The History and Development of Quantitative Structure-Activity Relationships (QSARs)

John C. Dearden, School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, Liverpool, UK

ABSTRACT

It is widely accepted that modern QSAR began in the early 1960s. However, as long ago as 1816 scientists were making predictions about physical and chemical properties. The first investigations into the correlation of biological activities with physicochemical properties such as molecular weight and aqueous solubility began in 1841, almost 60 years before the important work of Overton and Meyer linking aquatic toxicity to lipid-water partitioning. Throughout the 20th century QSAR progressed, though there were many lean years. In 1962 came the seminal work of Corwin Hansch and co-workers, which stimulated a huge interest in the prediction of biological activities. Initially that interest lay largely within medicinal chemistry and drug design, but in the 1970s and 1980s, with increasing ecotoxicological concerns, QSAR modelling of environmental toxicities began to grow, especially once regulatory authorities became involved. Since then QSAR has continued to expand, with over 1400 publications annually from 2011 onwards.

KEYWORDS

1816, Corwin Hansch, Crum Brown and Fraser, Descriptors, Environmental Sciences, Newer Approaches, Pharmacology, Statistics

INTRODUCTION: WHAT IS A QSAR?

Humans are inherently inquisitive. Even small children persistently ask “Why?”. So it is no surprise that for many years scientists have asked why some substances have a beneficial effect on the body, whilst others are toxic, and why some are more beneficial, or more toxic, than are others. That led Crum Brown and Fraser (1868-1869) to postulate that “there can be no reasonable doubt but that a relation exists between the physiologic action of a substance (Φ) and its chemical composition and constitution (C)” Hence Φ = f/C. They did not go on to suggest what functions of composition and constitution might be important. Nevertheless, their equation is a valid generic quantitative structure-activity relationship (QSAR). They also pointed out that “to discover f we produce a known change on the constitution by which it becomes C + ΔC, and examine the corresponding change of physiological action which has become Φ + ΔΦ. We thus obtain the relation between ΔC and ΔΦ, and by sufficiently varying C and ΔC, we may hope to get at all events an approximate solution of the problem”. That was a remarkably prescient statement, for it is exactly how QSAR modelling is performed (Kubinyi, 1993).

Hence, in Crum Brown and Fraser’s terminology, a QSAR equation would be:

Φ = c_1C_1 + c_2C_2 + c_3C_3 + … c_nC_n

(1)
where \( C_n \) represents one constitutional (structural) property, and \( c_n \) is its coefficient.

A QSAR is now defined as a mathematical relationship linking chemical structure and pharmacological activity or other property in a quantitative manner for a series of compounds. It should be noted that a correlation between a physico-chemical property (such as aqueous solubility) and some function(s) of chemical composition and constitution is usually called a quantitative structure-property relationship (QSPR), although strictly the term QSPR covers both structure-biological activity and structure-physico-chemical property relationships, as the title of this journal indicates.

As will be seen, the growth of QSAR since the 1970s has been huge. Consequently, it has not been possible in this brief review to provide more than a glimpse into modern aspects and achievements of QSAR. For those wishing to know more of those approaches, the recent excellent review by Cherkasov et al. (2014) is recommended.

**EARLY APPROACHES**

It is significant that for centuries there has been recognition that quantitation is an essential part of science. Leonardo da Vinci (1452-1519) commented that “there is no certainty in sciences where one of the mathematical sciences cannot be applied” (da Vinci). According to Galileo (1564-1642) “to study a given phenomenon, it was necessary to measure quantities, identify regularities, and obtain relationships representing mathematical descriptions as simply as possible” (Ponte, 1992). Gay-Lussac (1778-1850) optimistically said that “we are perhaps not far from the epoch when we will be able to submit to calculation the majority of chemical phenomena” (Gay-Lussac, 1809). Charles Babbage (1791-1871), the father of computing, could almost be thought to have had QSAR in mind when he stated that “all of chemistry…would become a branch of mathematical analysis which, like astronomy, taking its constants from observation, would enable us to predict the character of any new compound” (Babbage, 1837). A dissenting voice was that of Auguste Comte (1798-1857), who wrote that “every attempt to employ mathematical methods in the study of chemical questions must be considered profoundly irrational and contrary to the spirit of chemistry. If mathematical analysis should ever hold a prominent place in chemistry – an aberration which is happily almost impossible – it would occasion a rapid and widespread degeneration of that science” (Liang, Kvalheim, & Manne, 1993).

Sir William Thomson (1824-1907), later Lord Kelvin, was very forthright in his views on the importance of mathematics in science; “I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the state of Science, whatever the matter may be” (Thomson, 1884).

**THE BEGINNINGS OF CORRELATION (1816-1900)**

**Physicochemical Properties: The Periodic Table**

So far as is known, the earliest work on property prediction involved the elements, or as many of them as were known in the early nineteenth century. The first to publish on this topic appears to have been Döbereiner, who in a letter to Goethe in 1816 (Kauffman, 1999) mentioned what was to evolve into his Dreieheit (rule of triads); “The mineral coelestine [strontium sulphate] shows remarkable relationships: its specific weight is the mean of that of [calcium sulphate] and [barium sulphate], namely \((2.95 + 4.47)/2 = 3.71\)”. The rule of triads is in effect a read-across technique (Van Leeuwen, Schultz, Henry, Diderich, & Veith, 2009) whereby a property value of a chemical is predicted from known values of that property from one or more similar chemicals. Nonetheless, it is a valid predictive approach, and thus can be included as an historical use of QSPR.
Study of Pyrimidine-4-carboxamide Derivatives as HIV-1 Integrase Inhibitors Using QSAR and DFT Calculations
www.igi-global.com/article/study-of-pyrimidine-4-carboxamide-derivatives-as-hiv-1-integrase-inhibitors-using-qsar-and-dft-calculations/191197?camid=4v1a

Multi-Objective Modeling of Herbicidal Activity from an Environmentally Friendly Perspective
www.igi-global.com/article/multi-objective-modeling-of-herbicidal-activity-from-an-environmentally-friendly-perspective/181616?camid=4v1a

QSAR and Structure-Based Docking Studies of Aryl Pyrido[2,3-d]pyrimidin-7(8H)-ones: An Attempt to Anticancer Drug Design
www.igi-global.com/article/qsar-and-structure-based-docking-studies-of-aryl-pyrido23-dpyrimdin-78h-ones/191193?camid=4v1a