THE ISRAEL NATIONAL INSTITUTE FOR HEALTH POLICY RESEARCH

Title of research (in English)

External validity of randomized control trials

Title of research (in Hebrew)

תוקף חיצוני של מחקרים השוואתיים בהקצאה אקראית (Randomized controlled trials)

Section 4: Executive Summary- Hebrew

1. רקע מדעי: הוכחות מוצקות מראות כי מחקרים השוואתיים בהקצאה אקראית - Randomized controlled trials (RCTs) המתבצעים בשיטות אופטימליות, מהווים את מערך המחקר החופשי ביותר מהטיות. מן הראוי כי קבלת החלטות בתחום מדיניות בריאות תתבסס על העדויות המוצקות ביותר. בעוד שמחקרים רבים עסקו בשיפור התוקף הפנימי של ניסויים מבוקרים אקראיים (RCTs), מעט עבודות נעשו לשיפור תוקפם החיצוני [1-2].
התוקף החיצוני של אוכלוסיית המחקר מסתמך ראשית על קריטריוני הכללה ואי הכללה. שנית, על אוכלוסיית החולים שגויסו בפועל. יש להגדיר קריטריוני הכללה ואי הכללה באופן מדויק, ברור וחד משמעי [2]. מחקרים הראו כי חולים שגויסו ל- RCT היו לפעמים שונים מאלה שהיו מתאימים להכללה אך לא גויסו. חולים שלא יכלו לספק הסכמה מדעת, ולכן לא נכללו, לרוב סבלו ממחלה קשה יותר ותוצאותיהם היו לרוב גרועות יותר בהשוואה לחולים שנכללו בניסויים [3-4]. על מנת ליישם את תוצאות המחקר, עלינו להיות מסוגלים להעריך את תוקפו החיצוני.

2. מטרות המחקר: להעריך את הגורמים אשר עשויים להשפיע על טיב התוקף החיצוני של ארבעה מחקרים קלינים בתחום של מחלות זיהומיות. בסקירה השיטתית המטרה הייתה לבדוק האם מחקרים עם שיעור גבוה יותר של משתתפים הביאו לאומדנים מוטים של התוצא הראשי בהתאם להיפותזה שנחקרה.

3. שיטות המחקר: נערכה השוואה בין מטופלים שנכללו ב-4 RCTs למטופלים שלא נכללו באמצעות ארבעה מחקרים תצפיתיים, המחקרים יכונו מעתה כ- מחקר Colistin, מחקר MRSA, מחקר UTI ומחקר GNB. המטופלים בזרוע התצפיתית התאימו להכלל לפי קריטריוני ההכללה של המחקר אך לא נכללו מסיבות שונות. איתור המטופלים שלא נכללו נעשה במקביל לגיוס המטופלים

שנכללו בתקופת ה- RCTs. ההשוואה בין הקבוצות כללה נתונים דמוגרפים, מאפייני רקע, מאפייני זיהום ותוצאים. כמו כן נאספו נתונים אודות הסיבות לאי הכללה. המחקר התבצע במרכז הרפואי רבין, בבית החולים בילינסון בפתח תקווה ובביה"ח רמב"ם בחיפה. ההשוואה בין קבוצות החולים התבצעה בעזרת השיטות הסטטיסטיות המקובלות.

בסקירה ספרותית שיטתית שערכנו, נסקרו אקראית 184 RCTs בתחום של מחלות זיהומיות, סרטן וסכרת שפורסמו בשנת 2017. נבחנה טיב מתודולוגית המחקרים בין מחקרים שדיווחו על מספר המטופלים שהיו מועמדים להכללה לאלו שאינם. בנוסף נבחן הקשר בין אחוז גיוס גבוה של מטופלים לתוצא חיובי מוטה.

4. ממצאים: בשלושה מתוך ארבעה מחקרים (MRSA, GNB, Colistin) נכללו מטופלים מאושפזים עם זיהומים חמורים. על אף הדמיון בין האוכלוסיות בשלושת המחקרים, נצפו תוצאות שונות במחקרים התצפיתיים. במחקר Colistin, חולים שנכללו במחקר הקליני היו דומים לאלו שאינם נכללו במאפייני הבסיס שלהם, אם כי חולי ה-RCT הראו הבדלים קלים כלפי זיהום חמור יותר. חולים שלא נכללו במחקרי ה- MRSA וה- GNB היו בעלי תחלואות רקע משמעותיות יותר וסבלו מזיהומים קשים בהשוואה לאלו שגויסו ל- RCT.

ההבדל העיקרי בין מחקרים אלה היה תהליך ההסכמה. במחקר COLISTIN, בניגוד ל - MRSA ו- GNB, קיבלנו אישור מועדות האתיקה המקומיות לגייס חולים שלא היו מסוגלים לספק הסכמה מדעת ולא היה להם אפוטרופוס חוקי, בהסכמת רופא בלתי תלוי. בעוד שקבלת הסכמה מאפוטרופוס עלולה לעכב את גיוס החולים, הסכמה ע"י רופא בלתי תלוי אפשרה הכללה מיידית ומקיפה של חולים קשים המאפיינים את אוכלוסיית היעד.

בבחינת התוקף החיצוני של RCTs, אנו מודאגים מכך שבדרך כלל האוכלוסייה שנכללה במחקר צעירה יותר ובעלת פחות מחלות רקע מאשר האוכלוסיה שלא נכללה. במחקר UTI הצלחנו להראות שההפך הוא הנכון: מטופלות שהשתתפו במחקר הקליני היו מבוגרות יותר ובעלות היסטוריה של הישנות דלקות בשתן. בנוסף, שיעור הכשל הקליני היה נמוך משמעותית בקרב המטופלות שלא נכללו. ניתן להסביר ממצאים אלה בכמה דרכים. ראשית, אוכלוסיית היעד של מחקר זה היא נשים

צעירות ובריאות בדרך כלל. ברוב המקרים מטופלות אלה לא פנו לרופא לטיפול ב- UTI. נתון זה הוכח גם בשיעורים נמוכים יותר של ביצוע תרביות שתן בקרב מטופלות שלא נכללו. שני גורמים אלה מערערים את חשיבות התוצא הראשי שנבדק ב- RCT לאוכלוסיית המטרה. שנית, היו הבדלים ביו אופי המעקב אחר קבוצות המחקר. שלישית, מטופלות המחקר הקליני הוקצו באופן אקראי לקבלת ניטרופורנטואין או לפוספומיצין, מטופלות שלא נכללו במחקר, טופלו בעיקר בפלואורוקווינולונים הנחשבים כיעילים יותר קלינית ומיקרוביולוגית אך אינם מומלצים לשימוש בהנחיות המקובלות הממליצות להגביל את השימוש בהם מסיבות של בטיחות.

בסקירה הספרותית, מצאנו שמחקרים אשר לא דיווחו על מספר המטופלים שהיו מועמדים להכללה היו בעלי מאפיינים מתודולוגיים חלשים. מחקרים אשר הכלילו אחוז גבוה מאוד מהמטופלים שנסקרו (>90%) העריכו את ההתערבות כמועילה פי שתיים (ע"פ התוצא הראשי) ממחקרים אשר הכלילו אחוז נמוך יותר מהמטופלים.

5. מסקנות: כצפוי, נצפו הבדלים בתוקף החיצוני של האוכלוסייה בין ארבעת המחקרים אותם בחנו. המחקרים נבדלו הן באוכלוסיית היעד אליה פנו והן בחומרת הזיהום. התוצאות שלנו מעלות שאלה בדבר הכללת תוצאות שנצפו ב- RCTs בתחום של מחלות זיהומיות. אי הגבלה של קריטריוני הכללה והאפשרות לגייס את החולים הקשים ביותר למחקרים קלינים הם פקטורים חשובים לשיפור תוקפם החיצוני. הידיעה על שיעור גיוס מטופלים למחקר חשובה להערכת אפקט ההתערבות ותוקפו החיצוני של המחקר.

6. השלכות למדיניות והמלצות למקבל ההחלטות

- תוקף חיצוני של RCTs נפגע מהליך ההסכמה הנוכחי. בהליך זה, לא ניתן להכליל את החולים הקשים ביותר שלא ע"י אפוטרופוס מה שמביא לעיכובים בתחילת ההתערבות המחקרית. אחד הפתרונות האפשריים שהוצג במחקר זה הינו היכולת לגייס מטופלים ע"י רופא בלתי תלוי. באמצעות הליך זה מתאפשרת הכללתם של חולים קשים ותחילת ההתערבות ללא עיכוב.

- דיווח אודות נתוני קבוצת החולים שלא נכללו (אך היו יכולים להכלל על פי קריטריוני הכללה ואי הכללה) ישפר את הבנתנו בנוגע לתוקף החיצוני של המחקרים.
- העדר דיווח אודות מספר המטופלים שהיו מועמדים להכללה אך לא נכללו ועל הסיבות לכך עלול להוביל להטיות במחקר. דיווח זה דרוש בבסיסו של כל פרסום של RCT.

מקורות (לא יותר מ-4)

1. Green LW, Glasgow RE. Evaluating the relevance, generalization, and applicability of research: issues in external validation and translation methodology. Evaluation & the health professions. 2006 Mar; 29(1):126-53.
2. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMC medicine. 2010 Dec; 8(1):18.
3. Rothwell PM. Factors that can affect the external validity of randomised controlled trials. PLoS clinical trials. 2006 May 19; 1(1):e9.
4. Paul M, Bronstein E, Yahav D, Goldberg E, Bishara J, Leibovici L. External validity of a randomised controlled trial on the treatment of severe infections caused by MRSA. BMJ open. 2015 Sep 1; 5(9):e008838.

Section 5: Comprehensive scientific report - English

הנחיות:
הדוח המדעי המלא ייכתב **באנגלית** בצורה המקובלת של מאמר בעיתון מקצועי ולפי הסעיפים הבאים ככותרות משנה:

1. **Scientific background** (what is already known)
2. **Research objectives** (what does the work intents to solve)
3. **Methodology**
4. **Findings**
5. **Discussion and Conclusions**
6. **Policy Implications and Recommendations**
7. **References**

לתשומת לב
- הפירוט בסעיפים 1 – 6: לא יעלה על **עשרים** עמודים (כולל תרשימים)
- פונט: Arial מספר 12
- מרווח בין שורות – 1.5

הדוח המדעי המלא ישמר בארכיון המכון ויימסר לצד ג' רק לאחר קבלת הסכמה מפורשת של החוקר האחראי.

1. Scientific background

Randomized controlled trials (RCTs) are the gold standard for guidelines and evidence-based medicine. Internal validity of an RCT reflects the strengths to support a clinical decision based on study results and the extent to which the results are influenced by bias. Adequate randomization, allocation concealment, blinding, non-selective reporting of outcomes and intention-to-treat analysis, have been identified as important factors in study design to minimize bias in RCTs and increase internal validity [1-2]. External validity is defined as the extent and manner in which the results of an experimental study can be generalized to different subjects and settings. It has two components: population validity, the extent to which the results can be generalized from the specific sample to a defined population, and ecological validity, the extent to which the results can be generalized from a set of environmental conditions created by the researcher to other environmental conditions/settings [3].

The population external validity of RCTs relies firstly on the inclusion and exclusion criteria. Secondly, it relies on the population of patients actually recruited. Inclusion and exclusion criteria should be defined precisely, clearly and unambiguously [2].

Studies have shown that patients recruited into RCTs were sometimes different from those who were eligible but not recruited in terms of age, gender, educational status, socioeconomic status, place of residence, ability to provide informed consent and severity of disease. Patients that could not provide informed consent, and thus were not included, had more severe disease and their outcome was often worse compared to patients included in trials [4-6]. The problem of external validity is particularly relevant to registration trials, which typically specify numerous exclusion criteria. In order to apply a study's results, one should be able to assess its population external validity; however, few studies to date have done so [7-12].

We measured factors that might influence external validity of four pragmatic, investigator-initiated RCTs comparing: (i) fosfomycin vs. nitrofurantoin for treatment of uncomplicated lower urinary tract infection in female adults at increased risk of antibiotic-resistant bacterial infection. (ii) Long vs short duration of antibiotic treatment of Gram-negative bacilli bacteremia (iii) Colistin alone vs. colistin plus meropenem in patients with severe infections due to multi-drug resistant bacteria (iv) Trimethoprim-sulfamethoxazole versus vancomycin for severe infections caused by methicillin resistant Staphylococcus aureus [5]. In addition, we performed a systematic review of RCTs in the field of infectious diseases, diabetes and cancer.

2. Research objectives

We aimed to measure factors that might influence external validity by comparing patients included in the four mentioned clinical trials to patients that were candidates for inclusion but were not included (in each study separately). In the systematic review we aimed to explore whether studies with a higher rate of included participants resulted in biased estimates of beneficial intervention effect.

3. Methodology

Part I: observational cohort studies.

We compared patients randomized in each trial (interventional arm) to those fulfilling clinical and microbiological inclusion criteria who were not randomized due to exclusion from the trial (observational arm) during the RCTs recruitment period.

1. Fosfomycin vs. nitrofurantoin for treatment of uncomplicated lower urinary tract infection in female adults at increased risk of antibiotic-resistant bacterial infection (referred as "UTI trial").

   Study design and patients

   In an investigator-initiated multinational, open-label RCT adult women with lower UTI were randomized to nitrofurantoin or fosfomycin between 2013-2017 in Switzerland, Poland, and Israel [13]. In the observational study we compared women who were screened for enrolment in the RCT but excluded, to women who participated in the RCT- both groups in Israel.

   Excluded patients were included in the cohort if they had been excluded from the trial for the following reasons: logistics (staff unavailable for recruitment on evenings, weekends and holidays); army service precluding follow-up visits, antibiotic use in the preceding 4 weeks, pregnancy or lactation.

   Outcomes

   The primary outcome was incidence of emergency department index visits resulting in hospitalization within 28 days.

2. Long vs. short duration of antibiotic treatment of Gram-negative bacilli bacteremia (referred as "GNB trial").

   Study design and patients

   In investigator-initiated multinational open-label noninferiority RCT, adult inpatients with gram-negative bacteremia were randomized to receive 7 days (short treatment) or 14 days (long treatment) of covering antibiotic therapy between 2013-2017 in 3 centers in Israel and Italy [14].

   Excluded patients were included in the cohort if they had been excluded from the trial for the following reasons: participation in another trial, early discharge, patient/guardian refusal to participate or unwillingness of the treating physician to include the patient.

   Outcomes:

   The primary outcome at 90 days was a composite of all-cause mortality; clinical failure, including either relapse of the bacteremia, local suppurative complications, or distant complications; and readmission or extended hospital stay (>14 days) as defined in the RCT.

3. Colistin alone vs. colistin plus meropenem in patients with severe infections due to multi-drug resistant Gram-negative bacteria (referred as "colistin trial").

   Study design and patients

   An investigator-initiated multinational, open-label RCT was conducted between 2013-2017 in Greece, Israel and Italy. The RCT compared colistin-meropenem combination therapy to colistin monotherapy in the treatment of patients infected with carbapenem-resistant Gram-negative bacteria (CR GNB). Infections included bacteraemia, definite ventilator associated or hospital-acquired pneumonia, probable ventilator-associated pneumonia, and urosepsis [15]. Excluded patients were included in the cohort if they had been excluded from the trial for the following reasons: no consent; receipt of colistin for >96 hours before assessment for eligibility; and prior inclusion in the RCT.

   Outcomes

   The primary outcome was clinical failure at 14 days after the first positive culture was obtained. The outcome was a composite of: patient deceased, systolic blood pressure<90 mmHg or the need for vasopressor support, no stability or improvement in Sequential Organ Failure Assessment (SOFA) score, and for patients with bacteremia due to growth of the initial isolate in blood cultures taken on day 14.

4. Trimethoprim-sulfamethoxazole versus vancomycin for severe infections caused by methicillin resistant Staphylococcus aureus (referred as "MRSA trial").

   Study design and patients

   An open-label RCT conducted in four medical centers in Israel, included hospitalized patients with documented or highly probable invasive MRSA infections who were randomized to vancomycin versus trimethoprim-sulfamethoxazole (TMP-SMX) treatment, between 2007 and 2014 [16]. Excluded patients were included in the cohort if they had been excluded from the trial for the following reasons: no consent, meningitis, left-sided endocarditis, severe neutropaenia, chronic renal dialysis or treatment with study medications for longer than 48 hours.

Outcomes:

The primary outcomes were clinical failure at day 7 and 30-day mortality.

Statistical analysis

Data were expressed as frequencies (percentages) for categorical variables, mean ± standard deviation for normally distributed continuous variables and as median and interquartile range (IQR, 25–75 percentiles) for non-normally distributed continuous variables.

Univariate analysis was conducted for all independent variables for the comparison between patients included in the RCT to those who were excluded using the t-test or Mann–Whitney U-test (as appropriate based on their distribution) for continuous variables. The Chi-square test was used for categorical variables. Statistical analyses were performed using IBM SPSS Statistics.

Part II: Systematic review.

Eligibility criteria:

The main criterion for inclusion was randomized controlled trials (RCTs) that were published in 2017 in the field of; cancer, infectious diseases, and diabetes.

Search methods for identification of studies:

The search was restricted to humans and to studies that included adult patients (age ≥ 18). We did not restrict by type of intervention. Since the search yielded an extremely high number of studies, more than 18,000 articles, we have randomly reviewed 184 manuscripts, 55 randomized controlled trials on infectious diseases (representing acute diseases), 33 randomized controlled trials on diabetes (representing chronic diseases), and 85 randomized controlled trials on cancer (representing both).

Our search phrase include the Cochrane filter for randomized controlled trials and Medical Subject Headings (MESH) categories in the field of cancer, infectious diseases and endocrine system disease:

((((randomized controlled trial [pt] OR controlled clinical trial [pt] OR randomized [tiab] OR placebo [tiab] OR clinical trials as topic [mesh: noexp] OR randomly [tiab]

OR trial [ti]) NOT (animals [mh] NOT humans [mh]))))) AND ("Bacterial Infections and Mycoses"[Mesh] OR "Virus Diseases"[Mesh] OR "Neoplasms"[Mesh] OR "Endocrine System Diseases"[Mesh]).

Study selection:

Studies were reviewed and appraised by 2 investigators independently. On first screening, we reviewed the titles of the yielded articles and, when relevant, we proceeded to reading the abstracts. For the second screening, we read the full text and reviewed the reference lists.

Data extraction:

The following data were extracted: type of primary hypothesis (superiority, equivalence or non-inferiority), description of the primary hypothesis, number of included and non-included patients, study design characteristics (sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, sample size calculation and early termination), type of intervention (invasive vs. non-invasive procedure, drug vs. placebo, drug vs. device vs. other intervention), description of the intervention, intervention effects, primary outcomes, type of primary outcome (divided to: as all-cause mortality, other objectively assessed and subjectively assessed).

Statistical analysis:

Intervention effects were modeled as positive or negative effect according to the study hypothesis. We compared methodological characteristics between studies that reported the number of patients assessed for eligibility to studies that did not report the number of patients assessed for eligibility. We also compared methodological characteristics between randomized controlled trials demonstrating beneficial intervention effects versus randomized controlled trials demonstrating non-beneficial intervention effects. A sub-group analysis of methodological characteristics between randomized controlled trials demonstrating beneficial intervention effects in the fields of infectious diseases, diabetes, and cancer was conducted.

4. Findings

Part I: Observational cohort studies.

1. Fosfomycin vs. nitrofurantoin for treatment of uncomplicated lower urinary tract infection in female adults at increased risk of antibiotic-resistant bacterial infection.

   Of 288 adult women with symptoms of lower UTI screened for enrolment in the RCT, 127 women were enrolled and randomized while 110 women were included in the observational cohort. Patients with suspected upper UTI or immunocompromised patients were not included in this study. In the RCT, logistic difficulties and prior antibiotic use were the main reasons for exclusion [45.4% (50/110), 43.6% (48/110), respectively], followed by army service [3.6% (4/110)] and pregnancy or lactating [7.3% (8/110)].

   Patient characteristics

   Trial participants tended to be older than those excluded but this difference did not reach statistical significance [39 years (IQR 29-59) vs. 35.5 years (IQR 24-56.25); P= 0.073]. No significant differences between study groups were demonstrated for any background comorbidities (table 2). History of recurrent UTI was more common among trial participants [32.2% (41/127) vs. 13.6% (15/110); P<0.001].

   Infection characteristics and management

   Trial participants had more symptoms of lower UTI including urgency, frequency, or suprapubic tenderness compared to those excluded (table 1). Flank pain, a common symptom of upper UTI, was more frequent among excluded patients (40% [44/110] vs. 26.8% [34/127]; P=0.031).

   RCT patients were assigned to receive either fosfomycin or nitrofurantoin. Among excluded patients, only 15 were treated with one of the study drugs; nine were treated with nitrofurantoin and six with fosfomycin (8.2%, 5.5%, respectively). Remaining excluded patients were mostly treated with fluoroquinolones (55/110, 50%), though one in five (23/110) did not receive any antibiotic (Table 1).

While rates of positive urine cultures were similar among included and excluded patients (68% [73/107] vs. 64% [35/55]), only half of the excluded patients underwent urine culture (table 1).

Table 1. UTI trial - baseline patient characteristics

	RCT - Excluded N=110 (%)	RCT - Included N=127 (%)	P value
Age (median, IQR)*	35.5 (24-56.25)	39 (29-59)	0.073
Background comorbidities			
Metabolic syndrome	25 (22.7)	33 (26)	0.561
Cardiovascular disease	8 (7.3)	11 (8.7)	0.695
Cancer	1 (0.9)	3 (2.4)	0.386
Mental illness	4 (3.6)	5 (3.9)	0.904
Lung diseases	3 (2.7)	8 (6.3)	0.192
Gastro intestinal disorders	4 (3.6)	4 (3.1)	0.836
Autoimmune diseases	13 (11.8)	10 (7.9)	0.306
Renal diseases	5 (4.5)	4 (3.1)	0.575
History of recurrent UTI	15 (13.6)	41 (32.3)	<0.001
Symptoms			
Dysuria	82 (74.5)	98 (77.2)	0.638
Frequency	35 (31.8)	102 (80.3)	<0.001
Urgency	36 (32.7)	102 (80.3)	<0.001
Suprapubic tenderness	57 (51.8)	101 (79.5)	<0.001
Flank pain	44 (40)	34 (26.8)	0.031
Lower back pain	16 (14.5)	61 (48)	<0.001
Nausea	18 (16.4)	41 (32.3)	0.005
Vomiting	6 (5.5)	12 (9.4)	0.247
Fever (subjective)	20 (18.2)	13 (10.2)	0.078
Chills	19 (17.3)	34 (26.8)	0.08
Gross hematuria	36 (32.7)	33 (26)	0.254
Urine culture			
Negative	20/55 (36)	34/107 (32)	<0.001
Positive	35/55 (64)	73/107 (68)	
Not collected	55 (50)	20 (15.7)	
Antibiotic treatment			
Fosfomycin	6 (5.5)	64 (50.4)	<0.001
Nitrofurantoin	9 (8.2)	63 (49.6)	
Fluoroquinolones	55 (50)	0 (0)	
Cefuroxime	13 (11.8)	0 (0)	
TMP/SMX	4 (3.6)	0 (0)	
none	23 (20.9)	0 (0)	
Outcomes			
ESBL	7 (6.4)	5 (3.9)	0.395
Index visit resulted in hospitalization	15 (13.6)	4 (3.1)	0.003

Emergency department visits within 28 days.	12 (10.9)	22 (17.3)	0.16
Bacteriologic failure – bacteriuria recurrence	0 (0)	10 (7.9)	0.003
Clinical failure - Cystitis recurrence	2 (1.8)	33 (26)	<0.001

Data are presented as number. (%) unless otherwise indicated.
Abbreviations: RCT – randomized controlled trial; IQR – interquartile range; UTI – urinary tract infection; TMP/SMX – trimethoprim/sulfamethoxazole; ESBL – extended spectrum beta-lactamase.
*Age-per one year increment

Outcomes

Clinical failure at 28 days occurred in 1.8% (2/110) of excluded patients vs. 26% (33/127) of RCT patients, P<0.001. Bacteriologic failure also differed significantly between groups, with no events of bacteriuria recurrence documented in the excluded group vs. 10 cases (7.9%, n=127) reported in the RCT group (P=0.003) (table 1).

Among excluded patients, 13.6% (15/110) were hospitalized in the 28 days following screening compared to only 3.1% (4/127) of patients who were included in the RCT, P=0.003. Each patient was hospitalized once during follow up. All hospitalizations in both groups were for reasons other than UTI. In the excluded group, all hospitalization events were in the group of 23 patients who did not receive any antibiotics. The incidence of ED visits within 28 days was similar among included and excluded patients (table 1).

2. Long vs short duration of antibiotic treatment of Gram-negative bacilli bacteremia

Refusal to provide informed consent by the patient or legal guardian was the main reason for exclusion from the RCT [32.5% (199/613)] (table 2).

Table 2. GNB trial - reasons for exclusion of patients from the RCT (and inclusion in the observational study)

Reason for exclusion	**N (%) (Total N=613)**
Participation in other trial	122 (19.9)
Early discharge	180 (29.4)
Patient/ guardian refusal to participate	199 (32.5)
Treating physician unwillingness	112 (18.3)

Data are presented as no. (%)

Patients' characteristics:

Patients recruited into an investigator-initiated RCT were significantly different from patients not included and treated in usual clinical practice. Excluded patients differed from included patients in their functional and cognitive status. Almost 50% of excluded patients (288/613) were non independent at baseline compared to 37.7% of RCT participants (228/604). Patients with dementia, hemiplegia or cancer were significantly less represented in the RCT (table 3).

Table 3. GNB trial - baseline patient characteristics

Variable	Excluded from randomized controlled trial N=613, (%)	Included in randomized controlled trial N=604, (%)	P value
Demographics and background			
Age, y, median (IQR)	73 (61.5-82)	71 (61-80)	0.172
Gender, male	334 (54.5)	285 (47.2)	0.011
Residency - nursing home	66 (10.8)	28 (4.6)	<0.001
Functional capacity at baseline: needs assistance/dependent in ADL or bedridden	288 (47)	228 (37.7)	<0.001
comorbidities			
Congestive heart failure	41 (6.7)	59 (9.8)	0.050
Chronic pulmonary disease	50 (8.2)	82 (13.6)	0.002
Dementia	36 (5.9)	22 (3.6)	0.068
Hemiplegia	40 (6.5)	12 (2)	<0.001
Malignancy	230 (37.5)	191 (31.6)	0.031

Data are presented as no. (%) unless otherwise indicated.
Abbreviations: IQR, interquartile range; ADL, activities of daily living.

Infection characteristics:

At infection onset, excluded patients were significantly more ventilated and carried more catheters than those included in the RCT. Patients excluded from the RCT were more likely to acquire their infection in the hospital [53.3% (327/613) vs. 29.1% (176/604), respectively, P<0.001] and less likely to have urinary tract infection [322 (52.5%) vs. 411 (68%), P<0.001]. Excluded patients were more infected with Acinetobacter spp. than those included in the RCT (table 4).

Table 4. GNB trial - infection characteristics and management

Variable	Excluded from Randomized controlled trial N=613, (%)	Included in Randomized controlled trial N=604, (%)	P value
Devices prior to infection			
Endotracheal tube	54 (8.8)	16 (2.6)	<0.001
Urine catheter	218 (35.6)	100 (16.6)	<0.001
Central venous catheter	91 (14.8)	41 (6.8)	<0.001
Peripheral catheter	322 (52.5)	148 (24.5)	<0.001
Nasogastric tube	66 (10.8)	16 (2.6)	<0.001
Infection Characteristics			
Bacteria type			
Escherichia coli	312 (50.9)	380 (62.9)	<0.001
Klebsiella spp	116 (18.9)	80 (13.2)	
Other Enterobacteriaceae	73 (11.9)	83 (13.7)	
Acinetobacter spp	41 (6.7)	6 (1)	
Pseudomonas spp	54 (8.8)	48 (7.9)	
Other	17 (2.8)	7 (1.2)	
ESBL	138 (22.5)	99 (16.4)	0.007
Source of bacteremia - UTI	322 (52.5)	411 (68)	<0.001
Appropriate empirical therapy	521 (85)	502 (83.1)	0.371
Hospital-acquired infection	327 (53.3)	176 (29.1)	<0.001

Data are presented as no. (%) unless otherwise indicated.
Abbreviations: ESBL, extended-spectrum β-lactamase; UTI, urinary tract infection.

Outcomes:

There was a difference in the outcome event rates between the two cohorts. The primary composite outcome of mortality, clinical failure, readmissions, or extended hospitalization at 90 days occurred in 339 of 613 excluded patients (55.3%) compared to 284 of 604 in the RCT (47%). All-cause mortality at 90 days occurred in 68 (11.3%) patients in the RCT versus 117 (19.1%) excluded patients (p<0.001) (table 5).

Table 5. GNB trial - primary outcome

Variable	Excluded from Randomized controlled trial N=613, (%)	Included in Randomized controlled trial N=604, (%)	P value
Primary outcome	339 (55.3)	284 (47)	0.004
90-day all-cause mortality	117 (19.1)	68 (11.3)	<0.001

Readmissions	209 (38.9) N=537	246 (41.7) N=590	0.343
Extended hospitalization beyond 14 d	82 (13.4)	34 (5.6)	<0.001
Distant Complications	49 (8)	3 (0.5)	<0.001
Relapse of bacteremia	43 (7)	16 (2.6)	<0.001
Suppurative complications	26 (4.2)	26 (4.3)	0.957

Data are presented as no. (%) unless otherwise indicated.

The univariate analysis for primary composite outcome of mortality, clinical failure, readmissions, or extended hospitalization at 90 days is shown in table 4. On multivariable logistic regression, participation in the RCT was not an independent risk factor for primary outcome (as defined above).

3. Colistin alone vs. colistin plus meropenem in patients with severe infections due to multi-drug resistant bacteria

Analysis was performed on 701 patients, including 295 non-randomized patients in the observational arm and 406 RCT patients. The most common reason for not including suitable patients in the RCT was refusal to participate [62% (183/295)]. 20.7% (62/295) of patients were excluded due to treatment with colistin for more than 96 hours, and 16.9% (50/295) were excluded for prior inclusion in the RCT.

Patients' characteristics

Non-randomized and RCT patients were similar in most of the demographic and background parameters. There were more patients with dementia in the RCT [10.7% (49/406) vs. 5.8% (17/295), p=0.050]. Hematological malignancies were more common in non- randomized patients [8.5% (25/295) vs. 3.4% (14/406), p=0.004]**.** At infection onset, RCT patients had more arterial lines [37.2% (151/406) vs. 25.8% (76/295), p=0.001] central venous catheters [55.4% (225/406) vs. 40.3% (119/295), p=0.000] and urinary catheters [87.2% (354/406) vs. 77.3% (228/295), p=0.001] than non-randomized patients (table 6).

Table 6: Colistin trial - patients' characteristics

	Excluded from randomized controlled trial (N=295)	Included in randomized controlled trial (N=406)	P value
Demographics and background			

Age (Mean±SD), year	65±18	66±17	0.411
Gender (female)	101 (34.2%)	151 (37.2%)	0.421
Admitted from home	204 (69.2%)	276 (68%)	0.742
BMI, kg/m^2	27.1 (6.7)	27.4 (5.8)	0.610
Charlson Score (Mean±SD)	2±2	2±2	0.497
Dementia	17 (5.8%)	49 (10.7%)	0.050
Diabetes	61 (20.7%)	90 (22.2%)	0.636
Chronic kidney disease	71 (24.1%)	79 (19.5%)	0.129
Hematological Malignancy	25 (8.5%)	14 (3.4%)	0.004
Congestive heart failure	66 (22.4%)	92 (22.7%)	0.928
Chronic pulmonary disease	57 (19.3%)	91 (22.4%)	0.322
Immune suppressive therapy	54 (18.3%)	61 (15%)	0.247
Status at infection onset (culture taken time)			
Temperature, °C (SD)	37.9 (1.7)	38.0 (1.7)	0.655
Systolic blood pressure, mm Hg (SD)	106 (24)	109 (21)	0.054
Haemodynamic support	68 (24.2%)	75 (18.5%)	0.069
Mechanical ventilation (invasive)	198 (69.5%)	264 (65%)	0.221
Haemodialysis	11 (3.9%)	27 (6.7%)	0.118
SOFA score (Mean±SD)	6±3	6±3	0.755
Creatinine Clearance (Cockcroft-Gault Equation), mL/min (Percentiles 25-75)	59.79 (32.54-108.58)	69.95 (41.21-126.27)	0.012
Arterial line	76 (25.8%)	151 (37.2%)	0.001
Central venous catheter	119 (40.3%)	225 (55.4%)	0.000
Urinary catheter	228 (77.3%)	354 (87.2%)	0.001
Nasogastric tube	201 (68.1%)	285 (70.2%)	0.559

Infection characteristics

Severity of infection was similar in the two groups, as evidenced by similar SOFA scores, need for hemodynamic support, blood pressure and body temperature. Patients not randomized were less likely to acquire their infection in the intensive care unit [22.7% (67/295) vs. 30.5% (124/406). p=0.022], to be infected with Enterobacteriacaeae [35/295 (11.9%) vs. 73/406 (18%), p=0.027]; and more likely to have urinary tract infection [32/295 (10.8%) vs. 26/406 (6.4%), p=0.035]. The minimum inhibitory concentration (MIC) of >0.5 mg/L for colistin was more prevalent in randomized patients [24.3% (85/350) vs. 7.7% (18/233), p=0.000] (table 7).

Table 7: Colistin trial - infection characteristics

	Excluded from randomized controlled trial (N=295)	Included in randomized controlled trial (N=406)	P value
Infection characteristics			
Acquisition of infection in the intensive care unit	67 (22.7%)	124 (30.5%)	0.022

Pathogen			
Acinetobacter baumannii	236 (80%)	312 (76.8%)	0.318
Enterobacterales	35 (11.9%)	73 (18%)	0.027
Pseudomonas/other	24 (8.1%)	21 (5.2%)	0.114
Type of infection			
Bacteraemia	109 (36.9%)	173 (42.6%)	0.131
Ventilator-associated or hospital-acquired pneumonia	140 (47.5%)	182 (44.8%)	0.490
Probable ventilator-associated pneumonia	14 (4.7%)	25 (6.2%)	0.421
Urinary tract infection	32 (10.8%)	26 (6.4%)	0.035
Colistin MIC distribution >0.5 mg/L	18 (7.7%), n=233	85 (24.3%), n=350	0.000

*Numbers apply to all patients in the group unless stated otherwise.

Outcomes

More non-randomized patients met the criteria for the primary outcome, clinical failure at day 14, than randomized patients [82% (242/295) vs. 75.5% (307/406), p=0.042]. Mortality rates were higher in non- randomized patients [40.2% (117/295) vs. 33% (134/406 in the RCT patients, p=0.051]. The difference between the two groups waned at the end of study: 28-day mortality was 47.8% (138/295) in the non- randomized patients vs. 44.3% (180/406) in RCT patients.

4. Trimethoprim-sulfamethoxazole versus vancomycin for severe infections caused by methicillin resistant Staphylococcus aureus.

The RCT included 252 patients and the observational study, 220 patients. The most common reasons for exclusion from the RCT were inability or refusal to provide informed consent and treatment with study drugs for longer than 48 h, together accounting for more than 70% of exclusions, table 8.

Table 8. MRSA trial - reasons for exclusion of patients from the RCT (and inclusion in the observational study)

Reason for exclusion	**N (%) (Total N=220)**
Treatment with study drugs >48 h prior to identification	73 (33.2%)
Refusal to sign an informed consent	44 (20%)
Inability to provide informed consent and no legal guardian	40 (18.2%)
Chronic dialysis	29 (13.2%)
Resistance to one of the study antibiotics	14 (6.4%)
Left-side endocarditis	8 (3.6%)
Acute leucaemia with neutropaenia	7 (3.2%)
Hypersensitivity to one of the antibiotics in the trial	2 (0.9%)
Meningitis	2 (0.9%)

Participation in other trial	1 (0.5%)

Baseline patient characteristics

Excluded patients differed from included patients in their functional and cognitive status, conditions affecting the ability to provide informed consent (table 9). Patients with chronic renal failure were significantly less represented in the RCT and malignancy was slightly less common. Otherwise, there were no significant differences regarding baseline comorbidities. The total Charlson score was significantly higher among excluded patients.

Table 9: MRSA trial - Baseline patient characteristics

	RCT included N=252	Excluded N=220	p Value
Age, years (mean±SD)	65.8±17	67.9±17.2	0.192
Female sex	86 (34.1%)	90 (40.9%)	0.129
Admission from home	194 (77%)	145 (65.9%)	0.008
Functional capacity—bedridden	53 (21%)	115 (52.3%)	<0.001
Dementia	12 (4.8%)	41 (18.6%)	<0.001
Congestive heart failure	50 (19.8%)	40 (18.2%)	0.647
Ischemic heart disease	80 (31.7%)	63 (28.6%)	0.463
Cerebrovascular accident in the past	44 (17.5%)	57 (25.9%)	0.026
Chronic lung disease	35 (13.9%)	27 (12.3%)	0.604
Diabetes mellitus	102 (40.5%)	88 (40%)	0.916
Chronic renal failure	6 (2.4%)	39 (17.7%)	<0.001
Manifest malignancy	49 (19.4%)	58 (26.4%)	0.073
Charlson score (median, percentile)	2 (1–4)	3 (2–4)	0.008

Infection characteristics

Excluded patients were significantly more ventilated and carried more catheters than those included in the RCT (table 10). They had more septic shock at onset and a higher SOFA score. There were more patients with skin, soft tissue, bone and joint infections in the RCT and less CVC-related or primary infections. While bacteraemia occurrence was similar among included and excluded patients, there were more patients with highly probably non-microbiologically documented MRSA infections among the excluded patients.

Table 10. MRSA trial - Infection characteristics

	RCT included N=252	**Excluded N=220**	**P value**
Predisposition			
Hospital-acquired infection†	173 (68.7%)	138 (62.7%)	0.176
Nasogastric tube prior to infection	26 (10.3%)	80 (36.4%)	<0.001

Urine catheter prior to infection	80 (31.7%)	138 (62.7%)	<0.001
Central venous catheter prior to infection	32 (12.7%)	104 (47.2%)	<0.001
Foreign body prior to infection‡	84 (33.3%)	26 (11.8%)	<0.001
Surgery 30 days prior to infection	121 (48%)	77 (35%)	0.004
Mechanical ventilation at onset	27 (10.7%)	98 (44.5%)	<0.001
Infection characteristics and presentation			
Bacteraemia	91 (36.1%)	91 (41.4%)	0.242
Any microbiologically (MRSA)-documented infection	245 (97.2%)	167 (75.9%)	<0.001
Source of infection			
Central venous catheter-related	16 (6.3%)	53 (24.1%)	<0.001
Other endovascular	9 (3.6%)	9 (4.1%)	
Pneumonia	27 (10.7%)	30 (13.6%)	
Skin, soft tissue, bone or joint	168 (66.7%)	54 (24.5%)	
Other documented source	17 (6.7%)	4 (1.8%)	
Primary, unknown source	15 (6%)	70 (31.8%)	
Septic shock at onset	6 (2.4%)	23 (10.5%)	<0.001
SOFA score at onset (median, IQR)	2 (1–4)	3 (2–4)	<0.001

*Numbers apply to all patients in the group unless stated otherwise.

Outcomes

There was a very large difference in outcome events rates between the cohorts. Clinical failure was documented in 83/252 (32.9%) patients in the RCT versus 175/220 (79.5%) among excluded patients (OR 7.94, 95% CI 5.21 to 12.05, p<0.001). All-cause mortality at 30 days occurred in 32 (12.7%) patients in the RCT versus 64 (29.1%) excluded patients (OR 2.82, 95% CI 1.76 to 4.52, p<001).

Part II: Systematic review.

Only 66.8% (123/184) of manuscripts reported the number of patients that were assessed for eligibility for inclusion to the trial. Most of the trials resulted in beneficial intervention effects [65% (114/174)].

Studies that did not report the number of patients assessed for had other methodological deficiencies such as not providing a defined hypothesis [71.4% (40/56) compared to 52.1% (63/121), p=0.001], having inadequate generation of a randomized sequence [65.6% (40/61) compared to 81.3% (100/123), p=0.019], and owning a low percent of allocation concealment [49.2% (30/61) compared to 74.0% (91/123), p=0.015] (table 11).

Table 11: Systematic review - comparison of methodological characteristics according to reporting of patients assessed for eligibility.

Methodological characteristics	Number of participants assessed for eligibility not reported (n=61)	Number of participants assessed for eligibility reported (n=123)	P value
Beneficial intervention effect	66.1% (37/56)	65.3% (77/118)	0.916
Type of illness			
Infectious diseases	26.2% (16/61)	31.7% (39/123)	0.445
Diabetes	21.3% (13/61)	25.2% (31/123)	0.560
Cancer	52.5% (32/61)	43.1% (53/123)	0.230
No defined hypothesis	71.4% (40/56)	52.1% (63/121)	0.015
No information on funding	6.6% (4/61)	10.6% (13/123)	0.376
Funded by industry	29.5% (18/61)	35.0% (43/123)	0.460
Random-sequence generation	65.6% (40/61)	81.3% (100/123)	0.019
Allocation concealment	49.2% (30/61)	74.0% (91/123)	0.001
Lack of blinding of participants	26.2% (16/61)	37.4% (46/123)	0.131
Lack of blinding of trial personnel	27.9% (17/61)	35.0% (43/123)	0.334
lack of blinding of outcome assessors	31.1% (19/61)	39.0% (48/123)	0.296

*Numbers apply to all patients in the group unless stated otherwise.

Table 12 presents a comparison of methodological characteristics according to beneficial intervention effects versus non-beneficial intervention effects.
Trials that have reported the number of patients assessed for eligibility and included a high rate of these patients (>90%) resulted in a higher percentage of beneficial intervention effect: 30.9% (25/81) have reported on a beneficial intervention effect vs. 14.6% (23/270) that have reported a non-beneficial intervention effect, p=0.052. Another interesting difference was the lower rate of blinding of participants in trials that resulted with beneficial intervention effect [40.4% (46/114) in beneficial intervention effect vs. 25.0% (15/60) in non-beneficial intervention effect, p=0.044].

Table 12: Systematic review – a comparison of methodological characteristics according to beneficial intervention effects versus non-beneficial intervention effects.

	Non-beneficial intervention effect (n=60)	Beneficial intervention effect (n=114)	P value
Randomized : assessed for eligibility patient ratio over 0.9	14.6% (6/41)	30.9% (25/81)	0.052
No defined hypothesis	52.5% (31/59)	60.5% (69/114)	0.313
No information on funding	8.3% (5/60)	7.9% (9/114)	0.919

Funded by industry	36.7% (22/60)	31.6% (36/114)	0.499
Random-sequence generation	78.3% (47/60)	75.4% (86/114)	0.669
Allocation concealment	71.7% (43/60)	64.9% (74/114)	0.367
Lack of blinding of participants	25.0% (15/60)	40.4% (46/114)	0.044
Lack of blinding of trial personnel	26.7% (16/60)	36.8% (42/114)	0.176
lack of blinding of outcome assessors	33.3% (20/60)	40.4% (46/114)	0.365

*Numbers apply to all patients in the group unless stated otherwise.

There were no significant differences in the comparison between the three types of illnesses, though we have found that studies in the field of diabetes have a much lower rate of patient inclusion then infectious diseases and cancer.

5. Discussion and Conclusions

The population of studies in the field of infectious diseases often includes hospitalized patients with multiple comorbidities and severe infections. Hence, it is essential to plan the inclusion of these patients in advance.

Three of the four trials (MRSA, GNB and Colistin) included inpatients with severe infections. Despite of the similarity between the studies' populations, different results were observed in the observational cohorts. In the COLISTIN trial, patients not randomized were similar to randomized patients in their baseline characteristics, though RCT patients showed minor differences towards a more severe infection. Excluded patients from the MRSA and GNB trials had significantly more comorbidities and severe infections than those recruited to the RCT.

The main difference between those trials was the consent process. In the Colistin trial, in contrast to the MRSA and GNB trials, we were authorized by the local ethics committees to recruit patients who were not able to provide informed consent and did not have a legal guardian, with the consent of an approved independent physician. Consent by a legal guardian often delays the patients' recruitment. However, consent by independent physician allowed both immediate and

comprehensive inclusion of severely ill patients that characterize the target population.

Examining population external validity of RCTs, we are usually concerned that the included population are younger and have fewer underlying disorders than excluded patients. In the UTI trial, we were able to show that the opposite was true: included patients were older and had a history of recurrent UTI. In addition, the rate of clinical failure was significantly lower among excluded patients. These findings can be explained in several ways. First, the target population of this trial is young and generally healthy women. In most cases these patients did not approach a physician for the treatment of the UTI. This was also demonstrated in lower rates of cultures obtained in excluded patients. These two factors undermine the importance of the primary outcome examined in the RCT to the population of interest. Second, different types of follow-up between study groups. RCT patients attended 2 follow-up visits at 14 and 28 days after completion of antibiotic therapy. During these visits clinical data and urine cultures were obtained. In contrast, data regarding 28-day outcomes of excluded patients were obtained from computerized medical records. Third, RCT patients were randomly assigned to oral nitrofurantoin or fosfomycin, while excluded patients were not treated according to guidelines and were mostly treated with fluoroquinolones. Fluoroquinolones are restricted due to safety and epidemiological reasons, though are considered more effective in terms of clinical and microbiological success.

In our systematic review we found that trials that did not report the number of patients assessed for eligibility were inclined to methodological weaknesses. RCTs that included a very high percentage (>90%) of suitable patients estimated the RCTs' intervention as beneficial twice as often as RCTs with a lower inclusion rate.

An important factor to consider when examining risk of bias is that it is often based on what is reported in papers, and the reported methods do not fully reflect the actual conduct. An assessment of extremely high or low participant inclusion rates could be considered a more accurate estimate of risk of bias. The knowledge on the inclusion process is crucial to evaluate the external validity of studies.

Limitations

Our study has few limitations. First, this study focuses on one aspect of external validity- comparison of characteristics and outcomes of excluded and included patients. This aspect refers to the population validity component and addresses the question of whether the findings of a study can be generalized to patients with characteristics that are different from those in the study, or patients who are treated or followed up differently. For a broader evaluation of external validity, it will be interesting to test ecological validity which specifically examines whether the findings of a study can be generalized to different clinical settings in everyday life. Second, national informed consent regulations vary in different countries, thus conclusions might not be applicable globally. Third, two of the observational studies were based only on data from a single center out of all participating centers in the RCT.

Conclusions

As expected, differences in the population external validity between the four trials were observed. The trials varied in their target population and severity of infection. Our results raise a question on the generalizability of evidence garnered from RCTs in infectious diseases. Nonrestrictive inclusion criteria and access to recruiting the most severely ill patients into the trial population are key elements conferring high population external validity.

6. Policy Implications and Recommendations

- External validity of RCTs in patients with severe infections is hampered by the present consent procedure. The present process causes the most severe patients to be ineligible for inclusion; and causes delays in the start of the trial intervention. One solution as shown here is recruitment by the approval of an independent physician. Other solutions should be sought as well, so the most severe patients will be included, and trial treatment started on time.
- Reporting the data of the patients' cohort who were not included (but could have been according to inclusion and exclusion criteria) will improve our grasp of the external validity of trials.

- Non-reporting on the number and sub-groups of candidates that were not recruited might lead to bias in the study. Such reporting should be followed and demanded by all players.

7. References:

5. Green LW, Glasgow RE. Evaluating the relevance, generalization, and applicability of research: issues in external validation and translation methodology. Evaluation & the health professions. 2006 Mar; 29(1):126-53.
6. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMC medicine. 2010 Dec; 8(1):18.
7. Bracht GH, Glass GV. The external validity of experiments. American educational research journal. 1968 Nov; 5(4):437-74.
8. Rothwell PM. Factors that can affect the external validity of randomised controlled trials. PLoS clinical trials. 2006 May 19; 1(1):e9.
9. Paul M, Bronstein E, Yahav D, Goldberg E, Bishara J, Leibovici L. External validity of a randomised controlled trial on the treatment of severe infections caused by MRSA. BMJ open. 2015 Sep 1; 5(9):e008838.
10. Claessens YE, Aegerter P, Boubaker H, Guidet B, Cariou A. Are clinical trials dealing with severe infection fitting routine practices? Insights from a large registry. Critical Care. 2013 Jun;17 (3):R89.
11. Travers J, Marsh S, Williams M, Weatherall M, Caldwell B, Shirtcliffe P, et al. External validity of randomised controlled trials in asthma: to whom do the results of the trials apply? Thorax. 2007 Mar 1; 62(3):219-23.
12. Saunders C, Byrne CD, Guthrie B, Lindsay RS, McKnight JA, Philip S, et al. External validity of randomized controlled trials of glycaemic control and vascular disease: how representative are participants?. Diabetic Medicine. 2013 Mar; 30(3):300-8.
13. Steg PG, López-Sendón J, de Sa EL, Goodman SG, Gore JM, Anderson FA, et al. External validity of clinical trials in acute myocardial infarction. Archives of internal medicine. 2007 Jan 8;167(1):68-73.

14. Fortin M, Dionne J, Pinho G, Gignac J, Almirall J, Lapointe L. Randomized controlled trials: do they have external validity for patients with multiple comorbidities? Ann Fam Med. 2006; 4:104–8.
15. Yessaian A, Mendivil AA, Brewster WR. Population characteristics in cervical cancer trials: search for external validity. Am J Obstet Gynecol. 2005;192: 407–13.
16. Milojevic K, Beltramini A, Nagash M, Muret A, Richard O, Lambert Y. Esmolol Compared with Amiodarone in the Treatment of Recent-Onset Atrial Fibrillation (RAF): An Emergency Medicine External Validity Study. J Emerg Med. 2019 Mar; 56(3):308-318. doi: 10.1016/j.jemermed.2018.12.010. Epub 2019 Jan 30.
17. Huttner A, Kowalczyk A, Turjeman A, Babich T, Brossier C, Eliakim-Raz N, Kosiek K, Martinez de Tejada B, Roux X, Shiber S, Theuretzbacher U, von Dach E, Yahav D, Leibovici L, Godycki-Cwirko M, Mouton JW, Harbarth S. Effect of 5-Day Nitrofurantoin vs Single-Dose Fosfomycin on Clinical Resolution of Uncomplicated Lower Urinary Tract Infection in Women: A Randomized Clinical Trial. JAMA. 2018 May 1; 319(17):1781-1789.
18. Yahav D, Franceschini E, Koppel F, Turjeman A, Babich T, Bitterman R, Neuberger A, Ghanem-Zoubi N, Santoro A, Eliakim-Raz N, Pertzov B, Steinmetz T, Stern A, Dickstein Y, Maroun E, Zayyad H, Bishara J, Alon D, Edel Y, Goldberg E, Venturelli C, Mussini C, Leibovici L, Paul M; Seven versus fourteen Days of Antibiotic Therapy for uncomplicated Gram-negative Bacteremia: a Non-inferiority Randomized Controlled Trial. Clin Infect Dis. 2018 Dec 11.
19. Paul M, Daikos GL, Durante-Mangoni E, Yahav D, Carmeli Y, Benattar YD, Skiada A, Andini R, Eliakim-Raz N, Nutman A, Zusman O, Antoniadou A, Pafundi PC, Adler A, Dickstein Y, Pavleas I, Zampino R, Daitch V, Bitterman R, Zayyad H, Koppel F, Levi I, Babich T, Friberg LE, Mouton JW, Theuretzbacher U, Leibovici L. Colistin alone versus colistin plus meropenem for treatment of severe infections caused by carbapenem-resistant Gram-negative bacteria: an open-label, randomised controlled trial. Lancet Infect Dis. 2018 Apr;18(4):391-400.

20. Paul M, Bishara J, Yahav D, et al. Trimethoprim-sulfamethoxazole versus vancomycin for severe infections caused by methicillin resistant Staphylococcus aureus: randomised controlled trial. BMJ 2015;350:h2219.