One of the features of dabrafenib-induced pyrexia is that most patients eventually tolerate the treatment quite well, which raises the question as to the adaptive mechanism involved. Again, the inflammasome hypothesis offers an explanation in that inflammasome receptors are subject to ubiquitination and degradation by autophagy. Further support for their involvement in pyrexia might be obtained by successful treatment of patients with blocking antibodies against IL1-β, such as canakinumab. Canakinumab is usually given subcutaneously every 4 weeks, and a single injection might even be considered for prophylactic administration. Specific inhibitors of NLRP3, such as MCC950 and CY-09,[10] also exist, but whether or not dabrafenib interacts with the NLRP3 receptor remains under investigation. NLRP3 inhibitors would have utility in establishing the pathogenesis of the fevers and lead to further assessment of their use relative to the existing treatment of pyrexia with corticosteroids. New approaches to managing this distressing side-effect will be particularly important as this drug combination transitions into the adjuvant setting.

*Peter Hersey, Jessamy C Tiffen, Stuart J Gallagher
Melanoma Oncology Group, Centenary Institute, University of Sydney, Sydney 2050, NSW, Australia
peter.hersey@sydney.edu.au

We declare no competing interests.

1 Dreno B, Ribas A, Larkin J, et al. Incidence, course, and management of toxicities associated with cobimetinib in combination with vemurafenib in the coBRIM study. *Ann Oncol* 2017; **28:** 1137–44.
2 Vera J, Paludo J, Kottschade L, et al. Case series of dabrafenib-trametinib-induced pyrexia successfully treated with colchicine. *Support Care Cancer* 2019; **27:** 3869–75.
3 Menzies AM, Ashworth MT, Swann S, et al. Characteristics of pyrexia in BRAFV600E/K metastatic melanoma patients treated with combined dabrafenib and trametinib in a phase I/II clinical trial. *Ann Oncol* 2015; **26:** 415–21.
4 Gogas HJ, Flaherty KT, Dummer R. Adverse events associated with encorafenib plus binimetinib in the COLUMBUS study: incidence, course and management. *Eur J Cancer* 2019; **119:** 97–106.
5 Atkinson V, Long GV, Menzies AM, et al. Optimizing combination dabrafenib and trametinib therapy in BRAF mutation-positive advanced melanoma patients: guidelines from Australian melanoma medical oncologists. *Asia Pac J Clin Oncol* 2016; **12** (suppl 7)**:** 5–12.
6 Maeda T, Yoshino K, Yamashita C, et al. Dynamics of neutrophil and C-reactive protein reflect the clinical course of pyrexia during combination therapy with dabrafenib and trametinib. *J Dermatol* 2019; **46:** 716–19.
7 Kim HY, Duong JK, Gonzalez M, et al. Pharmacokinetic and cytokine profiles of melanoma patients with dabrafenib and trametinib-induced pyrexia. *Cancer Chemother Pharmacol* 2019; **83:** 693–704.
8 Kozak W, Kluger MJ, Soszynski D, et al. IL-6 and IL-1 beta in fever. Studies using cytokine-deficient (knockout) mice. *Ann N Y Acad Sci* 1998; **856:** 33–47.
9 Hajek E, Krebs F, Bent R, et al. *BRAF* inhibitors stimulate inflammasome activation and interleukin 1 beta production in dendritic cells. *Oncotarget* 2018; **9:** 28 294–308.
10 Jiang H, He H, Chen Y, et al. Identification of a selective and direct NLRP3 inhibitor to treat inflammatory disorders. *J Exp Med* 2017; **214:** 3219–38.

Multiplicity in oncology randomised controlled trials: a threat to medical evidence?

Randomised controlled trials offer several advantages over observational studies in establishing whether cancer treatments improve patient outcomes. First, randomised controlled trials minimise bias and confounding. Observational studies judge treatments that were chosen not at random but deliberately by physicians, and the decision to treat or not treat an individual might be affected by patient characteristics, a type of bias called confounding by indication. Randomised controlled trials prescribe therapy and avoid this limitation. Second, randomised controlled trials restrict multiple hypothesis testing. Randomised controlled trials typically address a specific research question with a prespecified primary outcome and a predefined statistical analysis plan. This approach reduces the ability of investigators to assess multiple outcomes, which in turn minimises the risk of a spurious (eg, false positive) finding. For these reasons, David Sackett placed randomised controlled trials at the apex of the hierarchy of evidence.[1] Yet in the past 10 years, one of these assumptions is increasingly being questioned. Randomised controlled trials have proliferated in number, often test similar compounds with similar molecular targets, and are often run in redundant and duplicative trial portfolios or agendas. For this reason, multiplicity has emerged as a new threat to the validity of conclusions drawn from randomised controlled trials. In this Comment, we explore this issue.

Multiplicity, or the number of times an analysis can be run, has long been recognised as a threat to causal inference from observational studies. Patel and colleagues[2] found that if researchers are afforded

many opportunities to analyse a dataset and allowed to choose the variables they can adjust for, often both positive and negative associations can be reached for a single exposure–outcome relationship. The concern is that many investigators might explore a similar question in parallel, and only the researchers who discover significant findings report results, leading to a literature that finds both positive and negative associations and one that rarely reports null findings. When researchers randomly picked ingredients from a cookbook and searched for observational studies linking those nutrients to cancer risk, they found this was the case.[3]

In the past 10 years, there has been substantial growth in the number of randomised controlled trials that evaluate individual anticancer drugs, including those of the same class and mechanism of action. Consider the immune checkpoint inhibitor class of medications. Drugs targeting PD-1 or PD-L1 are being tested in more than 2000 clinical trials with an estimated enrollment of more than 500 000 participants.[4] Multiple randomised controlled trials are being done to evaluate the combination of many types of chemotherapy plus a checkpoint inhibitor for many types of tumours, different checkpoint inhibitors from different companies, or combinations of targeted and immunotherapy drugs; the probability that some results will emerge as favourable by chance alone is high.

Multiplicity in randomised controlled trials might explain the results of clinical trials of bevacizumab. Bevacizumab, a monoclonal antibody against VEGF, has been tested in at least 50 randomised controlled trials for people with solid tumours.[5] The results of these studies have varied even within a single malignancy. In non-small-cell lung cancer, one randomised controlled trial reported improved outcomes when bevacizumab was added to carboplatin and paclitaxel, yet a similar trial in the same population with cisplatin and gemcitabine as the chemotherapy backbone did not.[6] The same is true in colorectal cancer; a randomised controlled trial evaluating the addition of bevacizumab to chemotherapy with irinotecan, fluorouracil, and leucovorin in front-line colorectal cancer, by Hurwitz and colleagues,[7] found an overall survival benefit, but a similar randomised controlled trial evaluating the addition of bevacizumab to the FOLFOX (fluorouracil, oxaliplatin, leucovorin) regimen, led by Saltz and colleagues,[8] did not find any benefit. When trials of bevacizumab have been reported as positive, they have generally used a primary outcome of progression-free survival, whereas few have shown improvement in overall survival. Even in the trials that are deemed positive, the effect size of bevacizumab has been modest. These findings raise the question of whether bevacizumab truly improves survival. The drug might increase radiographical tumour response, and might lengthen radiographical progression, but are the few instances of survival benefit merely the artifact of multiple testing?

In a recent analysis,[9,10] we did a multiplicity adjustment to all of the published randomised controlled trials that have evaluated chemotherapy with or without bevacizumab, regardless of the type of malignancy being treated. Such adjustments are often done if multiple outcomes are analysed within a clinical study, but here we applied that same principle to multiple studies examining the same drug. When we applied a Bonferroni correction, a stringent adjustment for multiplicity, most of the trials that reported a statistically significant improvement in overall survival lost that significance and only one remained positive. For 30 trials reporting substantial improvements in progression-free survival, applying the multiplicity correction changed that conclusion in nine of them. Others might prefer to apply different false-discovery adjustments that are less stringent than the Bonferroni correction, which assumes that all data are independent. However, all false-discovery methods would lead to conclusions that are less optimistic than those reported for each trial individually. Philosophically, there is little difference between a single mega-randomised controlled trial comparing bevacizumab in combination with a chemotherapy backbone versus the same chemotherapy backbone in all solid tumours and subsequently doing a subgroup analysis by tumour type versus doing individual studies in all of these cohorts. However, based on the current framework for interpretation of oncology trials, we would always adjust for multiplicity in the former analysis or consider results to be hypothesis generating, or both, yet in the latter analysis, we often take the results as gospel.

The risk of clinical practice being influenced by multiplicity is even greater for trials evaluating immune checkpoint inhibitors. Drugs targeting PD-1 and PD-L1

are generating high clinical interest. There are more than 20 such compounds in preclinical testing and they are being tested in various combinations. Many clinical trials are duplicative, running in clinical settings where PD-1 antibodies are thought to exert meaningful activity—for instance, non-small-cell lung cancer and melanoma. With so many trials testing similar hypotheses with similar drugs, the likelihood that any one trial will yield a significant result is increased by the large number of times that something has been tested. Here, we need to correct for the portfolio of trials not within a single pharmaceutical company but across all companies. Of course, robust results (eg, multiple positive trials in non-small-cell lung cancer and melanoma) represent real gains, but if only some immunotherapy trials are positive and others are not, particularly in the same clinical setting, then caution is warranted.

Historically, oncologists did not have to worry about multiple hypothesis testing in clinical trials for the simple reason that there were small numbers of similar drugs developed simultaneously, and restricted research budgets with which to invest, thus the risk of multiple hypothesis testing was minimal. Accordingly, it was quite unlikely to have concurrent randomised controlled trials addressing the same clinical question. However, the ever-escalating cost of cancer drugs and their high-profit margins has incentivised the pharmaceutical industry to pursue the evaluation of similar drugs in duplicative, large portfolios of clinical studies. This approach has been further facilitated by the reluctance of the US Food and Drug Administration to compel manufacturers to test their drugs against existing compounds or to restrict the number of accelerated approvals.

In the 20th century, randomised controlled trials leapt to the top of the hierarchy of evidence-based medicine because of their ability to minimise confounding and isolate the therapeutic effect of interventions. Studies were assessed individually on the basis of their conduct, transparency, and reporting. In the 21st century, large portfolios of clinical trials testing similar drugs for similar questions have raised the bar. For randomised controlled trials to remain atop the hierarchy, not only should each individual study be free from bias, but each individual result should be judged alongside the entire portfolio of trials. The rare so-called positive trial within a sea of negative studies is more likely a false positive than a true positive. The oncology community will need to be mindful of the risk of multiplicity leading to false conclusions from randomised controlled trials, especially for checkpoint inhibitors. To deliver high-quality cancer care in a sustainable health ecosystem, clinicians, investigators, and policy makers will need to identify therapies that offer benefits that are substantial in magnitude and not statistical artifacts.

*Vinay Prasad, Christopher M Booth
Department of Medicine and Center for Health Care Ethics, Oregon Health & Science University, Portland, OR 97202, USA (VP); Department of Oncology, Queen's University, Kingston, ON, Canada (CMB); and Division of Cancer Care and Epidemiology, Queen's Cancer Research Institute, Kingston, ON, Canada (CMB)
prasad@ohsu.edu

VP is supported by the Arnold Foundation. CMB is supported as the Canada Research Chair in Population Cancer Care. We gratefully acknowledge Elizabeth Eisenhauer and Ian Tannock for comments on an earlier draft of this manuscript. VP reports receiving royalties from his book *Ending Medical Reversal* and an advance for the forthcoming book *Malignant: How Bad Policy and Bad Evidence Work Against People With Cancer*. VP's work is funded by the Laura and John Arnold Foundation and VP has received honoraria for Grand Rounds/lectures from several universities, medical centres, and professional societies and payments for contributions to Medscape. VP also hosts the podcast Plenary Session, which has Patreon backers. CMB declares no competing interests.

1 Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ. Users' guides to the medical literature. IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group. *JAMA* 1995; **274:** 1800–04.
2 Patel CJ, Burford B, Ioannidis JP. Assessment of vibration of effects due to model specification can demonstrate the instability of observational associations. *J Clin Epidemiol* 2015; **68:** 1046–58.
3 Schoenfeld JD, Ioannidis JP. Is everything we eat associated with cancer? A systematic cookbook review. *Am J Clin Nutr* 2013; **97:** 127–34.
4 Tang J, Yu JX, Hubbard-Lucey VM, Neftelinov ST, Hodge JP, Lin Y. Trial watch: the clinical trial landscape for PD1/PDL1 immune checkpoint inhibitors. *Nat Rev Drug Discov* 2018; **17:** 854–55.
5 Roviello G, Bachelot T, Hudis CA, et al. The role of bevacizumab in solid tumours: a literature based meta-analysis of randomised trials. *Eur J Cancer* 2017; **75:** 245–58.
6 Reck M, von Pawel J, Zatloukal P, et al. Phase III trial of cisplatin plus gemcitabine with either placebo or bevacizumab as first-line therapy for nonsquamous non-small-cell lung cancer: AVAil. *J Clin Oncol* 2009; **27:** 1227–34.
7 Hurwitz H, Fehrenbacher L, Novotny W, et al. Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *N Engl J Med* 2004; **350:** 2335–42.
8 Saltz LB, Clarke S, Diaz-Rubio E, et al. Bevacizumab in combination with oxaliplatin-based chemotherapy as first-line therapy in metastatic colorectal cancer: a randomized phase III study. *J Clin Oncol* 2008; **26:** 2013–19.
9 Tao DL, Gay ND, Prasad V. Statistical significance of bevacizumab trials when considering the portfolio of all studies. *Proc Am Soc Clin Oncol* 2018 **36:** 6551 (abstr).
10 Tao DL, Gay ND, Prasad V. Statistical significance of bevacizumab trials when considering the portfolio of all studies. *Proc Am Soc Clin Oncol* 2018; **36** (suppl 5)**:** 6551–51.