category, since indication was unknown). ‡Modelled without a COVID-19 adjustment (effects plotted in green are therefore unadjusted).
There was no sustained increase in adjusted 30-day mortality, but this effect was not observed after excluding all follow-up from March 1, 2020 (appendix pp 13–15).

---

Figure 4: Effect on total antibiotic DDDs per admission
Overall immediate effect at implementation (A), overall effect on year-on-year trend after versus before implementation by implementation fidelity (C), and effect on year-on-year trend after versus before implementation by implementation fidelity (D). Sites are identified numerically by the order in which they were randomised to implement and are ordered by the number of fidelity criteria achieved (appendix pp 13–15). The size of the symbols in panels C and D reflects the precision of each estimate (inverse of the within-hospital variance). Weights are from random effects analysis. IRR—incidence rate ratio.
Figure 5: Adjusted 30-day mortality
Overall immediate effect at implementation (A), overall effect on year-on-year trend after versus before implementation (B), immediate effect at implementation by implementation fidelity (C), and effect on year-on-year trend after versus before implementation by implementation fidelity (D). Sites are identified numerically by the order in which they were randomised to implement and are ordered by the number of fidelity criteria achieved (appendix pp 13–15). Weights are from random effects analysis. The size of the symbols in panels C and D reflects the precision of each estimate (inverse of the within-hospital variance). OR=odds ratio.
We identified weak evidence for sustained reductions in 30-day mortality among sites that introduced the ARK categories into the prescribing process by implementation than among sites that did not (−7·2%, 95% CI −14·6 to 0·8), among sites implementing a hard stop than among sites implementing no soft or hard stop (−8·2%, −15·0 to −0·9), and among sites implementing in July–September than among sites implementing in January–March (−10·3%, −19·0 to −0·8; appendix pp 16–20).

We did not identify any evidence that sites with greater reductions in antibiotic DDDs per admission had larger immediate (r=0·044, p=0·79) or sustained (r=0·011, p=0·95) increases in 30-day mortality trends than did sites with smaller reductions in antibiotic DDDs per admission (figure 6).

Adjusted models showed no evidence of an immediate effect on total antibiotic DDDs per bed-day (−0·4%, 95% CI −3·2 to 2·5), with weak evidence of subsequent reductions (by −4·2% per year, −8·3 to 0·1; figure 3; appendix pp 13, 130), similar to effects on total DDDs per admission. At implementation, rates of broad-spectrum and Watch DDDs per admission decreased significantly (−6·3%, 95% CI −9·5 to −3·0 for broad-spectrum DDDs and −9·5%, −13·2 to −5·6 for Watch DDDs), as did quinolones (−15·5%, −21·9 to −8·5), whereas Access DDDs per admission increased slightly (4·4%, 0·3–8·8; figure 3; appendix pp 131), parenteral antibiotics (−0·9%, 95% CI −9·4 to −0·9; figure 3; appendix p 133), and oral antibiotics (−6·4%, −11·2 to −1·5; appendix pp 138–139).

We did not identify evidence of long-term effects on broad-spectrum antibiotics (−2·6%, 95% CI −8·5 to 3·6; appendix p 131), parenteral antibiotics (−0·9%, −5·2 to 3·6; appendix p 138), and antibiotics considered Access or Watch depending on indication (1·0, −5·2 to 7·6; appendix p 133). By contrast, year-on-year trends in DDDs per admissions increased faster after versus before implementation for carbapenems (12·3%, 2·3 to 23·2; appendix pp 140–144) with a similar trend for Reserve antibiotics (7·3%, −1·5 to 17·0; appendix pp 133–137), but from low levels and with wide CIs (table; appendix pp 48–50).

Mortality within 90 days of admission was 8·1% (571/737 of 7014694) overall (range 4·6–12·5 by site) and adjusted models showed weak evidence for an immediate decline (−3·9%, 95% CI −8·3 to 0·7) in 90-day mortality odds and evidence of a sustained year-on-year increase of 3·9% (0·5 to 7·4; appendix p 147). However, there was no evidence of a significant association after excluding admissions from the onset of the COVID-19 pandemic in March, 2020 (−3·9%, 95% CI −8·3 to 0·7 immediate decline in 90-day mortality odds; 3·2%, −1·5 to 8·2% sustained year-on-year increase; appendix pp 13–15).

Admission to critical care was uncommon (1·5% [97554 of 6499923] overall, range 0·4–4·2), and there was no evidence of an immediate (2·3%, 95% CI −1·9 to 6·7) or sustained implementation effect (−5·9%, −12·8 to 1·6; figure 3; appendix p 148). Similarly, we did not identify any evidence of association between length of stay (median 8·5 h, IQR 3·1–89·2) and the ARK intervention at implementation (−0·3% relative change in subhazard ratio per additional day in hospital, 95% CI −1·0 to 0·5) or year-on-year after versus before implementation (0·1%, −0·8 to 1·1; appendix pp 149–150). Emergency readmission to hospital (to any specialty) was 13·6% (903265 of 6625542; range 8·7–26·4) across sites, with no evidence of an immediate (−0·1%, 95% CI −2·6 to 2·5) or sustained implementation effect (−1·5%, −4·6% to 1·6; appendix p 151). Detection of *C difficile* infection (16475 [0·2%] of 6728437, range 0·1–0·6) and colonisation (27958 [0·5%] of 6092290, range 0·2–1·1)
within 90 days of admission was low. We did not identify any evidence of an intervention effect on C difficile infection (−4·6%, 95% CI −16·6 to 9·0) and colonisation (−5·6%, −16·9 to 7·2) within 90 days of admission at implementation, or after versus before implementation (5·8%, −8·1 to 21·9 and −6·9%, −17·5 to 5·0; appendix pp 152–153).

An exploratory analysis of the proportion of admissions with a stay longer than 48 h (215 245 [32·6%] of 679 4684, range 23·1–51·3) also showed no immediate (0·3%, 95% CI −2·6 to 3·2) or sustained (−1·2%, −5·5 to 3·2) implementation effect (appendix pp 13–15, 149–150). There were sustained reductions in year-on-year trends for quinolones (−13·8%, −22·5 to −4·0; appendix pp 140–144). We did not identify evidence of long-term effects on piperacillin–tazobactam (0·6%, −14·9 to 18·8; appendix p 142)

Discussion
Here, we have evaluated the ARK intervention,18,20 which aimed to safely reduce antibiotic consumption in adult acute general medical hospital admissions, in a stepped-wedge cluster-randomised trial. In our final model adjusting for COVID-19, the ARK intervention resulted in mean reductions in antibiotic use of 4·8% per year, but no immediate reduction. That the intervention changed prescribing over time rather than suddenly might be expected, given the different components, including training for use of the novel decision aid and audit and feedback to re-enforce learning.21 The change over time could also reflect increasing acceptance that completion of arbitrary antibiotic courses might not reduce risk of resistance.22 Although the trial was powered to detect a 15% immediate reduction associated with the intervention, the effect observed is potentially clinically significant given that the national standard contract for acute trusts in England sought a reduction of only 1% per year. Given the importance of sustainable effects from behaviour change interventions in antibiotic stewardship, it is notable that this reduction was seen over a median of 23 months (range 14–37). Notably, consistent reductions were seen in Access, Watch, narrow-spectrum, and oral antibiotics, but not in broad-spectrum or parenteral antibiotics, antibiotics considered Access or Watch depending on indication, and piperacillin–tazobactam, and there was a significant increase in DDDs for carbapenems and Reserve classes. Since the intervention was targeted at acute general medical admissions, unsurprisingly its effect was seen in narrow-spectrum and Access agents, which are typically used as first-line medication or for de-escalation. The significant increase in carbapenem use after intervention could suggest that decreased use of one set of agents increased use of others. This effect seems unlikely, because the differences measured are relative and the absolute increases are small (appendix pp 49–50). Furthermore, use of carbapenems increased disproportionately across the NHS during the study period, driven by their inclusion in national treatment guidelines for hospital-acquired pneumonia, shortages of piperacillin–tazobactam, and increasing resistance to other agents.23 Broad-spectrum agents, such as carbapenems, are typically prescribed when other agents have already been tried for the patient or when microbiology has identified a specific pathogen, and we might simply have observed an increase that the intervention would not be expected to affect.

We found no overall relationship between fidelity of implementation and the effect of the intervention. An absence of relationship might be because complex interactions between intervention elements and the implementation setting are difficult to measure quantitatively in a large-scale trial, or because we took an average of how many fidelity criteria were met, but some of the criteria were likely to have had more of an effect on fidelity than others. The ARK audit tools were designed to support frequent, light-touch feedback to prescribers, sometimes called handshake stewardship,24 which relies on interpersonal factors that we could not analyse but will be considered in forthcoming mixed-methods process analyses. Prescription audits began 12 weeks before intervention to generate baseline data for the intervention’s feedback element, so it is perhaps not surprising that rates of audit completion were generally higher before implementation than afterwards. Notably, among individual intervention components, implementing the decision aid into the prescribing process and greater uptake of the online learning were both associated with greater reductions in antibiotic use than were not implementing these elements of the intervention, suggesting that these are key elements in achieving sustained change.

The ARK intervention focuses on decisions to stop rather than decisions to start antibiotics, because this approach has the potential to reduce overall use without withholding empirical antibiotics from patients with acute illness. Nevertheless, we considered it important to evaluate whether introducing ARK was associated with excess mortality. Beginning in March, 2020, when 12 of 39 sites were still within 12 months of implementation, the COVID-19 pandemic was associated with substantial increases in mortality among acute hospital admissions (appendix pp 51–127). Adjusting for this effect, in most of the main models and through sensitivity analysis excluding these 12 sites, we identified no clear evidence of associations between the intervention and 30-day or 90-day mortality. Notably, implementing the decision aid with a hard stop of antibiotic prescriptions at 72 h if not revised was associated with decreased risk of death over time, despite prescribers reporting anxiety that hard stops could compromise clinical outcomes.27 This decrease might be explained by clinicians placing a greater emphasis on prescription reviews at sites that introduced hard stops, improving patient management more broadly. Furthermore, we found no evidence that
sites that achieved greater reductions in antibiotic DDDs per admission had larger increases in mortality than did sites with smaller reductions in antibiotic DDDs per admission (figure 6).

Our study has important limitations. First, there are intrinsic limitations of the cluster-randomised design. Although we included over a quarter of all acute hospitals in the UK health system, we cannot reliably exclude imbalance, particularly of time-dependent factors, as emphasised by the onset of the COVID-19 pandemic during the postimplementation period. There could be imbalance in other time-dependent organisational changes (e.g., staffing, clinical or stewardship practice, or case-mix), which might have changed antibiotic consumption at individual sites. We do not have data for antibiotic resistance rates, which might have varied between sites over time and are generally lower in the UK than in many other countries.

Second, although sites were robustly randomised with respect to the timing of intervention implementation, they might not be a random sample of UK acute hospitals. It is plausible that only sites with well constituted antimicrobial stewardship teams volunteered, and other sites might not see the same effect, particularly as effect was associated with some aspects of intervention fidelity. Alternatively, the intervention effect could be greater at sites with weaker stewardship teams.

Third, we measured antibiotic consumption indirectly from dispensing data to clinical areas, as individual-level antibiotic data could be provided by only four sites. This method means that we cannot explore mechanisms through which the intervention reduced antibiotic use. However, our mortality analysis included over 7 million admissions, so there was no ability to collect individual prescribing data other than electronically. Although richer data at the individual-patient level would have allowed more detailed exploration, collecting consent and antibiotic use data from the number of patients needed to conduct a robust analysis would be infeasible. Furthermore, stewardship interventions, such as ARK, are made at the organisation level and, as such, organisation-level antibiotic use is an appropriate outcome.

Fourth, it is probable that not all prescribing decisions in the patient population analysed were subject to the intervention (e.g., outlying surgical patients). Conversely, some patients for whom prescribing decisions were not subject to the intervention might have been included in analysis. These inclusions and exclusions are because acute general medical inpatients are not easily identified in electronic admission data, and we had to infer this population from specialty codes, which are used slightly differently across sites. Importantly, both these effects, and low implementation fidelity at some sites, would be expected to dilute the observed effect of the intervention on antibiotic use, suggesting that antibiotic reductions might have been even greater in targeted patients and in sites with high implementation fidelity.

In terms of potential clinical harms from the intervention, analysing routinely available electronic health records, we identified no consistent evidence of effect on mortality, admission to critical care, length of stay, or readmission. Although we cannot exclude the possibility of other harms related to shorter antibiotic treatment, our overall findings make substantial increases in treatment failure and recurrence unlikely. Equally, we were not able to measure potential direct benefits from reduced antibiotic treatment, but it is a reasonable assumption that reductions in antibiotic exposure will reduce antibiotic-associated harms, including resistance.

Despite its limitations, the cluster-randomised approach that we adopted allowed us to capture both the organisation-level effects of the intervention on antibiotic consumption and the patient-level effects on clinical outcomes. Our findings are entirely consistent with the three, much smaller, previous trials of hospital stewardship interventions, which showed the importance of intervention co-design with practitioners, educational theory and clinical relevant audit and feedback to clinicians. Our findings are also consistent with conclusion of the most recent Cochrane review that stewardship interventions can reduce unnecessary antibiotic use safely. Our approach to intervention design and evaluation addresses many of the limitations that have prevented the translation of previous research findings into hospital practice. Crucially, the wider ARK-Hospital programme has delivered practice-ready materials for implementation, which are freely available. Acute hospital providers should consider embedding the ARK-Hospital toolkit in their staff training, prescribing processes, and stewardship work to reduce antibiotic overuse in acute general medical inpatients and protect these patients from antibiotic-related harms.

Contributors
TEAP, ASW, IW, and MJL conceived the research. ELAC, KS, SW, MSa, AK, FM, KSH, DWC, LV, SH LY, TEAP, and ASW conceived and developed the intervention. ML-S, RA, SB, PC, GC-B, SD, ME, RF, KJF, VG-A, SG, CG, KG, CH, DH, TH, SI, AJ, NJ, PK, GK, DMac, CM, DMacw, BM, MM, RM, SN, AN, JO, AO, RP, NP, DP, ES, MSl, BS, CW, JW, MD, and MJL conducted the trial. EPB conducted the statistical analysis. EPB and ASW accessed and verified all the data. ASW, EPB, and MJL wrote the first draft. All authors reviewed and approved the final manuscript. All authors had full access to the data in the study and had final responsibility for the decision to submit for publication.

Declaration of interests
MJL, DWC, IW, TEAP, and ASW declare funding from the National Institute for Health Research (NIHR) for the ARK-Hospital programme. ASW is an NIHR Senior Investigator. All other authors declare no competing interests.

Data sharing
The de-identified patient-level electronic health records (on over 7 million admissions) and hospital-level antibiotic use data used for this analysis was obtained from individual hospital organisations without permission for onward data sharing. It can be accessed either directly from the participating organisations or through the trial team if the participating organisations provide permission. De-identified patient-level admission data can also be accessed directly through an application to NHS Digital. All enquiries should be sent to Prof Martin J Llewelyn.
Acknowledgments

The trial was funded by the NIHR under its Programme Grants for Applied Research Programme (reference number RP-054-20015). ASW and LY are NIHR Senior Investigators. ASW, DYC, TEAP, MJL, and JSJR are supported by the NIHR Biomedical Research Centre, Oxford, UK. LY is partly supported by NIHR Applied Research Collaboration-West and NIHR Health Protection Research Unit for Optimizing design of research to evaluate antibiotic stewardship strategies for respiratory tract infections in primary care: a systematic analysis. Lancet 2022; 399: 629–55.

References


