Special Issue: Where is Science Going?

Vol. 5, No. 2 (2022)
ISSN: 2532-5876

DOI: 10.13133/2532-5876/17620

Data, Knowledge, and Theory: A Biostatistician's Perspective

Stuart G. Baker[a*]

[a] *Division of Cancer Prevention, National Cancer Institute [SGB], United Stets of America*

***Corresponding author:** Stuart G. Baker, Division of Cancer Prevention, National Cancer Institute, 9609 Medical Center Dr, Room 5E606, MSC 9789, Bethesda, MD 20892-9789 USA; e-mail: sb16i@nih.gov

Abstract

This commentary, inspired by a recent opinion piece of noted biologist Paul Nurse, overviews the interplay between data and various types of scientific knowledge within the realms of prediction, data patterns, causal inference, and scientific theory.

Keywords: prediction, cancer prevention, carcinogenesis, causal inference, machine learning

Citation: Baker, S G, 2022, "Data, Knowledge, and Theory: A Biostatisticians' Perspective", *Organisms: Journal of Biological Sciences*, vol. 5, no. 2. DOI: 10.13133/2532-5876/17620

Introduction

In a recently published opinion piece, the noted biologist Paul Nurse (2021) laments that the biological sciences are awash in data but sorely lacking in knowledge and theory. Less one think this is a new concern arising from our age of ubiquitous data, fifty years earlier, another British scientist, Leslie Foulds (1969), wrote

> "…Some investigators are fond of saying 'What we need is more facts'. The truth is that we have more "facts" than we know what to do with. Experimental analysis produced an alarming mass of empirical facts without providing an adequate language for their communication or effective concepts for their synthesis."

In fact, over one hundred years ago, the mathematician and physicists Henri Poincare (1902) made a more fundamental point that "Science is built of facts the way a house is built of bricks; but an accumulation of facts is no more science than a pile of bricks is a house." In more contemporary terms, scientific knowledge is not just data, but how data are used to improve understanding or prediction. This commentary expands on the viewpoint in Nurse (2021), discussing the interplay between data and scientific knowledge in prediction, data patterns, causal inference, and theory.

1. Prediction

In clinical science, an important type of scientific knowledge is the prediction of a binary outcome, such as disease onset, based on non-modifiable predictors. Prediction models can involve traditional statistical models, such as logistic regression, or algorithms in machine learning. The predictors, called features in the machine learning literature, can include risk factors, such as age or smoking status, baseline biomarker values, and images.

Investigators fit prediction models in a training sample, measure reproducibility in an internal validation sample (a random sample from same

population used to obtain training sample) and gauge generalizability in an external validation sample (a sample from a population different from the population from which the training sample was drawn) (Steyerberg & Harrell 2016). A prediction model that performs well in an internal validation sample can perform poorly in an external validation sample (Bleeker *et al.* 2003). Types of generalizability for external validation include different time periods, geographic regions, data collection methods, and clinical settings (Justice *et al.* 1999). Classification performance varying by clinical setting, which is known as spectrum bias, is an important consideration that is often underappreciated (Ransohoff & Feinstein 1978). A good example involves prediction based on carcinoembryonic antigen (CEA). Although CEA in the blood almost perfectly classifies specimens as diagnosed colorectal cancer or no cancer (Thomson *et al.* 1969), it poorly predicts the development of colorectal cancer in asymptomatic persons (Thomas *et al.* 2015). External validation is a prerequisite for recommending a prediction model for clinical use (Ramspek *et al.* 2020). On a fundamental level, external validation is analogous to testing a scientific theory to see how well it makes a prediction.

An underappreciated aspect of many clinical prediction models is the importance of feature selection. In this regard, computer scientist Pedro Domingos (2012) remarked "At the end of the day, some machine learning projects succeed and some fail. What makes the difference? Easily the most important factor is the features used." In terms of prediction, investigators should not be concerned if different sets of features predict outcomes equally well in the validation sample. The occurrence of many models with good prediction or classification performance is called the Rashomon effect, after the Japanese movie *Rashomon*, which depicted an event from multiple viewpoints (Breiman 2001). A possible explanation for the Rashomon effect is that the observed features are likely imperfect proxies for unobserved true predictors.

In clinical prediction with well-defined features (such as biomarker level, age, and family history), investigators may favor the standard statistical approach of logistic regression over complex machine learning algorithms because both approaches often perform equally well (Christodoulou *et al.* 2019) while logistic regression is easier to interpret. When the predictor is an image, feature selection by humans can perform poorly (Le Cun 1998). Fortunately, the development of optimization methods that took advantage of more powerful computing led to deep learning methods with automatic feature selection, yielding substantially improved performance with imaging data (Le Cun 1998; Krizhevsky, Sutskever & Hinton 2012; Cao *et al.* 2018). These algorithms typically make predictions in bizarre ways, using features not visible to humans or secondary to human recognition (D'Amour 2021). The result is an unusual but useful form of scientific knowledge involving good prediction (Cao *et al.* 2018) but lacking scientific interpretation. Conclusions about the performance of a clinical prediction model depend on the metric used to summarize performance. In recent years, there has been a growing appreciation of the value of decision analytic metrics (Baker 2018; Vickers *et al.* 2019).

2. Data Patterns

Another type of scientific information is what I call data patterns. The goal is to identify relevant patterns in high-dimensional data that can suggest new theories. One example is correlation networks to understand responses or biological adaptations to stress (Gorban *et al.* 2021). Another example is principal components analysis, which uses linear models to reduce dimensionality and has numerous applications in biology (Giuliani 2017). A third example involves biologically relevant longitudinal response in high dimensional data. For example, Baker (2014) compared biologically relevant changes (linear, sigmoid, and impulse) among thousands of genes at 14 times in the embryonic development of two species of frogs. Sigmoid curves suggest saturation effects while impulse curves suggest a transient response leading to a new steady state. Types of comparative results were heteromorphy (curves with different shapes), heterochrony (curves with the same shape but different transition times), and heterometry (curves with the same shape but different magnitudes). The training data were odd numbered time points, and the validation data were the even numbered time points (Baker 2014).

3. Causal Inference

Another major type of scientific knowledge is causal inference, which involves drawing conclusions about outcome after changing a modifiable variable, such as treatment. Conceptually, causal inference tries to determine the outcome if one went back in time and gave subjects a different intervention (Rubin, 2005). Causal inference is usually divided into methods for analyzing observational data and methods for analyzing data from randomized trials.

Most causal inference methods applied to observational studies involve a multivariate adjustment using data from concurrent controls (participants enrolled simultaneously with the treatment group and followed over the same time period). The multivariate adjustment is needed to control for confounders, which are variables that affect both intervention and outcome. For example, in estimating the causal effect of increased exercise on cancer incidence, it is important to control for obesity, which is a confounder because it affects both exercise and the cancer outcome. One useful technique for improving causal inference in these studies is the method of propensity scores, which matches on estimated probabilities of receiving treatment (Rosenbaum & Rubin 1983; Austin 2011). Causal graphs can be useful for identifying observed confounders in complex scenarios (Pearl 2010). However, there is always a nagging concern that no matter how many observed confounders are included in the analysis, there may be an important confounder that was not observed, and that lack of adjustment for this unobserved confounder could lead to incorrect conclusions.

An example of another type of causal inference from observational studies are results from the paired availability design for historical controls (Baker & Lindeman 1994, Baker, Kramer & Lindeman 2019). Standard historical controls are subject to selection bias, as persons who receive treatment later often differ from those who receive treatment earlier. The paired availability design avoids selection bias by comparing outcomes, in each medical center, before and after treatment becomes more available, with a causal adjustment (Baker & Lindeman 1994; Imbens & Angrist 1994) for changes in the availability of treatment. However, as with all observational studies, there is no free lunch. The paired availability assumes no systematic temporal changes unrelated to treatment. If data are available, bias from such temporal changes can be mitigated using outcomes from medical centers with no change in treatment over time.

During the recent pandemic, there was frequent debate as to the quality of evidence from observational studies versus randomized trials. The most convincing form of scientific knowledge in causal inference studies is the randomized trial, which avoids some critical assumptions needed for causal inference with an observational study. However, as with many types of observational studies, assumptions are required to accommodate missing-data, noncompliance, and extrapolation to a target population. A good illustration of the value of a randomized trial over an observational study involves the Alpha-Tocopherol, Beta Carotene trial that randomized male smokers to either control alpha-tocopherol, beta carotene, or both supplements with an outcome of lung cancer incidence (Alpha-Tocopherol, Beta Carotene Cancer Prevention Study Group 1994). Prior to the trial, there was considerable evidence from observational studies that these supplements could prevent cancer (The ATBC Cancer Prevention Study Group 1994; Peto *et al.* 1981). However, the results of the trial contradicted the observational evidence—alpha-tocopherol had no effect on lung cancer incidence and beta-carotene increased lung cancer incidence. Because the randomized trial is the gold standard of causal inference, the results of the ATBC trial trumped the previous observational results.

4. Scientific Theories

Scientific theories provide explanations. For examples, inflammation, radiation, and viruses can cause cancer but a theory is needed to explain how they cause cancer. Debates about theories are crucial to scientific progress and applications. The dominant theory of carcinogenesis is the somatic mutation theory. However, the somatic mutation theory has not been "scientifically tested" (Loomans-Kropp & Umar 2019) and does not explain many puzzling experimental results (Baker 2021). Alternative theories of carcinogenesis (Sonnenschein & Soto 1999; Soto & Sonnenschein 2011; Brücher & Jamall 2014; Baker 2020; Carvalho 2020) are worth considering not only for scientific value but because an understanding of tumorigenesis is a foundation

for cancer prevention (Golemis *et al.* 2018; Loomans-Kropp & Umar, 2019).

Competition in theories, like competition in the marketplace, often leads to improvement. The 19th century geologist Thomas Chamberlin (1890) proposed the method of multiple working hypotheses, which involves evaluating several hypotheses and rejecting those that conflict with available data. One advantage of the method of working hypotheses is that

> "the reaction of one hypothesis upon another tends to amplify the recognized scope of each, and their mutual conflicts whet the discriminative edge of each" (Chamberlin, 1890).

A theory needs to be stated precisely to be useful. With a vague theory in which any outcome can be explained, there is no scientific knowledge (Feynman 1964). Band-aid approaches to modify a theory to explain puzzling phenomena are not convincing and will no longer hold as new challenges arise. For example, to try to explain puzzling aspects of the geocentric theory of planetary orbits, astronomers kept inventing new epicycles, small circles whose centers move around the circumference of larger circles, until the whole edifice became unwieldy (Maor 1998). The heliocentric theory of Copernicus (with later refinement by Kepler) obviated these ad hoc attempts to make theory fit the data. The somatic mutation theory of cancer, the prevailing paradigm of carcinogenesis, has become a patchwork of modifications to fit new observations (Soto & Sonnenschein 2007), which the mainstream cancer biology community has been slow to appreciate. For example, recent work on mutation fingerprints of cellular histories has shown that cells with three driver mutations are also readily found in normal tissue (Li *et al.* 2021; Moore *et al.* 2021; Naxerova 2021), a result that challenges the genetic definition of cancer. However, the mainstream response is not to consider non-genetic drivers of cancer but instead to speculate on new band-aids to the somatic mutation theory, namely, tissue-specific combinations of mutations in a more permissive microenvironment (Naxerova 2021).

Theories are useful in guiding experimentation and determining what quantities to observe. The noted economist Thomas Sowell (2012) said,

> "if there are two different theories, there should be some empirical evidence in principle that could distinguish what would happen under one theory from what would happen under the other."

For example, in cancer biology research, to "distinguish" the somatic mutation theory from the tissue organization field theory, Maffini *et al.* (2004) devised an elegant experiment involving rat mammary tissue recombination model. Their results showed that carcinogens target the stroma and not the epithelial tissue, which contradicts the somatic mutation theory of cancer and supports the tissue organization field theory.

Nurse (2021) advocates that scientists propose reasonable theories even if they later turn out to be incorrect. In a wonderful book on scientific investigation, Beverdige (1952) provides many examples of incorrect biological hypotheses that led to scientific progress. For example, the noted physiologist Claude Bernard hypothesized that nerve impulses induced chemical changes that heat the skin. To test this hypothesis, he experimented on rabbits, severing a cervical nerve to see if rabbit ear would become cooler. To his surprise, he found that ear became warmer, leading him to realize that nerves control the flow of blood through the arteries.

A challenge with proposing new theories is pushback related to the sunk cost fallacy. Individuals or institutions commit the sunk cost fallacy when they continue an endeavor solely as a result of previously invested resources including time, money, and effort (Arkes & Blumer 1985). Any unrecoverable sunk cost is irrelevant when deciding on future actions and ignoring the sunk cost fallacy can have dire consequences (Ronayne, Sgroi & Tuckwell 2021). Scientists are not immune to the sunk cost fallacy. A scientist who spends years designing experiments and writing papers based on particular theory would be inherently resistant to an alternative theory that jeopardizes the value of previous work. Proper skepticism is good, but it also important to be open to major shift in research directions if circumstances warrant. Government agencies can also suffer from the sunk cost fallacy. Because government agencies often value institutional interests above national interests (Sowell 2018), they may be reluctant to abandon a long-term research program for an initiative based on a compelling new theory. One organization that has successfully tackled the sunk cost fallacy is the research arm of the giant

tech company Alphabet, which rewards teams for discontinuing projects that are unlikely to succeed (Teller 2016).

Another type of pushback to new scientific theories comes from the "Machiavelli effect" (Hall 2021). In his famous political treatise, *The Prince*, Machiavelli wrote "the innovator has for enemies all those who have done well under the old conditions, and lukewarm defenders in those who may do well under the new." According to Hall (2021), scientific funding rewards established researchers and who are then resistant to new ideas.

Nurse (2021) also suggests that teaching that science has ideas, and is not just a litany of data, will motivate students. Even elementary school studies can be motivated by ideas and theories. In the popular American children's cartoon television show, "The Magic School Bus" (Cole & Degan 1994), the science teacher, Ms. Frizzle, tells her students to "Take chances, make mistakes, and get messy." Taking chances (trying new experiments), making mistakes (proposing reasonable theories that may be incorrect) and getting messy (engaging in the practical details of experimentation and data analysis) is a good advice for all current and future scientific researchers.

Conclusion

There is a productive interplay between data, knowledge, and theory. Accurate and informative data are essential to prediction, data patterns, causal inference, and scientific theory. Consider Kepler's trial-and-error method to discover the motion of the planets around the sun. At first, Kepler thought the motion was circular. But with accurate data from Tycho Brahe, Kepler realized that a circular motion did not fit the data, and an ellipsis was required (Feynman 1965). In clinical prediction, investigators need informative features to develop good models, and they need data from an external validation sample to investigate generalizability. Data patterns can suggest new theories in the fields of network correlations, principal component analysis, and biologically relevant longitudinal response curves. For causal inference with multivariate models, investigators need to measure and adjust for all important confounders. For causal inference with the paired availability design, investigators need to select medical centers with no changes in relevant protocols over time except for the increased availability of treatment.

According to Nurse (2021), theory is needed to capitalize on today's data rich world. Similarly, Hand (2016) noted that focusing only on data misses the point of science, which is to develop theories. Following Nurse (2021), investigators should be encouraged to propose reasonable theories. Theories should be stated as precisely as possible with multiple theories welcome, and, ideally with relatively little pushback from the sunk cost fallacy or the Machiavelli effect. Facilitating such a culture can make scientific investigation more exciting and dynamic for both researchers and students.

The most underappreciated aspect of the interplay among data, knowledge, and theory is that theory can guide experimentation and thereby lead to important new data which can then lead to new knowledge and better theories.

Declarations

Funding

This research was supported by the National Institutes of Health (NIH). Opinions expressed by the author are his own and this material should not be interpreted as representing the official viewpoint of the U.S. Department of Health and Human Services, the National Institutes of Health, or the National Cancer Institute. The study sponsor had no role in the design of the study; the collection, analysis, and interpretation of the data; the writing of the manuscript; and the decision to submit the manuscript for publication.

Competing interests

The author declares no competing interest.

References

Alpha-Tocopherol, Beta Carotene Cancer Prevention Study Group 1994, "The effect of vitamin E and beta carotene on the incidence of lung cancer and other cancers in male smokers", *N Engl J Med*, vol. 330, pp. 1029-1035.

Arkes HR & Blumer C 1985, "The psychology of sunk costs", *Organ Behav Hum Decis Process*, vol. 35, pp. 124-140.

Austin PC 2011, "An introduction to propensity score methods for reducing the effects of confounding in observational studies", *Multivariate Behav Res*, vol. 46, pp. 399-424.

Baker SG & Lindeman KS 1994, "The paired availability design: a proposal for evaluating epidural analgesia during labor", *Stat Med*, vol. 13, pp. 2269-2278.

Baker SG 2014, "Comparative analysis of biologically relevant response curves in gene expression experiments: heteromorphy, heterochrony, and heterometry", *Microarrays*, vol. 3, pp. 39-51.

Baker SG 2018, "Simple decision-analytic functions of the AUC for ruling out a risk prediction model and an added predictor", *Med Decis Making*, vol. 38, pp. 225-234.

Baker SG 2020, "Rethinking carcinogenesis: The detached pericyte hypothesis", *Med Hyp*, vol. 144: 110056.

Baker SG 2021, "The case for a cancer paradox initiative", *Carcinogenesis*, vol. 42, no. 8, pp. 1023-1025.

Baker SG, Kramer BS & Lindeman KL 2016, "Latent class instrumental variables. A clinical and biostatistical perspective", *Stat Med*, vol. 35, pp. 147-160 (correction 2019, vol. 38, p. 901).

Beverdige WVI 1952, *The art of scientific investigation*, London: Heinemann.

Bleeker SE, Moll HA, Steyerberg EW, Donders AR, Derksen-Lubsen G, Grobbee DE & Moons KG 2003, "External validation is necessary in prediction research: a clinical example", *J Clin Epidemiol*, vol. 56, pp. 826-832.

Breiman L 2001, "Statistical modeling: The two cultures (with comments and a rejoinder by the author)", *Stat Sci*, vol. 16, pp. 199–231.

Brücher BL & Jamall IS 2014, "Epistemology of the origin of cancer: a new paradigm", *BMC Cancer*, vol. 14: 331.

Cao C, Liu F, Tan H, *et al.* 2018, "Deep learning and its applications in biomedicine", *Genomics Proteomics Bioinformatics*, vol. 16, pp. 17-32.

Carvalho J 2020, "Cell reversal from a differentiated to a stem-like state at cancer initiation", *Front Oncol*, vol. 10: 541.

Chamberlin TC, 1890, "The method of multiple working hypotheses", *Science*, vol 15, pp. 92–96 (reprinted in Science, 1965, vol. 148, pp. 754–759).

Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY & Van Calster B 2019, "A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models", *J Clin Epidemiol*, vol. 110, pp. 12-22.

Cole J & Degan B (creators) 1994, *Magic school bus* [television series], 1994-1998.

D'Amour A 2021, "Revisiting Rashomon: a comment on 'the two cultures'", *Observational Studies*, vol. 7, pp. 59-63.

Domingos P 2012, "A few useful things to know about machine learning", *Communications of the ACM*, vol. 55, pp. 78–87.

Feymann R 1964, "Richard Feynman on the scientific method (1964)", from the Character of Physical Law. Messenger Lectures Cornell University [video]. Retrieved from https://www.youtube.com/watch?v=0KmimDq4cSU.

Feynman R 1965, *The character of physical law*, Cambridge: MIT Press.

Foulds L 1969, *Neoplastic development*, vol. 1, New York: Academic Press.

Golemis EA, Scheet P, Beck TN, *et al.* 2018, "Molecular mechanisms of the preventable causes of cancer in the United States", *Genes De*, vol. 32, pp. 868-902.

Gorban AN, Tyukina TA, Pokidysheva LI, Smirnova EV 2021, "Dynamic and thermodynamic models of adaptation", *Phys Life Rev*, vol. 37, pp.17-64.

Giuliani A 2017, "The application of principal component analysis to drug discovery and biomedical data", *Drug Discov Today*, vol. 22, pp. 1069-1076.

Hall JS 2021, *Where's my flying car*, San Francisco: Stripe Press.

Hand DJ 2016, "Editorial: 'Big data' and data sharing", *J Roy Stat Soc A*, vol. 179, pp. 629-631.

Imbens GW & Angrist JD 1994, "Identification and estimation of local average treatment effects", *Econometrica*, vol. 62, pp. 467-475.

Justice AC, Covinsky KE & Berlin JA 1999, "Assessing the generalizability of prognostic information", *Ann Intern Med*, vol. 130, pp. 515-524.

Krizhevsky A, Sutskever I & Hinton GE 2012, "Imagenet classification with deep convolutional neural networks", *Advances in neural information processing systems*, vol. 25, pp. 1097-1105.

LeCun Y, Bottou L, Bengio Y & Haffner P 1998, "Gradient-based learning applied to document recognition", *Proc IEEE*, vol. 86, pp. 2278–2324.

Li R, Di L, Li J. *et al.* 2021, "A body map of somatic mutagenesis in morphologically normal human tissues", *Nature*, vol. 597, pp. 398–403

Loomans-Kropp HA & Umar A 2019, "Cancer prevention and screening: the next step in the era of precision medicine", *NPJ Precision Onc*, vol. 3: 3.

Maffini MV, Soto AM, Calabro JM, Ucci AA & Sonnenschein C 2004, "The stroma as a crucial target in rat mammary gland carcinogenesis", *J Cell Sci*, vol. 117, pp. 1495-1502.

Maor E 1998, *Trigonometric delights*, Princeton: Princeton University Press.

Moore L, Cagan A, Coorens THH *et al.* 2021, "The mutational landscape of human somatic and germline cells", *Nature*, vol. 597, pp. 381–386.

Naxerova K 2021, "Mutation fingerprints encode cellular histories", *Nature*, vol. 597, pp. 334-336.

Nurse P 2021, "Biology must generate ideas as well as data", *Nature*, vol. 597, p. 305.

Pearl J 2010, "An introduction to causal inference", *Int J Biostat*, vol. 6: 7.

Peto R, Doll R, Buckley JD & Sporn MB 1981, "Can dietary beta-carotene materially reduce human cancer rates?" *Nature*, vol. 290, pp. 201-208.

Poincaré H 1902, *La science et l'hypothèse*, Paris: Ernest Flammarion.

Ramspek CL, Jager KJ, Dekker FW, Zoccali C & van Diepen M 2020, "External validation of prognostic models: what, why, how, when and where?", *Clin Kidney J*, vol. 24, pp. 49-58.

Ransohoff DF & Feinstein AR 1978, "Problems of spectrum and bias in evaluating the efficacy of diagnostic tests", N Engl J Med, vol. 299, pp. 926–930.

Ronayne D, Sgroi D & Tuckwell A 2021, "How susceptible are you to the sunk cost fallacy?", Harv Bus Rev, July 15.

Rosenbaum PR & Rubin DB 1983, "The central role of the propensity score in observational studies for causal effects", Biometrika, vol. 70, pp. 41–55.

Rubin DB 2005, "Causal inference using potential outcomes", J Am Stat Assoc, vol. 100, pp. 322-331.

Sonnenschein C & Soto AM 1999, The society of cells: Cancer and control of cell proliferation, New York: Springer-Verlag.

Sonnenschein C & Soto AM 2016, "Carcinogenesis explained within the context of a theory of organisms", Prog Biophys Mol Biol, vol. 122, pp.70-76.

Soto AM & Sonnenschein C 2004, "The somatic mutation theory of cancer: growing problems with the paradigm?", Bioessays, vol. 26, pp.1097-1107.

Soto AM & Sonnenschein C 2011, "The tissue organization field theory of cancer: a testable replacement for the somatic mutation theory", Bioessays, vol. 33, pp. 332-340.

Sowell T (interviewee) 2018, "Thomas Sowell on the myths of economic inequality", Uncommon Knowledge with Peter Robinson, Dec 3 [video]. Retrieved from: https://www.youtube.com/watch?v=mS5WYp5xmvI

Steyerberg EW & Harrell FE Jr. 2016, "Prediction models need appropriate internal, internal-external, and external validation", J Clin Epidemiol, vol. 69, pp. 245-247.

Teller A 2016, "The head of 'X' explains how to make audacity the path of least resistance", Wired, April 15. Retrieved from: https://www.wired.com/2016/04/the-head-of-x-explains-how-to-make-audacity-the-path-of-least-resistance/?utm_source=WIR_REG_GATE#.wzc0zaixn

The ATBC Cancer Prevention Study Group 1994, "The alpha-tocopherol, beta-carotene lung cancer prevention study: design, methods, participant characteristics, and compliance. The ATBC Cancer Prevention Study Group", Ann Epidemiol, vol. 4, pp. 1-10.

Thomas DS, Fourkala EO, Apostolidou S, *et al.* 2015, "Evaluation of serum CEA, CYFRA21-1 and CA125 for the early detection of colorectal cancer using longitudinal preclinical samples", Br J Cancer, vol. 113, pp. 268-274.

Thomson DM, Krupey J, Freedman SO & Gold P 1969, "The radioimmunoassay of circulating carcinoembryonic antigen of the human digestive system", Proc. Natl Acad. Sci. USA, vol. 64, pp.161-167.

Vickers AJ, van Calster B & Steyerberg, EW 2019, "A simple, step-by-step guide to interpreting decision curve analysis", Diagn Progn Res, vol. 3: 18.