

The “1-year-death number needed to treat” for comparing the impact of distinct interventions on patient outcomes

Brett N. Hryciw MD, Finlay A. McAlister MD MSc, Meltem Tuna MSc PhD, Carl van Walraven MD MSc

■ Cite as: *CMAJ* 2019 November 11;191:E1242-9. doi: 10.1503/cmaj.190367

After concluding that a randomized controlled trial (RCT) is both internally and externally valid, consumers of evidence may want to consider an intervention’s overall importance compared with other interventions or health technologies. Such an exercise contributes to our perception of an intervention’s importance to health. This process could also help health care organizations determine which of an assortment of new technologies should be introduced first, or identify processes of care that should be subjected to quality-of-care reviews and optimization. Finally, it might also help with treatment selection for an individual patient; for example, in a patient with newly diagnosed left ventricular systolic dysfunction, should one start treatment with angiotensin-converting-enzyme (ACE) inhibition or β -blockade?

However, making comparisons across trials and between different technologies can be difficult because studies differ with respect to patient populations, comparators (placebo v. active control interventions or active controls of differing efficacy), concomitant treatments, follow-up times and outcomes. Comparisons across trials from different eras are even more difficult given secular trends in health outcomes over time. Each of these factors can influence the baseline absolute risk of an event and, therefore, interstudy comparisons.

To help with this, we propose here a method to compare the impact of diverse interventions applied to different populations over varying observation times and with distinct outcomes using a summary statistic we call the 1-year-death number needed to treat (NNT). We illustrate the application of this method using a convenience sample of landmark RCTs. We then present the studied interventions in a league table of 1-year-death NNTs to show how the statistic can be used to gauge the impact of interventions on patient outcomes.

What is the 1-year-death NNT?

Although reported outcomes vary extensively between studies, all-cause mortality is commonly reported and is an outcome whose clinical importance is arguably independent of both the patient population and treatment type. We therefore used all-cause mortality as the basis for the 1-year-death NNT. This statis-

KEY POINTS

- Cross-study comparisons of health care interventions are difficult because of varying patient populations, comparators, concomitant treatments and outcomes, which all affect the absolute risk of events.
- The 1-year-death number needed to treat (NNT) standardizes comparisons of treatment impact by focusing on all-cause mortality 1 year from treatment initiation using NNT methodology to account for distinct baseline risks.
- It is interpreted as the annual number of people requiring treatment and observation to avoid 1 death, with most beneficial treatments having positive values closest to 0 and most harmful treatments having negative values closest to 0.
- The 1-year-death NNT permits the standardized comparison of diverse health care technologies, which could allow physicians to prioritize distinct treatment options for an individual patient or health care organizations to determine which new technologies to introduce first.

tic summarizes the influence of treatment on patient survival by accounting for between-study variations in baseline death risk and follow-up duration. It is calculated as

$$1\text{-year death NNT} = \text{outcome time} / \text{ARR}$$

in which ARR is the absolute risk reduction (calculated as death risk_{control group} – death risk_{intervention group}), with death risks in each treatment arm expressed as proportions from 0 to 1 and calculated as the number dying divided by the number randomly assigned). Outcome time is the number of years after randomization when death status was measured in the study. If the study did not follow all patients for the same amount of time, the outcome time is the average observation time. Some interventions have a finite treatment duration, with death status measured later on; in such cases, outcome time occurs at the end of the “treatment cycle,” which includes both treatment time and the subsequent lag time before measurement of death status.

The 1-year-death NNT can have values that range from negative to positive infinity. The most beneficial treatments will have

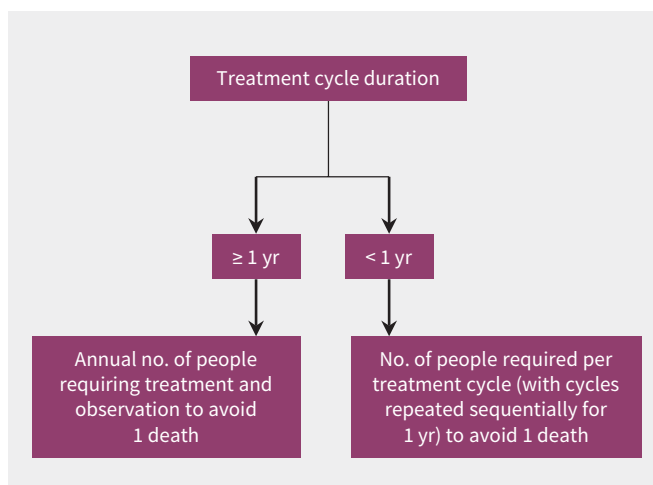


Figure 1: Interpretation of the 1-year-death number needed to treat. Note: treatment cycle duration = time from start of treatment to measurement of death status.

positive values closest to zero. Negative 1-year-death NNT values indicate interventions that increase death risk (with the absolute value of the 1-year-death NNT being interpreted as the “number needed to treat to harm”). Interpretation of the 1-year-death NNT depends on the treatment cycle duration (Figure 1). If the treatment cycle is 1 year or longer, the 1-year-death NNT is the annual number of people requiring treatment and observation to avoid 1 death. If the treatment cycle is less than a year, the 1-year-death NNT is the number of people required per treatment cycle (with cycles repeated sequentially for 1 year) to avoid 1 death. For example, consider a 1-year-death NNT of 5. If the treatment cycle is 2 years, this means that 5 people need treatment and observation for 1 year to avoid a single annual death; if the treatment cycle is 6 months, this means that 5 people need treatment and observation for 6 months twice (for a total of 10 treated people or 5 patient-years) to avoid a single annual death.

How can the 1-year-death NNT be applied?

To illustrate application of the 1-year-death NNT, we used as our sampling frame all RCTs included in *2 Minute Medicine's The Classics in Medicine: Summaries of the Landmark Trials*.¹ This is a collection of 190 “landmark clinical studies,” defined by the authors as those identified by academic clinicians in general internal medicine and its subspecialties as “practice-changing,” with a focus on common medical conditions (Dr. Andrew Cheung, St. Joseph’s Healthcare Hamilton, Hamilton, Ont.: personal communication, 2019). We limited our convenience sample to RCTs (since these provide the least biased evaluations of treatment effects) and those reporting all-cause mortality (regardless of whether it was the study’s primary outcome). Further methods and results are presented in Appendix 1 (available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.190367/-/DC1).

Table 1 summarizes the 31 RCTs reporting a significant difference in all-cause death risk between treatment groups. These trials included patients with congestive heart failure ($n = 8$), coro-

nary artery disease ($n = 5$), intensive care issues ($n = 4$), established (or at a high risk of attaining) cardiovascular disease ($n = 3$), infectious diseases ($n = 3$), cirrhosis and its complications ($n = 3$), arrhythmia ($n = 2$), and 1 each of diabetes, premature labour and cancer. Publication dates ranged from 1972 to 2013.

Five trials (16.1%) involved a procedural intervention, 13 trials (41.9%) involved continuous treatments (i.e., daily medication), and all other studies ($n = 13$, 41.9%) involved finite treatment durations ranging from 6 hours to 3 months (Table 2). Studies randomly assigned a mean of 5912 patients (range 61 to 45 852). The mean outcome time was 1.7 years after randomization (range 1 wk to 7.6 yr). Patient loss to follow-up was minimal in all studies except one.¹⁰ Death risks in control groups varied extensively between studies (range 2.8% to 88.5%, median 22.7%, mean 26.7%). A total of 27 studies reported on interventions that decreased death risk, and the remaining 4 identified interventions that increased death risk.

The median 1-year-death NNT (in absolute values) was 18. The intervention saving the greatest number of lives per year (in this case, neonatal lives) was betamethasone during premature labour, with a 1-year-death NNT of 0.16 (95% confidence interval [CI] 0.09 to 0.64).² Given this study’s treatment cycle duration of less than 1 year, this means (Figure 1) that 0.16 mothers in premature labour required 1 day of betamethasone every week for a year (for a total of $0.16 \times 52 = 8.3$ people treated and observed) to avoid 1 death. Eleven interventions were beneficial and associated with a 1-year-death NNT of less than 10 (betamethasone in premature labour, prone ventilation in severe adult respiratory distress syndrome, prednisolone in alcoholic hepatitis, goal-directed resuscitation in sepsis, albumin in spontaneous bacterial peritonitis, voriconazole instead of amphotericin B in immunocompromised invasive aspergillosis, dexamethasone in acute meningitis, hypothermia after resuscitation, transjugular intrahepatic portosystemic shunting in acute variceal bleeding, tranexamic acid in trauma, and enalapril in severe congestive heart failure).²⁻¹² The largest 1-year-death NNT (for prolonged tamoxifen after initial 5 years of treatment in early breast cancer) was 556 (95% CI 272 to ∞).

Four trials (12.9%) found a significantly increased death risk with the study’s intervention (including strict glucose control for critical care patients [1-year-death NNT -9.4 , 95% CI -63.8 to -5.1],¹³ encainide or flecainide for frequent premature ventricular contractions after myocardial infarction [1-year-death NNT -21 , 95% CI -47 to -13],¹⁸ warfarin for symptomatic intracranial arterial stenosis [1-year-death NNT -34 , 95% CI -148 to -19],²¹ and intensive glucose control in high-risk patients with type 2 diabetes [1-year-death NNT -334 , 95% CI $-\infty$ to -54]³⁰).

What can we learn from estimating the 1-year-death NNT?

By calculating the 1-year-death NNT, we were able to standardize both changes in death risk between treatment groups and the time after randomization when death risk was calculated. This let us directly compare the impact of very distinct health care technologies on survival (Table 2). These benefits ranged from a life

Table 1 (part 1 of 2): Summary of landmark randomized controlled trials in which treatments significantly reduced all-cause mortality

Study	Cochrane bias risk*	Run in†	Multi-centre‡	Industry funded§	Inclusion criteria	Intervention¶	Treatment duration	Control**
Liggins and Howie ²	L	N	N	Y	Premature labour 24–36 wk	Betamethasone phosphate plus betamethasone acetate 6 mg	1 d	Cortisone acetate 6 mg
Guérin et al. ³	U2	N	Y	N	Severe ARDS with ET for < 36 h	Prone ventilation > 16 h/d	4 d	Supine ventilation
Ramond et al. ⁴	L	N	Y	N	Alcoholic hepatitis on biopsy with DF > 32 or HE	Prednisolone 40 mg/d for 28 d	28 d	Placebo
Rivers et al. ⁵	L	N	N	Y	Age ≥ 18 yr, severe sepsis, septic shock or sepsis syndrome	Goal-directed resuscitation for 6 h	6 h	None
Sort et al. ⁶	U2	N	Y	N	Age 18–80 yr, ascitic PMN > 250	Albumin 1.5 g/kg day 1, 1.0 g/kg day 3	3 d	None
Herbrecht et al. ⁷	H(2,3)	N	Y	Y	Immunocompromised, definite or likely invasive aspergillosis	Voriconazole 4–6 mg/kg/d	12 wk	Amphotericin B 1–1.5 mg/kg
de Gans and van de Beek ⁸	L	N	Y	Y	Suspected meningitis	Dexamethasone 10 mg IV every 6 h	4 d	Placebo
Hypothermia after Cardiac Arrest Study Group ⁹	U3	N	Y	Y	Witnessed cardiac arrest from VF or VT, ROSC within 60 min	Hypothermia 32–34°C for 1 d	1 d	Normothermia
García-Pagán et al. ¹⁰	H(2,3)	N	Y	N	Cirrhosis, acute variceal bleeding, treatment with EGD and VA drugs	TIPS within 72 h	–	Endoscopic treatment
Williams-Johnson et al. ¹¹	L	N	Y	Y	Trauma, hemorrhage (or risk of such)	Tranexamic acid 1 g IV bolus, over 8 h	8 h	Placebo
CONSENSUS Trial Study Group ¹²	L	N	Y	Y	CHF, NYHA 4	Enalapril 2.5–20 mg twice daily	∞	Placebo
Finfer et al. ¹³	U3	N	Y	N	Admitted to ICU for anticipated > 3 d	Target glucose 4.5 to 6.0 mmol/L	4.2 d	Target glucose ≤ 10 mmol/L
Rose et al. ¹⁴	U3	N	Y	Y	CHF, NYHA 4, low peak oxygen consumption and a contraindication to transplantation	Left ventricular assist device	–	None
Malmberg et al. ¹⁵	H(1,2)	N	Y	N	Acute MI, treated diabetes or random glucose > 11.1 mmol/L	IV insulin every 24 h, then subsequently 4 times/d	≥ 3 min	“Conventional therapy”
Chen et al. ¹⁶	L	N	Y	Y	ST changes or LBBB within 24 h of symptoms	Clopidogrel 75 mg/d plus ASA 162 mg/d	15 d	ASA 162 mg/d
Pitt et al. ¹⁷	L	N	Y	Y	LVEF < 35, NYHA 3–4	Spirolactone 25–50 mg/d	∞	Placebo
Echt et al. ¹⁸	U1	Y	Y	N	MI last 2 years, ≥ 6 PVC/h, LVEF < 55%, response to treatment	Encainide or flecainide	∞	Placebo
MERIT-HF Study Group ¹⁹	L	Y	Y	Y	CHF and NYHA ≥ 2, LVEF < 40%	Metoprolol XL 12.5–200 mg/d	∞	Placebo

Table 1 (part 2 of 2): Summary of landmark randomized controlled trials in which treatments significantly reduced all-cause mortality

Study	Cochrane bias risk*	Run in†	Multi-centre‡	Industry funded§	Inclusion criteria	Intervention¶	Treatment duration	Control**
Moss et al. ²⁰	U2	N	Y	N	Prior MI, LVEF < 30	ICD	–	None
Chimowitz et al. ²¹	L	N	Y	Y	Age ≥ 40 yr, TIA or stroke from 50%–99% ICA stenosis, Rankin score ≤ 3	Warfarin INR 2–3	∞	ASA 1300 mg/d
Bardy et al. ²²	U2	N	Y	Y	Age ≥ 18 yr, NYHA 2–3 CHF, LVEF < 35%	Shock only, single-lead ICD	–	None
Yusuf et al. ²³	L	Y	Y	Y	CHF, LVEF < 35%	Enalapril 2.5–10 mg twice daily	∞	Placebo
β-Blocker Heart Attack Trial ²⁴	L	N	Y	Y	Age 30–69 yr, acute MI	Propranolol 40–80 mg 3 times/d	∞	Placebo
Pfeffer et al. ²⁵	L	Y	Y	Y	Age 21–80 yr, LVEF < 40%, 3–16 d post-MI	Captopril 6.25–25 mg 3 times/d	∞	Placebo
Wallentin et al. ²⁶	L	N	Y	Y	ACS presenting within 24-h onset	Ticagrelor 90 mg twice daily	∞	Clopidogrel 75 mg/d
Farkouh et al. ²⁷	H(2,3)	N	Y	Y	Diabetes, > 70% stenosis in ≥ 2 major epicardial vessels	CABG	–	PCI
Granger et al. ²⁸	L	N	Y	Y	(Atrial fibrillation or atrial flutter) + (age ≥ 75 yr or previous stroke or TIA or systemic embolism or CHF or diabetes or hypertension)	Apixaban 5 mg twice daily	∞	Warfarin INR 2–3
Yusuf et al. ²⁹	L	Y	Y	Y	Age > 55 yr, CAD, stroke, PVD or diabetes, and (hypertension, increased total cholesterol, low HDL, cigarette smoking or microalbuminuria)	Ramipril 10 mg/d	∞	Placebo
Gerstein et al. ³⁰	U2	N	Y	Y	Age 40–79 yr, type 2 diabetes and HbA _{1c} > 7.5%, high risk	Intensive treatment targeting HbA _{1c} < 6%	∞	Treatment targeting HbA _{1c} 7%–7.9%
Ridker et al. ³¹	U3	Y	Y	Y	LDL < 3.4, HS-CRP > 2 mg/L	Rosuvastatin 20 mg/d	∞	Placebo
Davies et al. ³²	U2	N	Y	Y	Early breast cancer, tamoxifen 5 yr, disease-free	Tamoxifen	5 yr	None

Note: ACS = acute coronary syndrome, ARDS = adult respiratory distress syndrome, ASA = acetylsalicylic acid, CABG = coronary artery bypass graft, CAD = coronary artery disease, CHF = congestive heart failure, DF = discriminant function (an algorithm to risk stratify severity of alcoholic hepatitis), EGD = esophagogastroduodenoscopy, ET = endotracheal tube, HDL = high-density lipoprotein, HE = hepatic encephalopathy, HbA_{1c} = glycated hemoglobin, HS-CRP = high-sensitivity C-reactive protein, ICA = intracranial artery, ICD = implantable cardioverter-defibrillator, ICU = intensive care unit, INR = international normalized ratio, IV = intravenously, LBBB = left bundle branch block, LDL = low-density lipoprotein, LVEF = left ventricular ejection fraction, MI = myocardial infarction, N = no, NYHA = New York Heart Association Functional Classification, PVD = peripheral vascular disease, PCI = percutaneous coronary intervention, PMN = polymorphonuclear cells, PVC = premature ventricular contractions, ROSC = return of spontaneous circulation, TIA = transient ischemic attack, TIPS = transjugular intrahepatic portosystemic shunt, VA = vasoactive, VF = ventricular fibrillation, VT = ventricular tachycardia, XL = extended release, Y = yes, – = 1 procedure or surgery, ∞ = infinite or ongoing.

*Based on the Cochrane Collaboration's tool for assessing risk of bias in randomized trials³³ focusing on domains most relevant to this study (allocation concealment, double-blinding and completeness of follow-up). L = low risk, U = uncertain risk, H = high risk. 1 = no allocation concealment, 2 = no double-blinding, 3 = completeness of follow-up not cited or < 95%.

†“Y” if study had prereandomization observation time followed by criteria necessary for randomization.

‡“Y” if patients were recruited at > 1 study site.

§“Y” if study cited financial support from any company.

¶Treatment for the intervention group.

**Treatment for the control (or comparator) group.

“being saved” annually after treatment of 8.3 mothers in premature labour with corticosteroids³⁴ to continuing 556 women with early breast cancer on tamoxifen (after an initial 5 years of treatment) for another year.³² Table 2 shows that, even after limiting ourselves to the “best of the best” landmark studies (i.e., recognized landmark RCTs of interventions that significantly influenced all-cause mortality), large variations remained in the interventions’ influence on patient survival. These results highlight that the critical review of RCTs should not stop with “significantly decreased all-cause mortality.”

The 1-year-death NNT and Table 2 help put RCT outcomes into perspective and better gauge the importance of interventions on patient survival. However, the interventions presented in these 31 studies should all be considered special. All-cause mortality is critical to patients and is almost invariably measured accurately by researchers, making it a key RCT outcome. Interventions that significantly change the likelihood of death beyond that expected by chance alone are important. If desired, the formula presented above could be modified to calculate 1-year standardized NNTs for any outcome that one cares to compare between studies.

What are the potential limitations of our example?

Any method used to sample all published studies evaluating treatments has its drawbacks. We used a collection of studies subjectively identified by experts to be important to and influential on clinical practice.¹ However, there is considerable overlap between studies included in *The Classics*¹ and other collections of key studies; almost a third of the trials published between 1991 and 2003 in *The Classics* were included in Ioannidis’ seminal report of landmark RCTs,³⁵ and more than 40% of RCTs in Hochman’s collection of key studies³⁶ were in *The Classics*.¹ Our study may have excluded key RCTs reporting important treatment effects on all-cause survival. Further work is required to systematically identify all other methodologically robust RCTs that report a significant influence of treatment on all-cause survival and adding them to our table.

Two issues arise from our results being based on RCT outcomes. First, with the exception of ACE inhibitors^{12,23,25} and implantable cardioverter-defibrillators^{20,22} for left ventricular systolic dysfunction, results in our league table are based on a single

Table 2 (part 1 of 2): All-cause mortality outcomes in landmark randomized controlled trials with statistically significant results

Study	Treatment, no.*	Control, no.†	Treatment duration‡	Outcome time, yr§	Death risk¶			p value§§	NNT-outcome time¶¶	1-year-death NNT (95% CI)***
					Treatment**	Control††	ARR‡‡			
Liggins and Howie ²	126	100	1 d	0.0192	0.063	0.18	0.117	0.02	8.6	0.16 (0.09 to 0.64)
Guérin et al. ³	237	229	4 d	0.0767	0.236	0.41	0.174	<0.001	5.7	0.4 (1.0 to 2.7)
Ramond et al. ⁴	32	29	28 d	0.4932	0.125	0.552	0.427	0.001	2.3	1.2 (0.8 to 2.3)
Rivers et al. ⁵	130	133	6 h	0.1644	0.385	0.526	0.142	0.03	7.1	1.2 (0.6 to 7.3)
Sort et al. ⁶	63	63	3 d	0.2466	0.222	0.413	0.19	0.03	5.3	1.3 (0.7 to 7.9)
Herbrecht et al. ⁷	144	133	12 wk	0.2308	0.292	0.414	0.122	0.02	8.2	1.9 (1.0 to 23)
de Gans and van de Beek ⁸	157	144	4 d	0.1538	0.07	0.146	0.076	0.04	13.2	2.0 (1.1 to 27)
Hypothermia after Cardiac Arrest Study Group ⁹	137	138	1 d	0.5	0.409	0.551	0.142	0.02	7	3.5 (1.9 to 20)
García-Pagán et al. ¹⁰	32	31	–	1.3333	0.125	0.387	0.262	0.01	3.8	5.1 (2.1 to 18)
Williams-Johnson et al. ¹¹	10 060	10 067	8 h	0.0769	0.145	0.16	0.015	0.0035	67.6	5.2 (3.1 to 16)
CONSENSUS Trial Study Group ¹²	127	126	∞	0.9863	0.362	0.524	0.162	0.003	6.2	6.1 (3.5 to 24)
Finfer et al. ¹³	3010	3012	4.2 d	0.2466	0.275	0.249	–0.026	0.02	–38.3	–9.5 (–64 to –5.1)
Rose et al. ¹⁴	61	61	–	2.1667	0.672	0.885	0.213	<0.001	4.7	10 (2.8 to 14)
Malmberg et al. ¹⁵	306	314	≥ 3 min	1	0.186	0.261	0.075	0.027	13.4	13 (7.1 to 104)
Chen et al. ¹⁶	22 961	22 891	15 d	0.0767	0.075	0.081	0.005	0.03	184.2	14 (7.4 to 147)
Pitt et al. ¹⁷	822	841	∞	2	0.345	0.459	0.113	0.001	8.8	18 (12 to 30)

Table 2 (part 2 of 2): All-cause mortality outcomes in landmark randomized controlled trials with statistically significant results

Study	Treatment, no.*	Control, no.†	Treatment duration‡	Outcome time, yr§	Death risk¶			p value§§	NNT-outcome time¶¶	1-year-death NNT (95% CI)***
					Treatment**	Control††	ARR‡‡			
MERIT-HF Study Group ¹⁹	1990	2001	∞	1	0.073	0.108	0.036	< 0.001	28.1	28 (19 to 56)
Moss et al. ²⁰	742	490	–	1.6667	0.142	0.198	0.056	0.016	17.7	30 (17 to 127)
Chimowitz et al. ²¹	289	280	∞	1.85	0.097	0.043	–0.054	0.02	–18.5	–34 (–148 to –19)
Bardy et al. ²²	829	847	–	3.7917	0.22	0.288	0.069	0.007	14.6	55 (45 to 185)
Yusuf et al. ²³	1,285	1,284	∞	3.45	0.352	0.397	0.045	0.0036	22	76 (42 to 428)
β-Blocker Heart Attack Trial ²⁴	1916	1921	∞	2.0833	0.072	0.098	0.026	< 0.005	38.7	81 (69 to 365)
Pfeffer et al. ²⁵	1115	1116	∞	3.5	0.204	0.246	0.042	0.019	23.8	83 (46 to 480)
Wallentin et al. ²⁶	9333	9291	∞	1	0.043	0.054	0.012	< 0.001	85.4	85 (56 to 181)
Farkouh et al. ²⁷	947	953	–	5	0.088	0.12	0.032	0.049	31.3	156 (84 to 1084)
Yusuf et al. ²⁹	4645	4652	∞	3.5	0.104	0.122	0.019	0.005	53.9	270 (159 to 881)
Gerstein et al. ³⁰	5128	5123	∞	1.9	0.05	0.04	–0.01	0.04	–95.3	–334 (∞ to –54)
Ridker et al. ³¹	8901	8901	∞	7.6	0.022	0.028	0.006	0.02	181.7	345 (99 to ∞)
Davies et al. ³²	6454	6440	5 yr	7.6	0.214	0.227	0.014	0.04	73.2	556 (272 to ∞)

Note: ARR = absolute risk reduction, CI = confidence interval, NNT = number needed to treat, – = 1 procedure or surgery, ∞ = ongoing treatment throughout observation.

*Number of people randomly assigned to intervention group.

†Number of people randomly assigned to control group.

‡Duration of active treatment for the intervention group.

§Outcome time is the time from randomization at which death count or risk was measured. In closed studies, it equals the time patients received active treatment plus subsequent observation time required to achieve observed death risk. In open studies (i.e., outcome time varies between patients in study), the average follow-up time was used.

¶Neonatal (not maternal) death risk.

**Proportion dead in intervention group (no. dead ÷ no. randomly assigned).

††Proportion dead in control group (no. dead ÷ no. randomly assigned).

‡‡Absolute risk reduction for death, calculated as death risk for control group – death risk for intervention group.

§§Two-sided *p* value for comparison of death risk in groups.

¶¶Number needed to treat to prevent 1 death by outcome time.

***Number needed to treat to prevent 1 death by 1 year. If negative, this is the number needed to treat to cause 1 death by 1 year (the 1-year-death number needed to harm).

RCT of interventions. Single studies, even those that are large and internally valid, can return exaggerated results that can be smaller (or sometimes even absent) when trials are replicated or when the intervention is applied in clinical practice.³⁵ This issue could be addressed using data from meta-analyses rather than single studies. Second, the intervention effect measured in RCTs may generalize poorly.³⁷

Other issues regarding our example should be kept in mind. First, some of the health care technologies compared in our study are quite distinct, making their direct comparison relevant only in a numerical sense regarding the 1-year-death NNT. Second, we reported risk using a simple proportion because this was reported in all studies. Risk-measurement methods that account for observation time will more accurately measure risk; such measurements can be used instead of proportions if calculable in the RCTs of all technologies that are to be compared. Finally, we used the *p* value for the intertreatment difference in death risk to

select treatments that were highlighted in Table 2. This dichotomization of a continuous measure (i.e., the likelihood that differences seen exceed that arising by chance) will unfairly exclude some studies (such as those with a *p* value of 0.05) whose importance is not materially distinct from those presented here.

What are the potential limitations of using the 1-year-death NNT statistic?

Several potential limitations of the number needed to treat (NNT) will also apply to the 1-year-death NNT.^{38,39} First, both statistics will tend to increase extensively when baseline risks decrease.³⁸ Second, baseline death risks tend to exhibit secular decreases over time as concomitant therapies and other aspects of medical care improve; for example, 1-year mortality for adults admitted to hospital in Ontario decreased by 20% between 1994 and 2009. This makes comparisons of NNT and

1-year-death NNT between years problematic.⁴⁰ Third, death risk estimates — and differences between treatment groups — get larger the longer patients are observed. Therefore, the NNT (the reciprocal of the absolute risk reduction) will decrease as time from randomization to outcome measurement increases.³⁸ The 1-year-death NNT avoids this limitation by standardizing all NNT time frames to 1 year after randomization. However, this standardization assumes that the relative influence of treatment on death risk is the same 1 year after randomization as at the study's reported outcome time. If treatment effect changes significantly over time, the 1-year-death NNT may return a biased estimate. The outcome times of almost half of the studies we analyzed ($n = 15$) were very close to 1 year (between 6 mo and 2 yr), thereby reducing the influence of this potential bias. Fourth, the 1-year-death NNT simplifies the comparison of absolute death risk with potentially important loss of information. Therefore, one should always examine the absolute death risks in each patient group to truly appreciate treatment effects. Fifth, the 1-year-death NNTs reported in Table 2 (just as with all NNT measures) represent the average value of patients in those trials; values could vary widely in particular subgroups. Finally, in RCTs having variable observation time for patients, death risks were most accurate when censoring is considered using survival analysis techniques.⁴¹ Since most RCTs in our worked example did not use survival analysis or provide the necessary data to permit recalculation, death risks were summarized using proportions. This made interstudy comparisons possible but could underestimate death risk in studies with long observation times and large treatment effects.

Several limitations are particular to the 1-year death NNT. First, in using all-cause mortality to permit cross-study and cross-technology comparisons, we assumed that all deaths are equal. This is arguably not true since some deaths are symptomatically worse than others (e.g., contrast a person dying in their sleep to someone dying with progressive, incapacitating respiratory failure). Further research might explore the possibility of weighting deaths using health utilities and health-related quality of life measures.⁴² Second, the 1-year-death NNT also does not account for life years lost owing to the death. Neonatal death during premature labour is more consequential with regard to population health than death in an older adult. Further work could explore weighting 1-year-death NNT by expected life years lost. Third, interventions might provide important benefits to patients without influencing the likelihood of death. Fourth, 1-year-death NNT is but one statistic that could be used to compare health care technologies. Other factors, such as cost and disease prevalence, should also be considered. Finally, the 1-year-death NNT will vary with the outcome time chosen in the RCT. Consider a study in which deaths in both treatment groups are primarily clustered close to the start of observation; the corresponding survival curves will be increasingly flat over time. If such a study chose an observation time that notably exceeded the period during which deaths occurred, 1-year-death NNT would return an inappropriately large value. This possibility supports the point made above: to truly understand the 1-year-death NNT (as with any summary statistic), one should always examine its individual data components.

Conclusion

Despite its limitations, the 1-year-death NNT represents a useful statistic to help consumers of medical evidence contextualize potential impacts of health care interventions. It could also help health care organizations prioritize technologies for evaluation or funding and help individual physicians discriminate between distinct treatment options.

References

- Succi MD, Carr LH, Cheung A. *2 Minute Medicine's the classics in medicine: summaries of the landmark trials*. 1st ed. Boston: 2 Minute Medicine Inc.; 2015.
- Liggins GC, Howie RN. A controlled trial of antepartum glucocorticoid treatment for prevention of the respiratory distress syndrome in premature infants. *Pediatrics* 1972;50:515-25.
- Guérin C, Reignier J, Richard JC, et al.; PROSEVA Study Group. Prone positioning in severe acute respiratory distress syndrome. *N Engl J Med* 2013;368:2159-68.
- Ramond MJ, Poynard T, Rueff B, et al. A randomized trial of prednisolone in patients with severe alcoholic hepatitis. *N Engl J Med* 1992;326:507-12.
- Rivers E, Nguyen B, Havstad S, et al.; Early Goal-Directed Therapy Collaborative Group. Early goal-directed therapy in the treatment of severe sepsis and septic shock. *N Engl J Med* 2001;345:1368-77.
- Sort P, Navasa M, Arroyo V, et al. Effect of intravenous albumin on renal impairment and mortality in patients with cirrhosis and spontaneous bacterial peritonitis. *N Engl J Med* 1999;341:403-9.
- Herbrecht R, Denning DW, Patterson TF, et al.; Invasive Fungal Infections Group of the European Organisation for Research and Treatment of Cancer and the Global Aspergillus Study Group. Voriconazole versus amphotericin B for primary therapy of invasive aspergillosis. *N Engl J Med* 2002;347:408-15.
- de Gans J, van de Beek D; European Dexamethasone in Adulthood Bacterial Meningitis Study Investigators. Dexamethasone in adults with bacterial meningitis. *N Engl J Med* 2002;347:1549-56.
- Hypothermia after Cardiac Arrest Study Group. Mild therapeutic hypothermia to improve the neurologic outcome after cardiac arrest. *N Engl J Med* 2002;346:549-56.
- García-Pagán JC, Caca K, Bureau C, et al.; Early TIPS (Transjugular Intrahepatic Portosystemic Shunt) Cooperative Study Group. Early use of TIPS in patients with cirrhosis and variceal bleeding. *N Engl J Med* 2010;362:2370-9.
- Williams-Johnson JA, McDonald AH, Strachan GG, et al. Effects of tranexamic acid on death, vascular occlusive events, and blood transfusion in trauma patients with significant haemorrhage (CRASH-2): a randomised, placebo-controlled trial. *West Indian Med J* 2010;59:612-24.
- CONSENSUS Trial Study Group. Effects of enalapril on mortality in severe congestive heart failure. Results of the Cooperative North Scandinavian Enalapril Survival Study (CONSENSUS). *N Engl J Med* 1987;316:1429-35.
- NICE-SUGAR Study Investigators; Finfer S, Chittock DR, Su SY, et al. Intensive versus conventional glucose control in critically ill patients. *N Engl J Med* 2009;360:1283-97.
- Rose EA, Gelijs AC, Moskowitz AJ, et al.; Randomized Evaluation of Mechanical Assistance for the Treatment of Congestive Heart Failure (REMATCH) Study Group. Long-term use of a left ventricular assist device for end-stage heart failure. *N Engl J Med* 2001;345:1435-43.
- Malmberg K, Rydén L, Efendic S, et al. Randomized trial of insulin-glucose infusion followed by subcutaneous insulin treatment in diabetic patients with acute myocardial infarction (DIGAMI study): effects on mortality at 1 year. *J Am Coll Cardiol* 1995;26:57-65.
- Chen ZM, Jiang LX, Chen YP, et al.; COMMIT (Clopidogrel and Metoprolol in Myocardial Infarction Trial) collaborative group. Addition of clopidogrel to aspirin in 45,852 patients with acute myocardial infarction: randomised placebo-controlled trial. *Lancet* 2005;366:1607-21.
- Pitt B, Zannad F, Remme WJ, et al. The effect of spironolactone on morbidity and mortality in patients with severe heart failure. *N Engl J Med* 1999;341:709-17.
- Echt DS, Liebson PR, Mitchell LB, et al. Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. *N Engl J Med* 1991;324:781-8.
- Effect of metoprolol CR/XL in chronic heart failure: Metoprolol CR/XL Randomised Intervention Trial in Congestive Heart Failure (MERIT-HF). *Lancet* 1999;353:2001-7.

20. Moss AJ, Zareba W, Hall WJ, et al.; Multicenter Automatic Defibrillator Implantation Trial II Investigators. Prophylactic implantation of a defibrillator in patients with myocardial infarction and reduced ejection fraction. *N Engl J Med* 2002;346:877-83.
21. Chimowitz MI, Lynn MJ, Howlett-Smith H, et al.; Warfarin-Aspirin Symptomatic Intracranial Disease Trial Investigators. Comparison of warfarin and aspirin for symptomatic intracranial arterial stenosis. *N Engl J Med* 2005;352:1305-16.
22. Bardy GH, Lee KL, Mark DB, et al. Amiodarone or an implantable cardioverter-defibrillator for congestive heart failure. *N Engl J Med* 2005;352:225-37.
23. SOLVD Investigators; Yusuf S, Pitt B, Davis CE, et al. Effect of enalapril on survival in patients with reduced left ventricular ejection fractions and congestive heart failure. *N Engl J Med* 1991;325:293-302.
24. A randomized trial of propranolol in patients with acute myocardial infarction: I. mortality results. *JAMA* 1982;247:1707-14.
25. Pfeffer MA, Braunwald E, Moyé LA, et al. Effect of captopril on mortality and morbidity in patients with left ventricular dysfunction after myocardial infarction. Results of the survival and ventricular enlargement trial. The SAVE Investigators. *N Engl J Med* 1992;327:669-77.
26. Wallentin L, Becker RC, Budaj A, et al. Ticagrelor versus clopidogrel in patients with acute coronary syndromes. *N Engl J Med* 2009;361:1045-57.
27. Farkouh ME, Domanski M, Sleeper LA, et al.; FREEDOM Trial Investigators. Strategies for multivessel revascularization in patients with diabetes. *N Engl J Med* 2012;367:2375-84.
28. Granger CB, Alexander JH, McMurray JJV, et al.; ARISTOTLE Committees and Investigators. Apixaban versus warfarin in patients with atrial fibrillation. *N Engl J Med* 2011;365:981-92.
29. Heart Outcomes Prevention Evaluation Study Investigator; Yusuf S, Sleight P, Pogue J, et al. Effects of an angiotensin-converting enzyme inhibitor, ramipril, on cardiovascular events in high-risk patients [published erratum in *N Engl J Med* 2000;342:748, 1376]. *N Engl J Med* 2000;342:145-53.
30. Action to Control Cardiovascular Risk in Diabetes Study Group; Gerstein HC, Miller ME, Byington RP, et al. Effects of intensive glucose lowering in type 2 diabetes. *N Engl J Med* 2008;358:2545-59.
31. Ridker PM, Danielson E, Fonseca FAH, et al.; JUPITER Study Group. Rosuvastatin to prevent vascular events in men and women with elevated C-reactive protein. *N Engl J Med* 2008;359:2195-207.
32. Davies C, Pan H, Godwin J, et al.; Adjuvant Tamoxifen: Longer Against Shorter (ATLAS) Collaborative Group. Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. *Lancet* 2013;381:805-16.
33. Higgins JPT, Altman DG, Goetzsche PC, et al.; Cochrane Bias Methods Group; Cochrane Statistical Methods Group. The Cochrane Collaboration tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928.
34. Bardy GH, Lee KL, Mark DB, et al.; Sudden Cardiac Death in Heart Failure Trial (SCD-HeFT) Investigators. Amiodarone or an implantable cardioverter-defibrillator for congestive heart failure. *N Engl J Med* 2005;352:225-37.
35. Ioannidis JPA. Contradicted and initially stronger effects in highly cited clinical research. *JAMA* 2005;294:218-28.
36. Hochman ME. *50 studies every doctor should know: the key studies that form the foundation of evidence based medicine*. Revised ed. Oxford (UK): Oxford University Press; 2013.
37. Bailey KR. Generalizing the results of randomized clinical trials. *Control Clin Trials* 1994;15:15-23.
38. McAlister FA. The “number needed to treat” turns 20 — and continues to be used and misused. *CMAJ* 2008;179:549-53.
39. Suissa D, Brassard P, Smiechowski B, et al. Number needed to treat is incorrect without proper time-related considerations. *J Clin Epidemiol* 2012;65:42-6.
40. van Walraven C. Trends in 1-year survival of people admitted to hospital in Ontario, 1994–2009. *CMAJ* 2013;185:E755-62.
41. Zee J, Xie SX. The Kaplan–Meier method for estimating and comparing proportions in a randomized controlled trial with dropouts. *Biostat Epidemiol* 2018;2:23-33.
42. Horsman J, Furlong W, Feeny D, et al. The Health Utilities Index (HUI): concepts, measurement properties and applications. *Health Qual Life Outcomes* 2003;1:54.

Competing interests: None declared.

This article has been peer reviewed.

Affiliations: Department of Medicine (Hryciw, van Walraven), The Ottawa Hospital, Ottawa, Ont.; Division of General Internal Medicine (McAlister), University of Alberta, Edmonton, Alta.; ICES uOttawa (Tuna, van Walraven); Ottawa Hospital Research Institute (Tuna, van Walraven), Ottawa, Ont.

Contributors: Brett Hryciw and Carl van Walraven contributed to the conception and data collection. Brett Hryciw, Carl van Walraven and Meltem Tuna contributed to the data analysis. All of the authors contributed to the data interpretation. Carl van Walraven and Finlay McAlister drafted the article, which all of the authors reviewed. All of the authors gave final approval of the version to be published and agreed to be accountable for all aspects of the work.

Correspondence to: Carl van Walraven, carlv@ohri.ca