Trials

REVIEW Open Access

Safety outcome trials in four mainline medical journals 30 years apart: a narrative review and the need for the transparency of informed consents

Sinem Nihal Esatoglu[1], Ayse Ozdede[1], Yesim Ozguler[1], Koray Tascilar[2] and Hasan Yazici[3*]

Abstract

Background It was our impression that safety outcome trials were getting more frequent, raising ethical issues mainly related to patient autonomy. We and others had also proposed this autonomy would be best served if wording of the informed consents would be in the public domain.

Methods Initially two observers and an arbiter tabulated the main aims of randomized controlled trials (RCTs) published in 1990–1991 vs. 2019–2020 as efficacy, safety, or undecided in four mainline medical journals, from the websites. A pragmatic design as well as other salient features was also tabulated. After noting too many trials were categorized as undecided, two additional independent observers and the arbiter did a reassessment.

Results In our reassessment of 889 RCTs, 309 in earlier and 580 in the later time period, 828 (93%) were categorized as efficacy and 47 (5%) as a safety trial. We were undecided in 14 (2%) trials. The proportion of safety outcome trials between the two time periods were similar [14/309 (5%) vs. 33/580 (6%)] while RCTs of any category conducted in the critical care settings notably increased in time [12/309 (4%) vs. 52/580 (9%) OR 2.4; 1.3–4.6]. Death was a primary outcome measure in 0/14 among the earlier and 16/33 (49%) among the later safety outcome trials. Stroke in 9 and myocardial infarction in 8 safety outcome trials were additional primary outcome measures in the same group. There were 2 pragmatic trials in the earlier and 93 in the later period.

Conclusion Although we did not observe a differential increase among the safety outcome trials of all categories, those in critical care settings had significantly increased in time. So did the safety outcome trials with primary outcome measures like death, myocardial infarction, and stroke as well as randomized controlled trials with a pragmatic design. These raise the issue of autonomy related to how clearly the sought-after safety through designing empirical studies primarily to quantitate harm had been worded to the trial participants in plain language. We maintain this issue cannot be adequately addressed unless the informed consent forms, especially for safety outcome trials, are in the public domain.

Keywords Randomized controlled trials, Informed consent, Study aim, Primary outcome measures, Nuremberg Code, Helsinki Code, Research integrity

*Correspondence:
Hasan Yazici
hasan@yazici.net
Full list of author information is available at the end of the article

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Background

The randomized controlled trial (RCT) remains a most important scientific tool as well as a tool for drug licensing. This tool is and has been widely and successfully utilized for primarily assessing efficacy of medical interventions while its use in primarily assessing harm has been debated [1–3] as we will expand in the "Discussion."

Our aim for starting this narrative review was our impression that RCTs mainly designed to assess harm, referred to as safety outcome trials (SOTs), were becoming frequent. In addition, we [4, 5] like others [6–8] had been vocal about the need for increased transparency of the written informed consent forms (ICs) of all RCTs. By ICs we mean the document by which the trialists explain to the potential volunteers the main aims and the procedural components of the trial as well as the potential gains and harms of participating in the said trial, in plain language. Since SOTs by definition mainly aim to assess safety by designing a study to quantitate harm, we reasoned that our impression of the increase in SOTs would further strengthen our plea for the transparency of the ICs.

The initial analyses of our data had supported our impression of the increase in SOTs as we had presented as abstracts [9, 10]. On the other hand, we were concerned that categorizing an RCT as primarily for efficacy or safety was difficult and imprecise as reflected by too many RCTs being assigned to the undecided category, 61/894 (7%) in our earlier analyses (Supplementary material 1). We thus repeated our review with two additional observers and reassessed our findings.

This manuscript is based on the results of this reassessment. We also analyzed, based on our observations other than the mere number of SOTs, features like the trial settings, designs, and the nature of the primary outcome measures (POMs). We then discuss why we consider it important to render the ICs more transparent.

We gratifyingly learn that the US Government recently mandated the formal posting of the ICs of RCTs conducted under government funds in public webpages [11] with no emphasis on special requirements for SOTs. We consider this mandate as an important step in the right direction and would propose that this mandate should also cover the ICs of RCTs sponsored by all sources.

Methods

Initial assessment

First, two of us (SNE and AO) independently conducted a search for all RCTs published in NEJM, JAMA, The Lancet, and BMJ in 1990–1991 and 2019–2020, using the websites of these journals to review each issue for manuscripts of RCTs. We excluded phase 1 and phase 2 trials and manuscripts reporting only post hoc analyses or long-term extensions of RCTs. COVID-19 trials published in 2020 were also excluded to make a more justifiable comparison between the two different time periods. The authors used a predesigned spreadsheet and extracted data independently.

Trial characteristics were tabulated as follows:

a. Journal, first author, title, and the year of publication.
b. Patient demography: adult, pediatric, and the infant-neonate-preterm group. Any trial where > 50% of the participants were adults was considered as an adult trial.
c. Intervention: drug, surgery, vaccine, and other (e.g., transfusions, intubations, and physical therapy).
d. Setting: outpatient, inpatient, critical care unit (intensive care and special neonatal care units), and other (e.g., among health professionals in a hospital setting).
e. IC: written, verbal, waiver of consent, deferred, and no reference to an informed consent. The statements to the effect "Informed consent was obtained" were considered as the presence of a written IC.

The primary outcome of interest of our study was whether an RCT had mainly been planned for efficacy or safety. This main intent could, in many instances, have been declared by the RCT authors, usually expressed as a POM or, if there were more than one POM, in an order of priority. However, for the purposes of this study, we decided what the main aim was only after going through each study in detail. To this end, we considered an RCT with a definite primary endpoint of efficacy without any easily predicted and important potential harm to either the probands or to the controls as an efficacy (group E) trial. An RCT with declared or what we judged as primary endpoint(s) of "safety," "safety and efficacy" or "efficacy and safety," or even "efficacy" alone in which both observers noted the presence of potential serious harm to patients in any one arm of the study, was a group S trial. When an agreement could not be reached to designate a trial as E or S, it was categorized in the undecided category (group U). After two independent reviewers finalized their evaluations, the senior author (HY), as an arbiter, also did a reevaluation. He revaluated the full-text manuscripts of all group S and group U trials, and a random sample of group E trials, blind to the previous categorizations. When unanimous agreement could not be reached among the three observers even after a second joint reevaluation to designate a trial as a group E or S, it was categorized as a group U trial.

Reassessment

After reconsideration of the results based on the methodology we just described and presented our results as abstracts in two international meetings [9, 10], we considered there perhaps were too many group U trials in our review. We then decided to ask two further independent observers (KT and YO) to reassess the thus far classified S and U groups, a second time. In this reassessment, the consideration of *the presence of potential serious harm to patients in any one arm of the study* was omitted when designating a trial as safety. A second reassessment was then made again with the presence of the arbiter (HY), before the final designation of a trial as S, E, or U. Trials which were labeled as E&S or S&E by their authors were classified as either solo E or solo S in this reassessment.

The flowchart for data curation in the two assessment stages is attached (Fig. 1).

After this recategorization and the numbers of S, E, and U trials were finalized, the specific type of POMs sought in the S trials was further tabulated.

A secondary target of our review was to tabulate the frequency and some of the salient features of pragmatic trials. The authors' designation of their work as pragmatic was our sole criterion for an RCT to be categorized as such. Since there had only been two pragmatic trials in the 1990–1991 group, only the findings related to the 2019–2020 group were tabulated along with the number

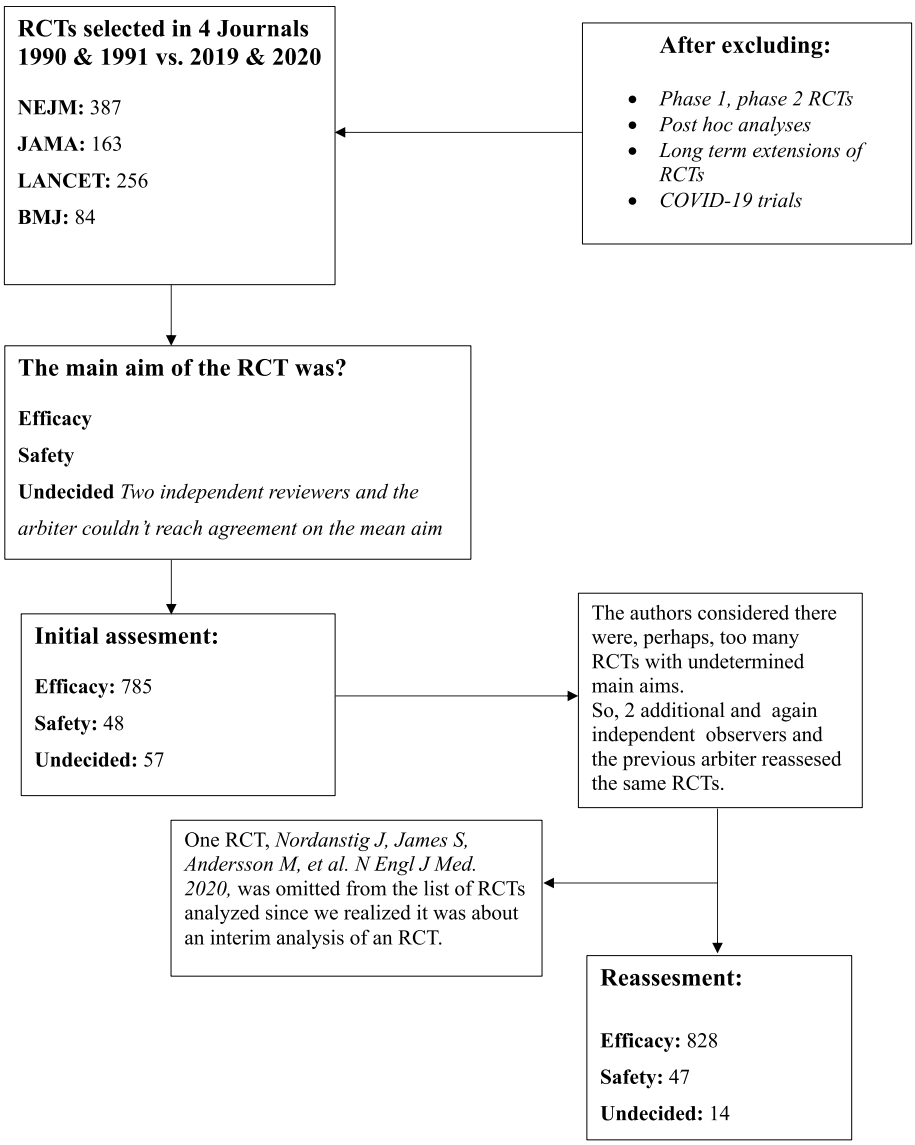

Fig. 1 Data curation

of trial participants, the trial durations, and whether a particular RCT was a S trial or not.

Statistics

Data were presented as frequencies and percentages. Odds ratios (OR) and 95% confidence intervals were estimated to describe, where we considered necessary, the relative frequency of a designated study type or a study feature in the period 2019–2020, versus the period 1990–1991.

Results

The results we present in this report are based mainly on the recategorization of the RCTs made after our reassessment as explained above and in Fig. 1, which is a flowchart of our data curation based on the two assessment stages.

Supplementary material 1 displays the table that was presented in both previous abstract presentations [9, 10] of our earlier assessment as we explained above. Supplementary materials 2, 3, and 4 give the raw data based after their categorization into the E, S, and U groups. Supplementary materials 2, 3, and 4 also indicate the initial assessment categorizations of the individual RCTs. Finally, the changes between efficacy, safety, and undecided trials between the initial and the reassessments for both trial period in detail are depicted in Supplementary material 5.

Table 1 shows the characteristics of the RCTs included in the two indicated time periods, 30 years apart. We see that there were more RCTs published in later years and there was a notable decrease in the proportion of drugs as the intervention in these trials paralleled by an increase in the proportion of surgical interventions and in interventions designated as "other." Table 1 also shows there was a notable increase in RCTs conducted in critical care settings. Finally, there was an increase in the written ICs in the later period.

Table 2 shows, while there were numerical increases in time in the total number of RCTs, there were no appreciable changes in the respective proportion of E and S trials. It also shows that the majority 44/52 (84%) of RCTs in the critical care settings were E trials.

Table 3 displays the POMs related to the individual RCTs in group S. It is to be noted that none of the 14 trials in the earlier period while 16/33 (49%) of the later period had death as a declared POM or one of several POMs, similarly declared. Stroke was an additional

Table 1 General characteristics of RCTs surveyed

	All RCTs *n* = 889	RCTs published in 1990–1991 *n* = 309	RCTs published in 2019–2020 *n* = 580	OR [95% CI] 1990–1991 vs. 2019–2020
Patient demographics				
Adult	756 (85)	251 (81)	505 (87)	1.6 [1.1–2.3]
Pediatric	99 (11)	45 (15)	54 (9)	0.6 [0.4–0.9]
Infant-neonate-preterm	34 (4)	13 (4)	21 (4)	0.9 [0.4–1.7]
Type of intervention				
Drug	624 (70)	241 (78)	383 (66)	0.6 [0.4–0.7]
Surgery	48 (6)	7 (2)	41 (7)	3.2 [1.5–7.4]
Vaccine	21 (2)	11 (4)	10 (2)	0.4 [0.2–1.1]
Other*	196 (22)	50 (16)	146 (25)	1.7 [1.2–2.5]
Setting				
Outpatient	592 (66)	221 (71)	371 (64)	0.7 [0.5–0.9]
Inpatient	230 (26)	76 (25)	154 (27)	1.1 [0.8–1.5]
Critical care	74 (8)	12 (4)	52 (9)	2.4 [1.3–4.6]
Other**	3	0	3	
Informed consent				
Written	821 (93)	263 (85)	558 (96)	4.5 [2.6–7.5]
Verbal	9 (1)	6 (2)	3 (1)	0.2 [0.1–1.1]
Waiver of consent	10 (1)	1	9 (2)	4.8 [0.6–38.4]
Deferred	12 (1)	3 (1)	9 (2)	1.6 [0.4–6.0]
No consent reported	37 (4)	36 (13)	1	0.01 [0–0.1]

Unless otherwise indicated, all parentheses indicate percentages

* Transfusions, intubations, physical therapy, etc.

** Among the health professionals in a hospital setting

Table 2 Tabulation of trials (both pragmatic and non-pragmatic) at two time periods as efficacy (E), safety (S), or undecided (U)

	All RCTs *n* = 889	RCTs published in 1990–1991 *n* = 309	RCTs published in 2019–2020 *n* = 580	OR [95% CI] 1990–1991 vs. 2019–2020
E trial	828 (93)	293 (95)	535 (92)	0.6 [0.4–1.2]
S trial	47 (5)	14 (5)	33 (6)	1.2 [0.7–2.5]
U trial	14 (2)	2	12 (2)	3.2 [0.7–15.0]
Critical care trials	*64 (7)*	*12 (4)*	*52 (9)*	*2.4 [1.3–4.6]*
E trial	55 (86)	11 (92)	44 (84)	2.2 [1.1–4.4]
S trial	6 (9)	1 (8)	5 (10)	2.7 [0.3–23.0]
U trial	3 (5)	0	3 (6)	–
Pragmatic trials	*95 (11)*	*2 (1)*	*93 (16)*	*29 [7.2–120.0]*
E trial	89 (94)	2	87 (94)	27 [7.0–110.0]
S trial	4 (4)	0	4 (4)	–
U trial	2 (2)	0	2 (2)	–

All parentheses indicate percentages

See text for E, S, or U designations

POM in 9/16 and myocardial infarction in 8/16 of these trials which also listed death as a POM.

Table 4 details the features of the pragmatic as compared to the other RCTs reviewed in the later period since only 2/309 of the RCTs had been declared by their authors as pragmatic in 1990–1991. Of the 580 RCTs in 2019–2020, 93 (16%) were pragmatic trials (Table 2). Table 4 shows that the median number of participants enrolled in these trials was somewhat greater compared to the non-pragmatic trials and this was also true for group S trials as compared to the groups E & U. Table 4 also shows the trial durations were rather dependent on whether a pragmatic trial was also an SOT, as the median trial duration in this group was only a month.

Discussion

Our main aim for undertaking this work was to substantiate our impression that SOTs were on the increase. We were particularly interested in substantiating this because of the previous proposal we and others had made for the transparency of the ICs [4–8]. Since SOTs mainly aim to quantitate harm through empirical research, we reasoned that if we could substantiate the increase in SOTs, our proposal for the transparency of the ICs would strengthen.

Our review did not support our impression that the SOTs were on the increase . On the other hand, some features of RCTs in general and that of SOTs in particular we reviewed, we believe, further strengthened our plea for the transparency of the ICs.

Namely:

I. More RCTs were being conducted in critical care settings where there are not infrequent issues in obtaining ICs [12].
II. POMs like death, stroke, or myocardial infarction were common POMs among the SOTs of the later years. We are not aware of formal studies about the frequency of these particular POMs in SOTs nor how they have been or should have been related to a potential trial participant in plain words in the related ICs.
III. A pragmatic study design, almost solely present among the later RCTs in our review, also potentially has ethical issues related to its ICs [13] as we will further discuss under a separate subheading.

It is important to note that our plea for the transparency of ICs only referred to the written ICs. In this line, it is important to note that the portion of such written ICs increased from 85% among the earlier to 96% among the latter group of trials (Table 1). This strengthens our assumption that once ICs are in the public domain, the frequency of the very risky POMs like death, stroke, and myocardial infarction will substantially lessen.

An important methodological issue was that the final decision was ours whether an RCT had the main aim of efficacy or safety. In many instances, these aims were in line with the POMs stated by the RCT authors. However, in some instances, an RCT might declare its POMs both as an E plus a S trial, or vice versa where the RCT authors use the stated order of POMs to prioritize their main aim. Also, the indicated main aim of an RCT might not be the same as what the reader might consider as the

Table 3 Primary outcome measures in group S trials

Journal	First author	Title	Year	Outcome
NEJM	Farwell JR	Phenobarbital for febrile seizures — effects on intelligence and on seizure recurrence	1990	Low IQ[2]
JAMA	Margolis KL	Frequency of adverse reactions to influenza vaccine in the elderly a randomized, placebo-controlled trial	1990	Adverse reactions
BMJ	Lucas A	Early diet of preterm infants and development of allergic or atopic disease: randomized prospective study	1990	Allergic reactions[2]
NEJM	Riggs BL	Effect of fluoride treatment on the fracture rate in postmenopausal women with osteoporosis	1990	Bone fracture
BMJ	Levesque H	Low molecular weight heparins and hypoaldosteronism	1990	Hypoaldosteronism
BMJ	Jones MJ	Cognitive and functional competence after anesthesia in patients aged over 60: controlled trial of general and regional anesthesia for elective hip or knee replacement	1990	Cognitive function
BMJ	Wolthers OD	Short term linear growth in asthmatic children during treatment with prednisolone	1990	Growth delay[2]
LANCET	Cook JD	Gastric delivery system for iron supplementation	1990	Gastrointestinal side effects
NEJM	Salusky IB	Aluminum accumulation during treatment with aluminum hydroxide and dialysis in children and young adults with chronic renal disease	1991	Aluminum levels[2]
LANCET	Meade TW	Medical Research Council European Trial of chorion villus sampling	1991	Safety of chorion sampling
BMJ	Wolthers OD	Growth of asthmatic children during treatment with budesonide: a double blind trial	1991	Growth suppression[2]
BMJ	Frenklen LA	Assessment of pain after subcutaneous injection of erythropoietin in patients receiving hemodialysis	1991	Pain
BMJ	Egger M	Influence of human insulin on symptoms and awareness of hypoglycemia: a randomized double blind crossover trial	1991	Hypoglycemia
BMJ	Solomon SA	Beta blockade and intermittent claudication: placebo controlled trial of atenolol and nifedipine and their combination	1991	Claudication
NEJM	Mariette C	Hybrid minimally invasive esophagectomy for esophageal cancer	2019	Surgical complication
NEJM	Llanos-Cuentas A	Tafenoquine versus primaquine to prevent relapse of *Plasmodium vivax* malaria	2019	Anemia
NEJM	Wiviott SD	Dapagliflozin and cardiovascular outcomes in type 2 diabetes	2019	Death, myocardial infarction, stroke[1,4,5]
NEJM	Silber JH	Patient safety outcomes under flexible and standard resident duty-hour rules	2019	Death[1,3]
NEJM	Ray KK	Safety and efficacy of bempedoic acid to reduce LDL cholesterol	2019	Cardiac effusion
NEJM	Landoni G	Volatile anesthetics versus total intravenous anesthesia for cardiac surgery	2019	Death[1,3]
NEJM	Lopes RD	Antithrombotic therapy after acute coronary syndrome or PCI in atrial fibrillation	2019	Death, any type of bleeding[1]
NEJM	Husain M	Oral semaglutide and cardiovascular outcomes in patients with type 2 diabetes	2019	Death, myocardial infarction, stroke[1,4,5]
NEJM	Gupta A	Isoniazid preventive therapy in HIV-infected pregnant and postpartum women	2019	Drug toxicity
NEJM	Mehran R	Ticagrelor with or without aspirin in high-risk patients after PCI	2019	Any type of bleeding
JAMA	Wildes TS	Effect of electroencephalography-guided anesthetic administration on postoperative delirium among older adults undergoing major surgery the ENGAGES randomized clinical trial	2019	Delirium[3]
JAMA	Wise RA	Effect of aclidinium bromide on major cardiovascular events and exacerbations in high-risk patients with chronic obstructive pulmonary disease the ASCENT-COPD randomized clinical trial	2019	Death, myocardial infarction, stroke[1,4,5]
JAMA	Watanabe H	Effect of 1-month dual antiplatelet therapy followed by clopidogrel vs 12-month dual antiplatelet therapy on cardiovascular and bleeding events in patients receiving PCI the STOPDAPT-2 randomized clinical trial	2019	Death, myocardial infarction, stroke[1,4,5]
JAMA	Rosenstock J	Effect of linagliptin vs glimepiride on major adverse cardiovascular outcomes in patients with type 2 diabetes the CAROLINA randomized clinical trial	2019	Death, myocardial infarction, stroke[1,4,5]
JAMA	Purisch SE	Effect of delayed vs immediate umbilical cord clamping on maternal blood loss in term cesarean delivery a randomized clinical trial	2019	Maternal blood loss

Table 3 (continued)

Journal	First author	Title	Year	Outcome
LANCET	Mehanna H	Radiotherapy plus cisplatin or cetuximab in low-risk human papillomavirus-positive oropharyngeal cancer (De-ESCALaTE HPV): an open-label randomized controlled phase 3 trial	2019	Severe toxicity
LANCET	Halliday BP	Withdrawal of pharmacological treatment for heart failure in patients with recovered dilated cardiomyopathy (TRED-HF): an open-label, pilot, randomized trial	2019	Heart failure
LANCET	McCann ME	Neurodevelopmental outcome at 5 years of age after general anesthesia or awake-regional anesthesia in infancy (GAS): an international, multicenter, randomized, controlled equivalence trial	2019	Neurologic development[2]
LANCET	Al-Shahi Salman R	Effects of antiplatelet therapy after stroke due to intracerebral hemorrhage (RESTART): a randomized, open-label trial	2019	Cerebral bleeding[3]
LANCET	Vranckx P	Edoxaban-based versus vitamin K antagonist-based antithrombotic regimen after successful coronary stenting in patients with atrial fibrillation (ENTRUST-AF PCI): a randomized, open-label, phase 3b trial	2019	Cerebral bleeding
LANCET	Lanz J	Safety and efficacy of a self-expanding versus a balloon-expandable bioprosthesis for transcatheter aortic valve replacement in patients with symptomatic severe aortic stenosis: a randomized non-inferiority trial	2019	Death, stroke[1,5]
NEJM	Brouwer J	Aspirin with or without clopidogrel after transcatheter aortic-valve implantation	2020	Any type of bleeding
JAMA	Wolf M	Effects of iron isomaltoside vs ferric carboxymaltose on hypophosphatemia in iron-deficiency anemia two randomized clinical trials	2020	Hypophosphatemia
NEJM	Nijenhuis VJ	Anticoagulation with or without clopidogrel after transcatheter aortic-valve implantation	2020	Any type of bleeding
NEJM	Cannon CP	Cardiovascular outcomes with ertugliflozin in type 2 diabetes	2020	Death, myocardial infarction, stroke[1,4,5]
NEJM	Landrigan CP	Effect on patient safety of a resident physician schedule without 24-h shifts	2020	Death[1,2,3]
JAMA	Kim BK	Effect of ticagrelor monotherapy vs ticagrelor with aspirin on major bleeding and cardiovascular events in patients with acute coronary syndrome the TICO randomized clinical trial	2020	Death, any type of bleeding[1]
NEJM	Olsen HT	No sedation or light sedation in critically ill, mechanically ventilated patients	2020	Death[1]
NEJM	Knops RE	Subcutaneous or transvenous defibrillator therapy	2020	Defibrillator related complication
JAMA	Arkell TR	Effect of cannabidiol and Δ9-tetrahydrocannabinol on driving performance a randomized clinical trial	2020	Cognitive and psychomotor impairment
LANCET	Mackenzie IS	Long-term cardiovascular safety of febuxostat compared with allopurinol in patients with gout (FAST): a multicenter, prospective, randomized, open-label, non-inferiority trial	2020	Death, myocardial infarction, stroke[1,4,5]
LANCET	Kim HS	Prasugrel-based de-escalation of dual antiplatelet therapy after percutaneous coronary intervention in patients with acute coronary syndrome (HOST-REDUCE-POLYTECH-ACS): an open-label, multicenter, non-inferiority randomized trial	2020	Death, myocardial infarction, stroke[1,4,5]
LANCET	Short TG	Anesthetic depth and complications after major surgery: an international, randomized controlled trial	2020	Death[1]

[1] Death; [2]Non-adult population; [3]Pragmatic trial; [4]Myocardial infarction; [5]Stroke

main aim after carefully going over the same RCT [14]. In our review where the RCT authors declared a dual aim, it was solely our judgment what the overriding aim was regardless of what the authors stated. We felt this was justified since our main concern was to identify SOTs where potential harms to the study patients were perhaps not adequately explained to the study participants during the IC procedure.

Ethical considerations

It seems the older consideration that the harm of a medical intervention, either in the official post-approval state or otherwise in routine clinical use, like many older medical remedies are, is best assessed by observational data [1] is currently being seriously challenged by SOTs. It is important to remember that the Nuremberg Code explicitly states that "No experiment should be conducted where there is an a priori reason to believe that death

Table 4 Comparison as to pragmatic vs. non-pragmatic of the median number of participants and the median trial durations in all RCTs, S, and E & U trials* published in 2019–2020

	Median number of participants (IQR)	Median trial duration in months (IQR)
All trials ($n=580$)	702 (303–1739)	8 (3–18)
Pragmatic ($n=93$)	886 (481–2445)	6 (1–12)
E & U trials ($n=89$)	885 (450–2384)	6 (1–12)
S trials ($n=4$)	3316 (885–124,790)	1 (1–30)
Non-pragmatic ($n=487$)	665 (303–1672)	9 (3–18)
E & U trials ($n=458$)	654 (303–1605)	6 (3–18)
S trials ($n=29$)	956 (330–5324)	12 (6–24)

* See text for E, S, or U designations

or disabling injury will occur; except, perhaps, in those experiments where the experimental physicians also serve as subjects" [15]. We do not find such strong discouragement against SOTs in the ensuing Helsinki Declarations. These are eleven in number starting in 1964 and ending in 2022. However, starting with the 2000 version, in the last five versions of the Helsinki Declaration the following statement appears: "Physicians may not be involved in a research study involving human subjects unless they are confident that the risks have been adequately assessed and can be satisfactorily managed." [16]. The exact wording here discouraging SOTs might not be as strong as that in the Nuremberg Code. Nevertheless, to reconcile a POM of death, being present in almost a half of the latter period SOTs we reviewed, with the quoted paragraph from the Helsinki Code is clearly difficult. Here we also draw attention to the fact that the same paragraph aimed at research in general and not necessarily to SOTs, almost as a rule related to officially approved medical remedies.

We like to emphasize that a meaningful majority of efficacy trials are currently conducted in critical care facilities (Table 2) where as we described the issue of IC is more involved. This is why we want to reiterate that ICs should better be transparent for all types of RCTs.

A related and a historically important document is the report of an expert committee of the National Academy of Sciences on the ethical issues in studying the safety of approved drugs. This report was prepared upon a request from the FDA specifically for post-approval RCTs of safety. The report initially appeared in a letter form [2] and later as a book. These documents point out that an SOT is only justified when observational studies have not provided or cannot adequately provide an answer to the existence and the importance of a newly observed/suspected treatment harm from an already approved medical product. The report also emphasizes that the related informed consent in an SOT is expected to be substantially different from an ordinary efficacy trial. Those safeguards should include "…. the determination by an appropriately constituted review committee that the additional risk is small enough for it to be ethical to ask people whether they are willing to accept it *solely* [sic] to contribute to the public good." [2]. Personal vs. public gain is perhaps at the core of the ethics of IC process related to an RCT [17]. This concern in planning an SOT has also been an important drive behind our conducting this narrative review.

Furthermore, some contemporary authorities posit that SOTs are never justified. "Although the outcome of a trial is uncertain, the primary response variable should always be one where either benefit or noninferiority is achievable" [3].

In brief, we appreciate that great care is indeed currently taken to prepare the ICs by expert institutional review boards under the scrutiny of drug licensing authorities with due concern for and respect to the patient autonomy. However, these considerations do not deter us from wanting to know how exactly POMs like probable death, myocardial infarction, or a stroke are conveyed, in plain language, to a potential study participant during the consent procedure.

Pragmatic trials

Recently, a case for pragmatic trials has been made that they are more easily conducted in a *real life* setting and among greater number of patients and for longer time periods [18]. Our findings do support (Table 4) that the trial durations and the numbers of participants are indeed higher in trials described as pragmatic by their authors. Interestingly, as we gratifyingly note, the trial durations get notably shorter in SOTs with a pragmatic design (Table 4) giving less time for potential harm.

The reason we focused on the pragmatic trials as a separate group in this review was the fact that there are well recognized potential ethical issues with pragmatic trials, particularly related to what constitutes an actual minimal risk to a trial participant, in the consent procedure of such trials [18]. As we point out in Table 3, there were five pragmatic trials among the SOTs. Among these five trials, three had death and one cerebral bleeding as a POM. Admittedly the numbers at hand are yet small while we are concerned that such pragmatic trials with such POMs will get more frequent in time.

Study limitations

We consider our study to have two potential limitations. Paramount was accurately classifying the RCTs we reviewed as efficacy or safety trials. This was mainly the

reason we recruited two additional colleagues into our review as we summarized in the methodology section. As is shown in Fig. 1 and can be seen in detail in Supplementary materials 2, 3, and 4, the main result of utilizing two additional observers was a significant decrease in the number of U trials the majority of which were reclassified under the E group (Supplementary material 5). We reason this notable decrease in the U trials in the second assessment the result of the observers not being bound by the generalization that *the presence of potential serious harm to patients in any one arm of the study* also qualified an RCT as an SOT in the reassessment.

It is interesting to note while we call the difficulty to tell an SOT from an efficacy trial as a limitation of our work, on further consideration, this very limitation can be viewed as supporting our call for transparency. Our observers and our arbiter are academics particularly interested and versed in RCTs. The consideration that even with this background we experienced serious difficulties in classifying a portion of completed RCTs published in our high impact medical journals as whether testing efficacy or safety, makes us feel more justified in our wanting to know how the main aims of these trials were conveyed to potential trial participants in plain language. We envisage that our not formally tabulating the POMs among the E and the U groups may also be considered as a potential limitation. On the other hand, our not tabulating death, stroke, or myocardial infarction as POMs among the E and U group of trials would, we reason, might even be a bias in the right direction in that this would only strengthen or plea to know how these particular POMs were, again, vocalized to the trial subjects.

The second potential limitation of our survey was in identifying the pragmatic trials. The pinpointing of an RCT as pragmatic can indeed be difficult. It has been said that the distinction between an explanatory, be it efficacy or safety, and a pragmatic trial can indeed be difficult and according to one scheme can range from a very explanatory to very pragmatic [19]. For this reason, we chose to use the authors' naming their own trials as pragmatic, in our tabulations.

Conclusion

We admit that an SOT may be an efficient way to assess potential harm of a medical remedy already used in clinical practice. On the other hand, whether this approach is superior to a well-conducted observational study needs to be further and, we suggest, urgently assessed. In this line, the powerful tool of machine learning, so-called artificial intelligence, for inferring on safety using open-source or administrative data [20] might prove to be especially useful. SOTs can probably continue to be designed provided they particularly target short term and truly reversible harms.

We maintain that the ICs must be in the public domain as we and others previously proposed. The SOT might be considered a useful or a *good way* to collect safety data but whether it is the *right way* remains, to us, open to debate. Finally, this right way we propose, that of putting the ICs in the public domain, is one of the ways we can maintain or, according to some [21], reestablish the public trust in science.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13063-025-08732-x.

Supplementary Material 1.
Supplementary Material 2.
Supplementary Material 3.
Supplementary Material 4.
Supplementary Material 5.

Acknowledgements
We thank Angela T. Yazici and Yusuf Yazici MD for critical reading of our manuscript.

Authors' contributions
Nihal S. Esatoglu: conceptualization, methodology, investigation, data curation, formal analysis, writing—review, editing. Ayse Ozdede: conceptualization, methodology, investigation, data curation, formal analysis, writing—review, editing. Yesim Ozguler: formal analysis, validation, writing—review and editing. Koray Tascilar: formal analysis, validation, writing—review, editing. Hasan Yazici: conceptualization, methodology, investigation, data curation, formal analysis, writing—original draft.

Funding
None.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
None.

Author details
[1]Division of Rheumatology, Department of Medicine, Cerrahpasa Medical School, Istanbul University-Cerrahpasa, Istanbul, Turkey. [2]Department of Internal Medicine 3, Friedrich-Alexander University Erlangen-Nuremberg and Universitätsklinikum Erlangen, Erlangen, Germany. [3]Internal Medicine (Rheumatology), Academic Hospital, Istanbul, Turkey.

Received: 27 August 2024 Accepted: 14 January 2025
Published online: 25 January 2025

References

1. Vandenbroucke JP. What is the best evidence for determining harms of medical treatment? CMAJ. 2006;174:645–6.

2. Institute of Medicine (US) Committee on Ethical and Scientific Issues in Studying the Safety of Approved Drugs: A Letter Report. Washington (DC): National Academies Press (US); 2010.
3. Friedman, LM et al. editors. Fundementals of clinical trials. 5th ed. Springer, International Publishing Switzerland; 2015. p. 28.
4. Yazici Y, Yazici H. Informed consent: time for more transparency. Arthritis Res Ther. 2010;12:121.
5. Ozdede A, Yazıcı H. Cardiovascular and cancer risk with tofacitinib in rheumatoid arthritis. N Engl J Med. 2022;386:1766.
6. Kotz D, Viechtbauer W, Spigt M, Crutzen R. Details about informed consent procedures of randomized controlled trials should be reported transparently. J Clin Epidemiol. 2019;109:133–5.
7. Tugwell P, Knottnerus JA. Need for greater transparency in documenting informed consent. J Clin Epidemiol. 2019;109:v–vii.
8. Dal-Ré R, Voo TC, Holm S. Leading pharmaceutical companies lacked transparency on clinical trials' informed consent forms. J Clin Epidemiol. 2023;162:187–9.
9. Ozdede A, Esatoglu SN, Yazici H. AB1614 the increase in safety outcome trials and the issue of informed consent [abstract]. Ann Rheum Dis. 2023;82:2041–2.
10. Ozdede A, Esatoglu S, Yazici H. An increase in safety outcome trials and the issue of informed consent [abstract]. Arthritis Rheumatol. 2023; 75 (suppl 9). http://acrabstracts.org/abstract/an-increase-in-safety-outcome-trials-and-the-issue-of-informed-consent/. Accessed 21 Jan 2025.
11. US Department of Health & Human Services. Informed consent posting instructions (2022). General instructions on the informed consent posting requirement (45 CFR 46.116(h)) 2022. Available at https://www.hhs.gov/ohrp/regulations-and-policy/informed-consent-posting/informed-consent-posting-guidance/index.html. Accessed 2 Nov 2024.
12. Ecarnot F, Quenot JP, Besch G, Piton G. Ethical challenges involved in obtaining consent for research from patients hospitalized in the intensive care unit. Ann Transl Med. 2017;5(Suppl 4):S41.
13. Nicholls SG, Carroll K, Zwarenstein M, et al. The ethical challenges raised in the design and conduct of pragmatic trials: an interview study with key stakeholders. Trials. 2019;20:765.
14. Yazici H. Is the STAR trial really a safety trial? J Rheumatol. 2004;12:2528.
15. The Nuremberg Code https://en.wikipedia.org/wiki/Nuremberg_Coden. Accessed 21 Jan 2024.
16. World Medical Association. Ethical principles for medical research involving human subjects. Eur J Emerg Med. 2001;8:221–3.
17. Nardini C. The ethics of clinical trials Ecancermedicalscience. 2014;8:387.
18. Ford I, Norrie J. Pragmatic trials. N Engl J Med. 2016;375:454–63.
19. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. BMJ. 2015;8(350):h2147. https://doi.org/10.1136/bmj.h2147. (PMID: 25956159).
20. Smith GF. Artificial intelligence in drug safety and metabolism. Methods Mol Biol. 2022;2390:483–501.
21. Hotez PJ. The deadly rise of anti-science: a scientist's warning. JHU Press; 2023.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.