Surrogate Endpoints in Cardio-Thoracic Trials: A Call for Better Reporting and Improved Interpretation of Trial Findings

Oriana Ciani, PhD [1]

Anthony Muchai Manyara, PhD [2]

Rod S. Taylor, [2, 3]

1. *SDA Bocconi School of Management, Center for Research on Health and Social Care Management, Milan, IT.*
2. *MRC/CSO Social and Public Health Sciences Unit, Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK.*
3. *Robertson Centre for Biostatistics, Institute of Health and Well Being, University of Glasgow, Glasgow, UK.*

Corresponding Author: Oriana Ciani, PhD

Mailing address: SDA Bocconi School of Management, Center for Research on Health and Social Care Management
MEO Building, Room W210, II Floor
Via Sarfatti 10, 20136 Milan (IT)

Tel: + 39 02 5836 2596

Fax: +39 02 5836 2598

E-Mail: oriana.ciani@unibocconi.it

KEYWORDS

Clinical trials; surrogate endpoints; reporting guidelines; research waste.

Evidence to inform regulatory, funding, and clinical decisions about health interventions should come from well-conducted clinical trials. Despite the challenges associated to their design, conduct and generalizability, randomised controlled trials (RCTs) are typically considered as the studies providing the least biased estimates of treatment effects [1].

The selection of outcomes to be evaluated in RCTs and what specific measures to use for them – the trial endpoints – is a complex yet pivotal step in the process. Outcomes should allow to reasonably monitor the trial objectives and provide meaningful information for the different stakeholders and evidence-based decision-making at different levels [2].

For instance, in an RCT evaluating a robot-assisted vs video-assisted thoracic surgery in the early-stage lung cancer, the clinicians might prioritise overall or disease-free survival, whilst the patients might be highly interested in the effect of treatment on their functional status or health-related quality of life. Both concepts are typically measured through patient-reported outcome measures (PROMs), which can be either generic (e.g., EQ-5D-5L [3]) or condition-specific (e.g., EORTC QLQ-LC29 [4]). Healthcare managers and policymakers might look for information on the cost-effectiveness profile of the experimental approach, or the budget impact expected from its implementation within 3 to 5 years [5].

The different needs and perspectives represented by these groups make usually necessary to collect multiple endpoints in trials, which are categorised as primary, secondary or exploratory. The primary endpoint in an RCT informs the sample size calculation. It is recognised as the basis to assess the comparative efficacy/effectiveness of an experimental intervention vs the alternative. Secondary endpoints generally complement the results of the trial on the primary endpoint with additional elements that may relate to other facets of effectiveness, further mechanisms of action, or safety. Exploratory endpoints, when introduced, may cover very rare outcomes or outcomes that may be relevant to test novel hypotheses [2].

There is a common definition initially proposed by the National Institutes of Health in the United States and frequently used in the research community for primary outcomes relevant for patients and patient care, i.e., characteristics or variables that reflect how patients feel, function or how long they survive [6]. In the drive towards faster patient access to innovation, surrogate endpoints have been increasingly adopted as biomarkers (i.e., objective indicators of a biological or pathogenic process) to substitute and predict for how a person feels, functions, or how long she survives. Surrogate endpoints observed in cardiovascular or

thoracic trials are the time to chest drain removal [7], sinus rhythm restoration [8], graft patency [9] or percentage change in computed tomography pleural fluid volume [10] (Table 1). As these examples reveal, surrogate endpoints could be laboratory or instrumental measures, imaging findings, or physiological measures.

Results from surrogate endpoints generally accrue more quickly than from patient-centred outcomes, thus allowing for clinical trials with shorter follow-up periods and smaller sample sizes. In cardiovascular disease, for example, a trial powered on a surrogate primary endpoint, such as carotid artery intima-media thickness, might involve a few hundred patients followed up for weeks or months, whereas the same trial powered on cardiovascular mortality or major cardiovascular morbidity might require thousands of patients followed up for years [11]. Reducing trial sample size and duration means that trials are less expensive, which make surrogate endpoints attractive to manufacturers and research sponsors. Despite the potential appeal of surrogates, their use remains controversial because they may not capture the full benefit-risk profile of an intervention. For example, Schneider and colleagues have recently considered as unreliable the identification of irreversible myocardial injury in cardiac surgery based only on biomarker release (troponin I and T and creatine kinase myocardial band) [12]. In addition, superiority on a surrogate endpoint may not translate into benefits for patients, or if it did, the health care system may not judge the benefits to be a good value for money. Ridker and Torres observed, after reviewing 324 consecutive cardiovascular trials, that surrogate primary endpoints were more likely to be associated to positive treatment effects (77 of 115 trials; 67%) than trials that reported final patient-relevant primary outcomes (113 of 209 trials; 54%, p = 0.02) [13]. In this respect, we have previously shown that trials using surrogate primary outcomes are more likely to report larger treatment effects (adjusted ratio of odds ratio: 1.46, 95% confidence interval = 1.05 – 2.04, p = 0.03) than trials reporting final patient-relevant primary outcomes and this overestimation is not explained by differences in the risk of bias or characteristics of these trials (such as blinding or outcome attrition) [14]. As well as exaggerating treatment effects, this can lead to overestimation of the cost-effectiveness of treatments.

Therefore, surrogate endpoints magnify the uncertainty around the interpretation of treatment effects observed in clinical trials and their meaning in terms of true patient benefit; hence, their adoption is only helpful if they are validated predictors of patient-relevant outcomes [15]. A valid surrogate endpoint is such that the effect of the intervention on it

reliably predicts the effect of the intervention on the patient-centred outcome. Several statistical approaches have been developed to establish this link, the most common requiring to perform a regression-based analysis within the context of a meta-analysis of (preferably patient-level data) from several RCTs that have measured both for the surrogate and the final patient-relevant outcome of interest [16].

Evidence around the validity of a putative surrogate endpoint should inform the trial protocol design, conduct, reporting and interpretation of trial findings. However, there exist no consensus guidelines specific for the reporting of trials and protocols with a surrogate primary endpoint. Hence, reporting standards often lack transparency, thus potentially encouraging misleading interpretation of trials findings. In 2010, a review of a sample of published randomised clinical trials revealed that one in five used a surrogate as a primary outcome. Of these trials, 57% clearly reported that the primary outcome was a surrogate. Only 35% also discussed the validity of the surrogate measure [17]. The authors concluded that better reporting on surrogate endpoints is urgently needed to limit unwarranted conclusions and uncritical acceptance of new treatments based on inappropriate documentation and the use of surrogate measures.

In January 2022, we launched a Medical Research Council (MRC) funded project to develop CONSORT (Consolidated Standards of Reporting Trials) and SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) extensions for RCTs with a primary surrogate endpoint ('SPIRIT-SURROGATE' and 'CONSORT-SURROGATE'). These extensions are important. Incomplete, unclear, or inaccurate reporting of trial design, conduct and analysis can hinder interpretability, lead to erroneous conclusions about the long-term safety and efficacy of interventions and impact on decisions relying on those poorly reported research findings. Development of the extensions will closely follow the Enhancing Quality and Transparency of Health Research (EQUATOR) methodology [18]. To make this process inclusive and increase the acceptability and usefulness of these tools for the wide research community, we would like to invite various stakeholders, clinicians and surgeons, particularly those with an interest or experience in using surrogate endpoints in cardio-thoracic trials, to contribute to the e-Delphi in preparation that will be launched in fall 2022. Project progress and updates are available at https://www.gla.ac.uk/spirit-consort-surrogate.

Improved reporting of surrogate endpoint-based trials should enable more effective scrutiny and interpretation of evidence by the end-users of trial findings and more efficient use of

resources dedicated to clinical research.

Surrogate endpoint	**Final patient-relevant endpoint that is substituted and predicted for**
Time to chest drain removal [7]	Length of hospital stay
Sinus rhythm restoration [8]	Overall survival
Graft patency [9]	Major cardiovascular events
Percentage change in computed tomography pleural fluid volume [10]	Readmission for heart failure

Table 1. Example of surrogate endpoints used in cardiothoracic randomised controlled trials and the patient-relevant endpoints they are intended to substitute and predict for.

REFERENCES

[1] Kramer BS, Wilentz J, Alexander D, Burklow J, Friedman LM, Hodes R, Kirschstein R, Patterson A, R odgers G, Straus SE. Getting it right: being smarter about clinical trials. PLoS Med. 2006 Apr;3(6):e144.

[2] McLeod C, Norman R, Litton E, Saville BR, Webb S, Snelling TL. Choosing primary endpoints for clin ical trials of health care interventions. Contemp Clin Trials Commun. 2019 Nov 12;16:100486.

[3] EuroQoL Research Foundation. *EQ-5D-5L | About.* https://euroqol.org/eq-5d-instruments/eq-5d-5l-abo ut/ (09 August 2022, date last accessed).

[4] Koller M, Shamieh O, Hjermstad MJ, Hornslien K, Young T, Chalk T, Ioannidis G, Harle A, Johnson C D, Tomaszewski KA, Serpentini S, Pinto M, van der Weijst L, Janssens A, Morag O, Chie WC, Arraras JI, Pomp ili C, Jungraithmayr W, Hechtner M, Katsochi D, Müller K, Gräfenstein L, Schulz C, Bottomley A; European Or ganisation for Research and Treatment of Cancer (EORTC) Quality of Life Group; EORTC Lung Cancer Group; European Society of Thoracic Surgeons. Psychometric properties of the updated EORTC module for assessing q uality of life in patients with lung cancer (QLQ-LC29): an international, observational field study. Lancet Oncol. 2020 May;21(5):723-732.

[5] Patel YS, Hanna WC, Fahim C, Shargall Y, Waddell TK, Yasufuku K, Machuca TN, Pipkin M, Baste J M, Xie F, Shiwcharan A, Foster G, Thabane L. RAVAL trial: Protocol of an international, multi-centered, blinde d, randomized controlled trial comparing robotic-assisted versus video-assisted lobectomy for early-stage lung c ancer. PLoS One. 2022 Feb 2;17(2):e0261767.

[6] De Gruttola VG, Clax P, DeMets DL, Downing GJ, Ellenberg SS, Friedman L, Gail MH, Prentice R, W ittes J, Zeger SL. Considerations in the evaluation of surrogate endpoints in clinical trials. summary of a Nationa l Institutes of Health workshop. Control Clin Trials. 2001 Oct;22(5):485-502.

[7] Holbek BL, Christensen M, Hansen HJ, Kehlet H, Petersen RH. The effects of low suction on digital dr ainage devices after lobectomy using video-assisted thoracoscopic surgery: a randomized controlled trial†. Eur J Cardiothorac Surg. 2019 Apr 1;55(4):673-681.

[8] Sharples LD, Mills C, Chiu YD, Fynn S, Holcombe HM, Nashef SAM. Five-year results of Amaze: a ra ndomized controlled trial of adjunct surgery for atrial fibrillation. Eur J Cardiothorac Surg. 2022 Mar 26:ezac181 .

[9] Houlind K, Fenger-Grøn M, Holme SJ, Kjeldsen BJ, Madsen SN, Rasmussen BS, Jepsen MH, Ravkilde J, Aaroe J, Hansen PR, Hansen HS, Mortensen PE; DOORS Study Group. Graft patency after off-pump coronar y artery bypass surgery is inferior even with identical heparinization protocols: results from the Danish On-pump Versus Off-pump Randomization Study (DOORS). J Thorac Cardiovasc Surg. 2014 Nov;148(5):1812-1819.e2.

[10] Hooper CE, Edey AJ, Wallis A, Clive AO, Morley A, White P, Medford AR, Harvey JE, Darby M, Zaha n-Evans N, Maskell NA. Pleural irrigation trial (PIT): a randomised controlled trial of pleural irrigation with nor mal saline versus standard care in patients with pleural infection. Eur Respir J. 2015 Aug;46(2):456-63.

[11] Buhr KA. Surrogate end points in secondary analyses of cardiovascular trials. Prog Cardiovasc Dis. 201 2 Jan-Feb;54(4):343-50.

[12] Schneider U, Mukharyamov M, Beyersdorf F, Dewald O, Liebold A, Gaudino M, Fremes S, Doenst T. The value of perioperative biomarker release for the assessment of myocardial injury or infarction in cardiac sur gery. Eur J Cardiothorac Surg. 2022 Mar 24;61(4):735-741.

[13] Ridker PM, Torres J. Reported outcomes in major cardiovascular clinical trials funded by for-profit and not-for-profit organizations: 2000-2005. JAMA. 2006 May 17;295(19):2270-4.

[14] Ciani O, Buyse M, Garside R, Pavey T, Stein K, Sterne JA, Taylor RS. Comparison of treatment effect s izes associated with surrogate and final patient relevant outcomes in randomised controlled trials: meta-epidemio logical study. BMJ. 2013 Jan 29;346:f457.

[15] Gallo L, Eskicioglu C, Braga LH, Farrokhyar F, Thoma A. Users' guide to the surgical literature: how to assess an article using surrogate end points. Can J Surg. 2017 Aug;60(4):280-287.

[16] Ensor H, Lee RJ, Sudlow C, Weir CJ. Statistical approaches for evaluating surrogate outcomes in clinic al trials: A systematic review. J Biopharm Stat. 2016;26(5):859-79.

[17] la Cour JL, Brok J, Gøtzsche PC. Inconsistent reporting of surrogate outcomes in randomised clinical tr ials: cohort study. BMJ. 2010 Aug 18;341:c3653.

[18] Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guide lines. PLoS Med. 2010 Feb 16;7(2):e1000217.