frontiers | Frontiers in Complex Systems

TYPE Original Research
PUBLISHED 16 June 2023
DOI 10.3389/fcpxs.2023.1208066

OPEN ACCESS

EDITED BY
Marta C. Gonzalez,
University of California, United States

REVIEWED BY
Soumyajyoti Biswas,
SRM University, India
Rudolf Marcel Füchslin,
Zurich University of Applied Sciences, Switzerland

*CORRESPONDENCE
R. Craig Herndon,
herndonrc@upmc.edu

RECEIVED 18 April 2023
ACCEPTED 05 June 2023
PUBLISHED 16 June 2023

CITATION
Herndon RC (2023), Information guided adaptation of complex biological systems.
Front. Complex Syst. 1:1208066.
doi: 10.3389/fcpxs.2023.1208066

Information guided adaptation of complex biological systems

R. Craig Herndon*

Hillman Cancer Center, Radiation Oncology, University of Pittsburgh Medical Center, Williamsport, PA, United States

Introduction: Functional information transmission through a complex biological system is introduced as a method for biological response (bioresponse) adaptation using qualified biological marker (biomarker) data.

Methods: This information guided adaptation methodology traverses the series of complex connections, defined by disparate bioresponse and biomarkers data sets, by placing both data sets on the common platiorm defined by information. The absolute uncertainty associated with these data sets can be defined in terms of decimal digits of information. This relationship between the data's absolute uncertainty and its information entropy is used to decompose information entropy into functional and relative uncertainty components, where the functional component quantifies the function or meaning of a data set in units of information.

Results: Application of functional information to adapt patient medical treatments using the information values for the bioresponse model and the biomarker model are outlined in detail and presented tabularly.

Discussion: Functional information provides a mathematical connection between a bioresponse model and a biomarker model by quantifying both data sets in the units of information and thereby providing the means to implement precision therapeutic plans by quantitatively adapting patient treatments using their biomarker measurements.

KEYWORDS

information theory, complexity, bioinformatics, functional information, complex adaptive systems, computational biology

1 Introduction

The human body is the quintessential complex system with its multiple interacting heterogeneous components that display emergent macro-level behavior due to multiple non-linear interconnections that are typically unknown and inseparable, and therefore cannot be modeled (Johnson, 2006; Sheard and Mostashari, 2009; Earl and Nicholson, 2021; San Miguel, 2023). Medical diagnosis of the bodies' biological response (bioresponse) to disease and treatment based on biological markers (biomarkers) is an effective method for understanding, treating, and managing complex disease processes (National Cancer Institute). Understanding the complexity inherent in cancer progression presents a profound challenge to researchers and is being actively investigated through the application of biomarkers in the field known as precision medicine (Chatterjee and Zetter, 2005). Precision medicine is a methodology that uses patient information to understand and manage the health of the complex biological system that is the human body.

Precision medicine uses patient information about their genes or proteins to prevent, diagnose, or treat their disease (National Cancer Institute). The prominence of precision medicine has increased in recent years because of advances in basic research areas including

molecular biology, genomics, and bioinformatics (Collins and Varmus, 2015). These advances have enabled more precise targeting of subgroups of disease with new therapies (Ashley, 2016). The intrinsic link between precision medicine and biomarkers is illustrated by the Precision Medicine Initiative (PMI) launched by the US National Institutes of Health (NIH) with the goal of improving health care by combining clinical data and multi-omic biomarker measurements on a large scale (Vargas and Harris, 2016; Olivier et al., 2019). Multi-omics combines multiple omic data sets, (e.g., genomics, epigenomics, transcriptomics, proteomics, metagenomics), during data analysis to determine the mechanism of a biological process (Kim et al., 2012; Urbanowicz et al., 2018; Krassowski et al., 2020; Momeni et al., 2020; Peng et al., 2020; Shi et al., 2021). This paper extends the application of information to the PMI goal of incorporating a patients' bioresponse and biomarker data into the planning, execution, and adaptation of their treatments during their prescribed course of therapy. This is accomplished by placing both the bioresponse and the biomarker data in the same analytic data space through the conversion of both data sets into functional information, i.e., all data are analyzed in the units of information. Thus, biomarker data can be fed directly back into bioresponse models to adapt patient therapies based on the patient's current biological profile.

Information-theoretic analysis began 75 years ago when Claude Shannon presented a mathematical measure of the amount of information that can be transmitted over a potentially noisy communication channel (Shannon, 1948). This mathematical theory optimizes the transmission of message data by quantifying how many bits of information are contained in the message, however, it does not determine if the data have meaning. This is termed the semantic problem because two messages can have identical information even though one message has meaning, and the other is nonsense (Weaver). Applications of information entropy in biomedical informatics include semantics, genetic selection, feature selection, and biomarkers (Li et al., 2004; Cohen and Widdows, 2009; Saha et al., 2009; Liu et al., 2010; Bolón-Canedo et al., 2014; Chen et al., 2017; Sato and Akimoto, 2017; Bakal et al., 2018; Jadon, 2020). Recently, through the application of measurement theory, the definition of information has been extended beyond the data streams associated with communications to include functional data associated with curves and graphs thus quantifying data meaning in terms of information bits (Herndon, 2022). The absolute uncertainty associated with measured data, in the form $y = f(x)$, has been defined in terms of real-valued decimal digits (dits) of information (Herndon, 2017; 2021). This relationship is used to decompose information into functional and relative uncertainty components, where the functional component quantifies the function or meaning of a process data set, $y = f(x)$, in units of information (Herndon, 2022).

Functional information, I_f, offers an unprecedented way to quantify system response when using surrogate measurements to assess system status. A biological system response, or bioresponse, model characterizes system response to an event, but if it is not measurable then a surrogate biological system measurement, or biomarker, is used because it is an objective indicator of the bioresponse to a therapeutic intervention (Strimbu and Tavel, 2010; Califf, 2018). The biomarker contains no parametric data from the system model and cannot be used to calculate the bioresponse status. Functional information removes this constraint between bioresponses and biomarkers by quantifying a value for the number of information-carrying decimal digits (dits) in each data set, thus providing a common space for analysis. Bioresponse will be adapted using functional information, I_S, acquired from the bioresponse function, S, and functional information, I_V, acquired using a biomarker function, V.

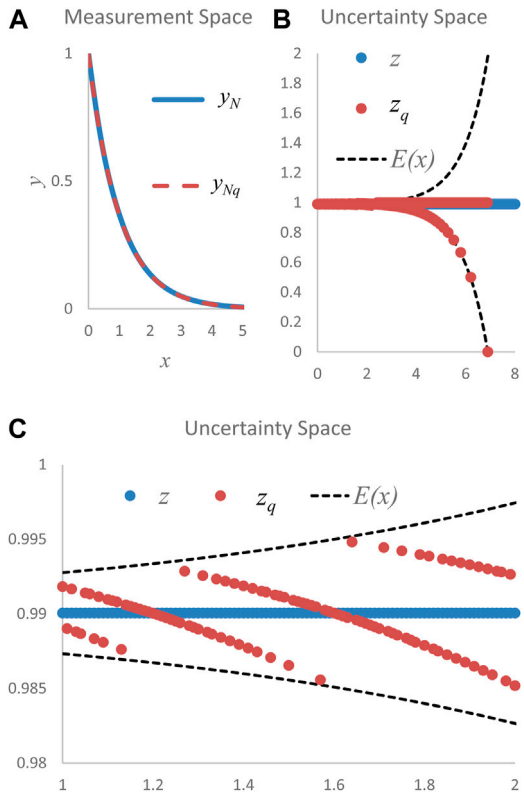

FIGURE 1
(A). Measurement space normalized exponential data, $y_N = f(x)_N = \exp(-x)_N$, and quantized normalized exponential data, $y_{N,q}$, are graphically indistinguishable **(B)**. Differences between uncertainty space data, z (straight line of blue dots) and z_q (spread of orange dots), calculated using normalized data, y_N, and normalized quantized (q = 3) measurement space data, $y_{N,q}$, are apparent. The envelope function, $E(x)$, provides an upper and lower boundary for the quantized data, z_q **(C)**. A zoomed closeup of uncertainty space shows all quantized points, z_q (orange data points), are contained within the envelope and illustrates that the spread of the envelope defines the spread of the data, z_q. All z_q data points are single valued for each input value of x. Note that the ideal, or unquantized uncertainty data, z (blue data points), have no spread.

Applications of this method include individualizing patient drug dosages based on their molecular profiles. This could be as routine as adapting an antibiotic drug prescription to fight a specific infection based on a patient's blood test or as specialized as prescribing immunotherapy uniquely to each patient's disease. Functional information is used to create the mathematical machinery that permits ongoing adaptative therapy during a patient's treatment course.

2 Materials and methods

2.1 Functional information

Separation of Shannon information entropy, H, into functional (meaningful) and relative uncertainty (noisy) information

components has been mathematically developed and presented (Herndon, 2022). Functional information forms the mathematical basis for the adaptation methodology which is the focus of the paper. Therefore, functional information is reviewed first, then the information guided adaptation methodology is presented, following by applications to demonstrate the range of the method's applicability and provide a detailed outline of the process.

Quantization uncertainty associated with the data measurement process was linked to the information associated with the function that maps to the measured data (Herndon, 2017; 2021). Information associated with measured data is defined in terms of the data's real-valued decimal digit (dits) or binary digit (bits) accuracy and is based on the absolute uncertainty of the data set. Real-valued digit accuracy, q, is the number of dits, or bits, required to describe each data point, i.e., the real-valued digit accuracy, q, is equivalent to the Shannon information, h, associated with each data point ($q = h$) (Herndon, 2017; 2021). A relationship between data uncertainty and information is obtained and used to decompose the information entropy, H, into functional, I_f, and relative uncertainty, I_u, information components (Herndon, 2022). The functional information component, I_f, quantifies mathematical model data, $f(x)$, into a value in the units of information that describes the entire function.

The exponential, y (Figure 1A), illustrates a data set acquired with m equiprobable measurements. The exponential function, ubiquitous in nature, graphically illustrates the useful transition from measurement data to uncertainty data that is detailed below (Figures 1B, C). Information entropy, H (Eq. 1), is the average of a data set's information components, h, when determined from these m equiprobable measurements (Cover and Thomas, 2006; Stone, 2015; Çengel, 2021).

$$H = 1/m\sum_{j=0}^{m} h_j \tag{1}$$

Data constituting this exponential function have information components, h, that are defined in terms of the data's real-valued digit accuracy and are determined in uncertainty space (Herndon, 2017; 2021). Measured data in the form, $y = f(x)$, are transformed from measurement space data, y, into uncertainty space data, z, by the γ operation in Eq. 2 (Figure 1B). Data, y, are normalized (subscript N), y_N, before transformation to uncertainty space data. Normalizing refers to preprocessing steps necessary to ensure all data values range from zero to one while maintaining the form of the measured data. After normalization, the data are quantized (subscript q) to the whole number that corresponds to the integer-valued digit accuracy of the measured data (Herndon, 2017). Uncertainty space data, z_q, determined by the γ operation, are defined by the sequence, z_q (Eq. 2), where u_r is the data's relative uncertainty (Figure 1B). Unquantized uncertainty space data, z, is ideal and therefore is not quantized and produces no data spread (z plotted as blue dots in Figures 1B, C).

$$z_q = \gamma y_{N,q} = f(x + \Delta x)_{N,q} \Big/ f(x)_{N,q} \pm u_r \tag{2}$$

There is also an envelope function, $E(x)$ (Eq. 3), in uncertainty space (Figures 1B, C) that is characterized by the normalized function, $y_N = f(x)_N$, and the information component, h, which was defined by the real-valued digit accuracy, q (Herndon, 2017; 2021).

$$E(x) = f(x + \Delta x)_N \Big/ f(x)_N \pm 10^{-h} \Big/ f(x)_N \tag{3}$$

An equivalency exists in uncertainty space (Eq. 4) between the relative uncertainty, u_r, associated with the spread of the data, z_q, and the spread of the envelope function, $E(x)$ (Herndon, 2017; 2021; 2022).

$$u_r = \frac{10^{-h}}{f(x)_N} \tag{4}$$

The equivalency of Eq. 4 is illustrated graphically in Figure 1B where the spread of the uncertainty space data, z_q, is bounded by the spread of the envelope function, $E(x)$, which is further demonstrated in the closeup of Figure 1C; Eq. 4 is used to decompose the information entropy, H, into separate components that quantify data function ($f(x)_N$) information, I_f, and the data relative uncertainty (u_r) information, I_u (Eq. 5) (Herndon, 2022).

$$H = -\frac{1}{m}\sum_{j=0}^{m}\left(\log\left(f\left(x_j\right)_N\right)\right) + \log\left((u_r)_j\right) = I_f + I_u \tag{5}$$

Information, I_f, quantifies a value for the number of functional information-carrying real-valued decimal digits (dits) in a data set by determining the average of all individual data point functional information values, i (Eq. 6) (Herndon, 2022). This deterministic value, I_f, assigns meaning to the model function, $f(x)$, in terms of information, thus any function can be identified in terms of information. Functional information becomes the common platform for comparison of data sets that were previously considered disparate.

$$I_f = -\frac{1}{m}\sum_{j=0}^{m}\log\left(f\left(x_j\right)_N\right) = \frac{1}{m}\sum_{j=0}^{m} i_j \tag{6}$$

Next, the information formulation for adapting bioresponses using biomarkers, based on Eq. 6, is presented.

2.2 Information guided adaptation methodology

An event, x, acts upon a complex biological system and causes a response or effect, S, that can be as modeled by $S = f_1(x)$. Typically, this biological system response model or bioresponse, $S = f_1(x)$, cannot be directly measured, so a valid surrogate biological system marker model or biomarker, V, is measured to obtain an indirect quantification of system response (Strimbu and Tavel, 2010; Califf, 2018). If the mathematical chain of intermediate functions, C_i (Eq. 7), connecting the biological system response, S, to the biomarker response, V, were known, then a cause-effect feedback mechanism based upon calculated and measured biomarker units would be available for adaptive therapies. However, this functional chain is unknown which means the input, C_{n-1}, to the biomarker, V, is unknown, therefore biomarker data is measured against an appropriate independent variable, y (Eq. 8), and modeled by function g.

$$x \xrightarrow{f_1} S \xrightarrow{f_2} C_2 \cdots \xrightarrow{f_{n-1}} C_{n-1} \xrightarrow{f_n} V \tag{7}$$

$$y \xrightarrow{g} V \tag{8}$$

The bioresponse and biomarker models, $S = f(x)$ and $V = g(y)$, must be used to analyze the system since the chain of models

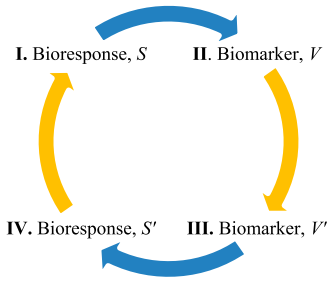

FIGURE 2
Information cycle, **I-IV**, used to adapt bioresponse data to biomarker-specified changes. Functional information can cross the knowledge gap represented by the blue arrows and permit changing the bioresponse model, *S*, using the adapted biomarker data, *V′*.

represented by $f_2 \rightarrow f_n$ in Eq. 7) are not available in complex biological systems. The formula used to relate bioresponse and biomarker information, I_S and I_V, is determined from the informational relationship between a function's current state, f, and its adapted state, f', defined as $I_{S'} = k_S I_S$ and $I_{V'} = k_V I_V$ (illustrated by the yellow arrows in Figure 2). The linear adaptation formula, (Eq. 9), determined from the bioresponse/biomarker ratio, $I_{S'}/I_{V'}$, is based on these relationships, different criteria will result in more complicated adaptation formulas. The values of the adaptation constants, k_i, specify the amount of change defined by the investigator's adaptive goal for the biomarker. This fractional change in information, k_V, between current and adapted biomarker information is equated to the fractional change, k_S, between current and adapted bioresponse information because they are describing the same biological system, therefore, $k_S = k_V$. Eq. 9 expects that the bioresponse and biomarker models do not change, other than parameters, between a function's current state, f, and its adapted state, f'. The information criterion $I_{V'} = kI_V$ is based on the measured biomarker data and is primary because it used by the investigator to set the value for the adaptation constant, k, assigned to the bioresponse relationship, $I_{S'} = kI_S$. The information criterion $I_{V'} = kI_V$ is applicable to a range of models including polynomials, power, and exponential functional groups (see Section 3).

$$I_{S'} = I_S \frac{I_{V'}}{I_V} \tag{9}$$

Functional information can be used to adapt treatments using measured outcomes because information is transmitted through the complex connections forming the bioresponse/biomarker system (illustrated by the blue arrows in Figure 2). Biomarkers must be investigator qualified as legitimate surrogates of the bioresponse to minimize uncertainties when modeling information flow through a biological system (Eq. 9). Next, the adapted bioresponse information (Eq. 9) will be used to determine the bioresponse input, x, that is necessary to adapt the bioresponse to the desired goal, i.e., a linear relationship, based on information, is used to solve a previously intractable nonlinear problem. A functional bioresponse-biomarker chain, based on process (7), will be synthesized to demonstrate that the mathematical foundation provided by functional information (Eq. 6) can be successfully applied to bioresponse/biomarker models (Eq. 9).

3 Results and discussion

3.1 Adaptation process outline

Exponential and power/polynomial examples are presented to demonstrate the range of functional groups applicable to the adaptation formulation developed in Section 2.2. For example, the adaptation formula (Eq. 9) robustly applies across exponential bioresponse models regardless of parametric changes to the models. Robustness is quantified by determining whether the information criterion $I_{V'} = kI_V$ (Section 2.2) is maintained when model parameters change.

The implementation process is outlined to show how functional information is transmitted through the complex connections of a bioresponse/biomarker system and can be used to adapt bioresponses using biomarker measurements. Adaptation of the bioresponse using biomarker information will be demonstrated using different sets of models based on the exponential, power, and polynomial functions. Exponential functions are used primarily because of their natural ubiquity in the biological sciences and power/polynomial functions are selected to demonstrate the flexibility of this methodology. These example applications demonstrate that information transferred through a complex biological system can be leveraged for adaptation purposes (Eq. 9).

The process of information flow from bioresponse, *S*, to biomarker, *V*, (Eq. 7), will be synthesized to demonstrate the utility of applied functional information. Information flow described in Eq. 9 can be visualized in the information cycle in Figure 2, where information is the mechanism that creates the bioresponse-biomarker data connection. Application of this information cycle begins with synthesizing a composition of functions that connects step I and II. Input changes in the bioresponse model, *S*, will automatically propagate to changes in the biomarker model, *V*. Next, the biomarker data of II will be modeled independently and then changed to affect the desired response (III). At step III the adapted biomarker data, V', does not equal the biomarker data of step II, $V' \neq V$. Information determined from steps I-III are then used to determine the adapted bioresponse information, $I_{S'}$ (Eq. 9), in step IV. This information, $I_{S'}$, is then used to determine the adapted input to the next iteration of the cycle (step I). If the adaptive information formula (Eq. 9) is valid then the next iteration of biomarker data, *V*, in step II will equal the previous adapted biomarker data of step III, i.e., $V = V'$ verifying the method. Equivalently, if functional information adaptation is viable, the next cycle of biomarker functional information (II) will equal the previous adapted biomarker functional information (III), i.e., $I_V|_{cycle\,2} = I_{V'}|_{cycle\,1}$.

Step I and II of Figure 2 and the processes connecting them are modeled by the composition of functions (Eq. 10) that automatically calculate the current biomarker, *V*, from the current bioresponse, *S*. This synthetic process is a simplified version of Eq. 7, where the bioresponse is modeled as $S = S_0 \exp(-(a_1 + a_2 x^2))$, the intermediate data is condensed to $C = C_0 \ln(b_1 S)$, and the biomarker is $V = V_0 \exp(-(c_1 + c_2 x)$ (Figure 3). Selection of an input parameters, a_1 and a_2, determine the bioresponse, *S* (Figure 3A), which cascades through the nonlinear functions (Eq. 10) to automatically determine the biomarker, *V* (Figures 3B, C).

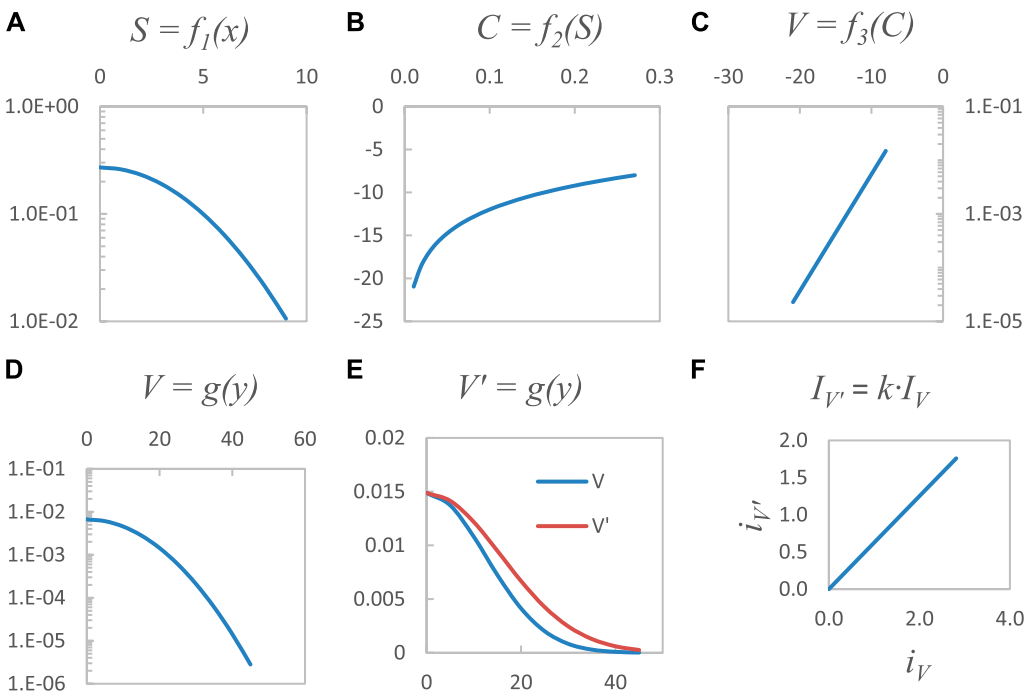

FIGURE 3
The functional composition $x \xrightarrow{f_1} S \xrightarrow{f_2} C \xrightarrow{f_3} V$ **(plots A–C)** connecting **Steps I** and **II** (Figure 2) is a synthesized process used to validate the methodology. Any change to the input, *x*, in the bioresponse model, *S*, results in immediate change in the biomarker model, *V* **(Plot C)**. Plot **(D)** simulates the clinical process of modeling the biomarker data from Plot **(C)** against a logically chosen independent variable, *y*, because the functional connections between Plot **(A)** and Plot **(C)** are unknown. The biomarker model, *V*, is plotted against the adapted model, *V'*, in plot **(E)** and the linear relationship between their information, $I_{V'} = kI_V$, is shown in Plot **(F)**. This linear relationship is maintained even when model parameters, a_i, b_i, and c_i, change.

$$x \xrightarrow{f_1} S \xrightarrow{f_2} C \xrightarrow{f_3} V \tag{10}$$

This biomarker data, *V*, is then independently modeled following Eq. 8, where $V = V_0 \exp(-d_1 y)$ (Figure 3D). Biomarker data, *V*, is changed to V' in step III using Eq. 8; (Figure 3E). The adapted biomarker functions, *V* and V', in Figure 3E, which are modeled using quadratic exponentials, maintain the linear criterion, $I_{V'} = kI_V$, shown in Figure 3F. This linear relationship, $I_{V'} = kI_V$, is robust, i.e., linearity is maintained even when model parameters, a_i, b_i, and c_i, change. The basis for determining the adapted input for the next cycle is the adapted bioresponse functional information, $I_{S'}$, calculated using the information adaptation formula (Eq. 9). Demonstration of the processes, graphically illustrated in Figure 3, are detailed next in a step-by-step format.

3.2 Adaptation process details

Application of the information adaptation Formula 9 presented in the following steps uses exponential data for the bioresponse and biomarker models. The models and the example parameters are listed in the steps below and in Table 1.

1) The bioresponse model is $S = S_0 \exp(-(a_1 + a_2 x^2))$, where $S_0 = 2$, $a_1 = 2$, and $a_2 = 0.04$. Normalize the data and determine the bioresponse information, I_S, using Eq. 6. Details are listed in Table 1A, where calculated bioresponse data are generated given $x = \{0, 1, 2, \ldots, 9\}$ (this corresponds to **Step I** of Figure 2). The parameter-of-interest is a_2 because it controls the bioresponse information content. The adaptation formula (Eq. 9) will be used to adapt a_2 in step 5 below.
2) Biomarker data, *V*, is automatically determined by the function composition $x \xrightarrow{f_1} S \xrightarrow{f_2} C \xrightarrow{f_3} V$ (10). Each step is graphed in Figures 3A–C. Calculated data in Table 1B for $C = C_0 \ln(b_1 S)$ and $V = V_0 \exp(-(c_1 + c_2 x)$ are determined using $C_0 = 4.0$, $b_1 = 0.5$, $V_0 = 6.0$, $c_1 = 2.0$, and $c_2 = -0.5$ (Table 1A). Functional information for *V* is calculated using Eq. 6. (**Step II** of Figure 2).
   a) Biomarker data, *V*, is modeled in linear increments of the independent variable, *y*, simulating the laboratory environment where the function $V = f_3(C)$ is unknown. The investigator now has biomarker data in terms of the function $V = g(y)$, where the independent variable $y = \{0, 5, 10, \ldots, 45\}$ (Figure 3D). The biomarker data in terms of function *g* are listed in Table 1B and the information, I_V, is identical to I_V in Table 1A because it is dependent on *x* from Eq. 10 not *y* from Eq. 8.
3) Current biomarker data, *V*, is adapted to the investigator's achievable goal, V' using the model $V' = V_0' \exp(-d_1 y^2)$, where $V_0' = 0.01487$, and $d_1 = 0.0032$ (Table 1B). Adaptation is accomplished by changing d_1 to $d_1' = 0.002$ to produce the desired change in the biomarker model, $V' = V_0 \exp(-d_1' y)$ (Figure 3E). This adaptation could also be accomplished by choosing a value for *k* that produces the desired fractional change in information, $I_{V'} = k_V I_V$. **Step III** (Figure 2).
   a) The parametric change, $d_1 = 0.0032$ to $d_1' = 0.002$, results in a change in biomarker information from $I_V = 0.9902$ dits to $I_{V'} = 0.6189$ dits.

TABLE 1 Exponential function application data.

A. Bioresponse information, I_S (Figure 2 Step I) and biomarker information, I_V (Step II). Data derivations are described in Section 3.2 steps 1-2.

x	S	S_N	C	V	V_N	i_S	i_V	I_S	I_V
0	0.2707	1.0000	−8.000	1.487E-2	1.0000	0.0000	0.0000	0.4951	0.9902
1	0.2601	0.9608	−8.160	1.372E-2	0.9231	0.0174	0.0347		
2	0.2307	0.8521	−8.640	1.080E-2	0.7261	0.0695	0.1390		
3	0.1888	0.6977	−9.440	7.239E-3	0.4868	0.1563	0.3127		
4	0.1427	0.5273	−10.560	4.135E-3	0.2780	0.2779	0.5559		
5	0.0996	0.3679	−12.000	2.012E-3	0.1353	0.4343	0.8686		
6	0.0641	0.2369	−13.760	8.348E-4	0.0561	0.6254	1.2508		
7	0.0381	0.1409	−15.840	2.950E-4	0.0198	0.8512	1.7024		
8	0.0209	0.0773	−18.240	8.887E-5	0.0060	1.1118	2.2236		
9	0.0106	0.0392	−20.960	2.281E-5	0.0015	1.4071	2.8142		

B. Adapted biomarker information, $I_{V'}$ (Step III). Data derivations are described in Section 3.2 step 3.

y	V	V'	V_N	V'_N	i_V	$i_{V'}$	I_V	$I_{V'}$
0	0.01483	0.0149	1.0000	1.0000	0.0000	0.0000	0.9902	0.6189
5	0.01373	0.0141	0.9231	0.9512	0.0347	0.0217		
10	0.01080	0.0122	0.7261	0.8187	0.1390	0.0869		
15	0.00724	0.0095	0.4868	0.6376	0.3127	0.1954		
20	0.00414	0.0067	0.2780	0.4493	0.5559	0.3474		
25	0.00201	0.0043	0.1353	0.2865	0.8686	0.5429		
30	0.00084	0.0025	0.0561	0.1653	1.2508	0.7817		
35	0.00030	0.0013	0.0198	0.0863	1.7024	1.0640		
40	8.89E-5	0.0006	0.0060	0.0408	2.2236	1.3897		
45	2.28E-5	0.0003	0.0015	0.0174	2.8142	1.7589		

C. Adapted bioresponse information, $I_{S'}$ (Step IV cycle 1), becomes the new I_S ($I_S = I_{S'}$) in the next information cycle (Step I, cycle 2) listed below. The next iteration of biomarker information, I_V =0.6189, is determined (Step II cycle 2) and compared to the target biomarker information, $I_{V'}$ = 0.6189 from Table 2B (Section 3.2 steps 4-5).

x	S	S_N	C	V	V_N	i_S	i_V	I_S	I_V
0	0.2707	1.000	−8.00	1.487E-2	1.0000	0.0000	0.0000	0.3094	0.6189
1	0.2640	0.9753	−8.10	1.415E-2	0.9512	0.0109	0.0217		
2	0.2449	0.9048	−8.40	1.218E-2	0.8187	0.0434	0.0869		
3	0.2161	0.7985	−8.90	9.483E-3	0.6376	0.0977	0.1954		
4	0.1814	0.6703	−9.60	6.683E-3	0.4493	0.1737	0.3474		
5	0.1449	0.5353	−10.50	4.261E-3	0.2865	0.2714	0.5429		
6	0.1100	0.4066	−11.60	2.458E-3	0.1653	0.3909	0.7817		
7	0.0795	0.2938	−12.90	1.283E-3	0.0863	0.5320	1.0640		
8	0.0546	0.2019	−14.40	6.062E-4	0.0408	0.6949	1.3897		
9	0.0357	0.1320	−16.10	2.591E-4	0.0174	0.8794	1.7589		

TABLE 2 Power/Polynomial function application data.

A. Bioresponse information, I_S (Figure 2 Step I) and biomarker information, I_V (Step II). Refer to Section 3.2 step 6 for details.									
x	S	S_N	C	V	V_N	i_S	i_V	I_S	I_V
0	5.000	0.1066	11.18	7.34	0.3350	0.9722	0.4750	0.4566	0.2413
1	6.000	0.1279	14.70	7.83	0.3573	0.8930	0.4469		
2	8.249	0.1759	23.69	8.87	0.4045	0.7548	0.3931		
3	11.47	0.2446	38.86	10.23	0.4668	0.6115	0.3308		
4	15.56	0.3317	61.36	11.83	0.5398	0.4793	0.2678		
5	20.43	0.4355	92.32	13.61	0.6208	0.3610	0.2071		
6	26.03	0.5550	132.8	15.52	0.7082	0.2557	0.1499		
7	32.33	0.6894	183.8	17.56	0.8010	0.1615	0.0964		
8	39.30	0.8379	246.3	19.70	0.8984	0.0768	0.0465		
9	46.90	1.0000	321.2	21.92	1.0000	0.0000	0.0000		
B. Adapted biomarker information, $I_{V'}$ (Step III). Refer to Section 3.2 step 6 for details.									
y	V	V'	V_N	V'_N	i_V	$i_{V'}$	I_V	$I_{V'}$	
0	7.344	0.335	7.078	0.2890	0.4750	0.5391	0.2413	0.2688	
5	7.834	0.357	8.013	0.3271	0.4469	0.4853			
10	8.867	0.405	9.198	0.3755	0.3931	0.4254			
15	10.23	0.467	10.63	0.4341	0.3308	0.3624			
20	11.83	0.540	12.32	0.5029	0.2678	0.2985			
25	13.61	0.621	14.25	0.5819	0.2071	0.2351			
30	15.52	0.708	16.44	0.6711	0.1499	0.1732			
35	17.56	0.801	18.87	0.7705	0.0964	0.1132			
40	19.70	0.898	21.56	0.8802	0.0465	0.0554			
45	21.92	1.000	24.49	1.0000	0.0000	0.0000			
C. Adapted bioresponse information, $I_{S'}$ (Step IV cycle 1), becomes the new I_S ($I_S = I_{S'}$) in the next information cycle (Step I, cycle 2) listed below. The next iteration of biomarker information, I_V =0.2674, is determined (Step II cycle 2) and compared to the target biomarker information, $I_{V'}$ = 0.2688 from Table 2B. Refer to Section 3.2 step 6 for details.									
x	S	S_N	C	V	V_N	i_S	i_V	I_S	I_V
0	4.000	0.0778	8.000	6.828	0.2942	1.1092	0.5313	0.5085	0.2674
1	5.000	0.0972	11.18	7.344	0.3164	1.0123	0.4997		
2	7.379	0.1434	20.04	8.477	0.3653	0.8433	0.4374		
3	10.88	0.2117	35.92	9.994	0.4306	0.6744	0.3659		
4	15.42	0.2997	60.53	11.780	0.5076	0.5233	0.2945		
5	20.89	0.4062	95.51	13.77	0.5935	0.3913	0.2266		
6	27.27	0.5302	142.4	15.93	0.6866	0.2756	0.1633		
7	34.51	0.6709	202.7	18.24	0.7859	0.1734	0.1047		
8	42.57	0.8276	277.8	20.67	0.8905	0.0822	0.0504		
9	51.44	1.0000	368.9	23.21	1.0000	0.0000	0.0000		

i) Linearity between the biomarker and adapted biomarker information, $I_{V'} = kI_V$, is maintained regardless of changes to the input coefficients, a_i, b_i, and c_i (Figure 3F).

4) The information adaptation formula (Eq. 9) is used to determine the adapted information value, $I_{S'}$, for the adapted state of the bioresponse model, S'. **Step IV** (Figure 2).
   a) The adapted bioresponse information value, $I_{S'} = 0.3094$ dits (Table 1C), is determined using the initial bioresponse information, I_S, the initial biomarker information, I_V, and the adapted biomarker information, $I_{V'}$, from Tables 1A, B ($I_S = 0.4951$ dits, $I_V = 0.9902$ dits, and $I_{V'} = 0.6189$ dits) and Eq. 9.
5) Finally, the parameter-of-interest, $a_2 = 0.04$, is iteratively adjusted to $a_2' = 0.025$ until the next cycle of bioresponse information, I_S, is equal to the target value, $I_{S'} = 0.3094$ dits. The cycle begins again at **Step I** by setting $a_2 = a_2'$, and calculating the next iteration of biomarker data, V. If this current iteration of biomarker data equals the adapted biomarker data, V', of **Step IV** (from #4 above), then the adaptation process is finished. An alternate method to determine if adaptation is finished occurs when there is an equivalency between the new cycle of biomarker information, $I_V = 0.6189$ dits (Table 1C), and the adapted biomarker information, $I_{V'}$ (Table 1B).
   a) Adaptation of a complex biological system based on functional information has been demonstrated when $I_V|_{cycle\,2} = I_{V'}|_{cycle\,1}$.
6) The application process outlined above in steps 1–5 is applied to a power function model, where the bioresponse model, $S = a_1 + x^{a_2}$ ($a_1 = 5$, $a_2 = 1.7$), $C = S^{b_1}$ ($b_1 = 1.5$), $V = c_1 + C^{c_2}$ ($c_1 = 4$, $c_2 = 0.5$). The power function data resulting from following process steps 1-2) above are listed in Table 2A. The model, $V = g(y)$, used for adaptation is the polynomial $V = d_1 + d_2 y + d_3 y^2$ ($d_1 = 7.0776$, $d_2 = 0.162$, $d_3 = 0.0038$). The quadratic coefficient is changed to $d_3 = 0.005$ to simulate adaptation to the target goal, V', following the procedure in Section 3.2 step 3 (Table 2B). The polynomial model, $V = g(y)$ (Figure 3D), may not offer any mechanistic explanatory value, but it still serves the purpose of transferring information and maintaining the criterion $I_{V'} = k{\cdot}I_V$ and therefore the validity of the adaptation formulation (Eq. 9). Following Section 3.2 steps 4-5 the adapted bioresponse information, $I_{S'}$, is determined and the power function variable is iteratively modified from $a_2 = 1.7$ to $a_{2'} = 1.7565$ so that $I_{S'}|_{\text{cycle }1} = I_S|_{\text{cycle }2} = 0.5085$ (Table 2C). The biomarker information, $I_V|_{\text{cycle }2} = 0.2674$ dits, that corresponds to this bioresponse input, $I_S|_{\text{cycle }2} = 0.5085$ dits, is 0.5% difference from the $I_V|_{\text{cycle }1} = 0.2688$ dits. This uncertainty (0.5%) is a measure of the robustness of the application using this power function model. In contrast, the linearity criterion for the exponential function had 0% uncertainty ($I_V|_{cycle\,2} = I_{V'}|_{cycle\,1} = 0.6189$ dits, Table 1), however, the power function adaptation process is considered complete because $I_V|_{cycle\,2} = I_{V'}|_{cycle\,1}$ is within an acceptable uncertainty range (0.5%).

Functional information transmission through the complexities of the bioresponse/biomarker relationship provides a mechanism for precision medicine. Application to precision medicine has been demonstrated using the reproducible processes outlined and detailed in Section 3.1, Section 3.2. The adaptation formulation (Eq. 9) is applicable to a range of bioresponse, $S = f(x)$, and biomarker, $V = g(y)$, models including exponential-based and power/polynomial-based models. Adaptation is performed at the fundamental level of information where the effect the prescribed therapy has on the patient is analyzed in the same units (dits) as the patient's measured outcomes. This information foundation permits quantitative adaptions of individual patient treatment responses based on changes defined by the patients' biomarker measurements. Functional information assists current efforts in the bioinformatics community to classify and verify candidate biomarkers (Shin et al., 2008; Wang et al., 2009; Ganchev et al., 2011; Asgari et al., 2018; Liu and Gao, 2018; Yaghoobi et al., 2021; Ding et al., 2023; Nazari and Zinati, 2023). Information guided precision medicine also provides a mechanism for solving inherently complex biological problems like those encountered in the active research areas of targeted and adaptive oncology therapies (Sawyers, 2004; Aggarwal, 2010; Baudino, 2015; Padma, 2015; Wang et al., 2007; Bhullar et al., 2018; Yang et al., 1038; Zhong et al., 2021; Gatenby et al., 2009; Pazarentzos and Bivona, 2015; Neri et al., 2007). Oncology applications are emphasized because it is the focus of the NIH PMI, however, information-guided analysis is broadly applicable to any functional analysis of signal data.

4 Conclusion

Functional information's mathematical origin in metrology has been presented and applied to complex biological systems by adapting bioresponses using biomarker information. Information-guided analysis is directly applicable to the bioresponse/biomarker relationships found in immunotherapy, chemotherapy, and radiotherapy and is extensible within the health sciences to include bioinformatics and computational biology. In addition, it has application to systems engineering where it provides a feedback mechanism for complex control systems. Functional information also impacts basic sciences because it defines the wave function in units of information.

Functional information guided adaptation of complex biological systems has been introduced as a promising analytic tool for investigators in the growing field of precision medicine. Information guided analysis of complex oncological processes will augment ongoing research efforts by combining information from the collective knowledge of oncology researchers in support of the near-term focus on cancer research established by the NIH Precision Medicine Initiative (PMI). Generally, functional information has potential application in any research area where signal function analysis is relevant.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcpxs.2023.1208066/full#supplementary-material

References

Asgari, Y., Khosravi, P., Zabihinpour, Z., and Habibi, M. (2018). Exploring candidate biomarkers for lung and prostate cancers using gene expression and flux variability analysis. *Integr. Biol.* 10 (2), 113–120. doi:10.1039/c7ib00135e

Ashley, E. A. (2016). Towards precision medicine. *Nat. Rev. Genet.* 17 (9), 507–522. doi:10.1038/nrg.2016.86

Aggarwal, S., "Targeted cancer therapies," *Nat. Rev. Drug Discov.*, vol. 9, no. 427, 428. doi:10.1038/nrd31866, 2010.

Bakal, G., Talari, P., Kakani, E. V., and Kavuluru, R. (2018). Exploiting semantic patterns over biomedical knowledge graphs for predicting treatment and causative relations. *J. Biomed. Inf.* 82, 189–199. doi:10.1016/j.jbi.2018.05.003

Baudino, T. A. (2015). Targeted cancer therapy: The next generation of cancer treatment. *Curr. Drug Discov. Technol.* 12 (1), 3–20. doi:10.2174/1570163812666150602144310

Bhullar, K. S., Lagarón, N. O., McGowan, E. M., Parmar, I., Jha, A., Hubbard, B. P., et al. (2018). Kinase-targeted cancer therapies: Progress, challenges and future directions. *Mol. Cancer* 17, 48–20. doi:10.1186/s12943-018-0804-2

Bolón-Canedo, V., Sánchez-Marono, N., Alonso-Betanzos, A., Benítez, J. M., and Herrera, F. (2014). A review of microarray datasets and applied feature selection methods. *Inf. Sci.* 282, 111–135. doi:10.1016/j.ins.2014.05.042

Califf, R. M. (2018). Biomarker definitions and their applications. *Exp. Biol. Med.* 243 (3), 213–221. doi:10.1177/1535370217750088

Çengel, Y. A., "On entropy, information, and conservation of information," *Entropy*, vol. 23, no. 6, 779, 2021.

Chatterjee, S. K., and Zetter, B. R. (2005). Cancer biomarkers: Knowing the present and predicting the future. *Future Oncol.* 1 (1), 37–50. doi:10.1517/14796694.1.1.37

Chen, Y., Zhang, Z., Zheng, J., Ma, Y., and Xue, Y. (2017). Gene selection for tumor classification using neighborhood rough sets and entropy measures. *J. Biomed. Inf.* 67, 59–68. doi:10.1016/j.jbi.2017.02.007

Cohen, T., and Widdows, D. (2009). Empirical distributional semantics: Methods and biomedical applications. *J. Biomed. Inf.* 42 (2), 390–405. doi:10.1016/j.jbi.2009.02.002

Collins, F. S., and Varmus, H. (2015). A new initiative on precision medicine. *N. Engl. J. Med.* 372 (9), 793–795. doi:10.1056/nejmp1500523

Cover, T. M., and Thomas, J. A. (2006). *Elements of information theory*. 2nd. Hoboken, NJ, USA: John Wiley and Sons.

Ding, Y., Lei, X., Liao, B., and Wu, F. X. (2023). Biomarker identification via a factorization machine-based neural network with binary pairwise encoding. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* 20, 2136–2146. doi:10.1109/tcbb.2023.3235299

Earl, R., and Nicholson, J. (2021). *The concise oxford dictionary of mathematics*. Oxford, England: Oxford University Press.

Ganchev, P., Malehorn, D., Bigbee, W. L., and Gopalakrishnan, V. (2011). Transfer learning of classification rules for biomarker discovery and verification from molecular profiling studies. *J. Biomed. Inf.* 44, S17–S23. doi:10.1016/j.jbi.2011.04.009

Gatenby, R. A., Silva, A. S., Gillies, R. J., and Frieden, B. R. (2009). Adaptive therapy. *Cancer Res.* 69 (11), 4894–4903. doi:10.1158/0008-5472.can-08-3658

Herndon, R. C. (2022). Data entropy decomposition into functional and uncertainty components. *Measurement* 188, 110541.

Herndon, R. C. (2021). Determining signal entropy in uncertainty space. *Measurement* 178, 109336. doi:10.1016/j.measurement.2021.109336

Herndon, R. C. (2017). Measurement analysis in uncertainty space. *Measurement* 105, 106–113. doi:10.1016/j.measurement.2017.04.012

Jadon, Shruti. (2020). "A survey of loss functions for semantic segmentation," in Proceedings of the IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), Via del Mar, Chile, December 2020, 1–7.

Johnson, C. W. (2006). What are emergent properties and how do they affect the engineering of complex systems? *Reliab. Eng. Syst. Saf.* 91 (12), 1475–1481. doi:10.1016/j.ress.2006.01.008

Kim, D., Shin, H., Song, Y. S., and Kim, J. H. (2012). Synergistic effect of different levels of genomic data for cancer clinical outcome prediction. *J. Biomed. Inf.* 45 (6), 1191–1198. doi:10.1016/j.jbi.2012.07.008

Krassowski, M., Das, V., Sahu, S. K., and Misra, B. B. (2020). State of the field in multi-omics research: From computational needs to data mining and sharing. *Front. Genet.* 11, 610798. doi:10.3389/fgene.2020.610798

Li, Haifeng, Zhang, Keshu, and Jiang, Tao (2004). "Minimum entropy clustering and applications to gene expression analysis," in Proceedings. 2004 IEEE Computational Systems Bioinformatics Conference, Stanford, CA, USA, August 2004, 142–151.

Liu, H., Liu, L., and Zhang, H. (2010). Ensemble gene selection by grouping for microarray data classification. *J. Biomed. Inf.* 43 (1), 81–87. doi:10.1016/j.jbi.2009.08.010

Liu, Z. P., and Gao, R. (2018). Detecting pathway biomarkers of diabetic progression with differential entropy. *J. Biomed. Inf.* 82, 143–153. doi:10.1016/j.jbi.2018.05.006

Momeni, Z., Hassanzadeh, E., Saniee Abadeh, M., and Bellazzi, R. (2020). A survey on single and multi omics data mining methods in cancer data classification. *J. Biomed. Inf.* 107, 103466. doi:10.1016/j.jbi.2020.103466

National Cancer Institute Dictionary of cancer terms. https://www.cancer.gov/publications/dictionaries/cancer-terms (Accessed April 17, 2023).

Nazari, L., and Zinati, Z. (2023). Gene expression classification for biomarker identification in maize subjected to various biotic stresses. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* 20, 2170–2176. doi:10.1109/tcbb.2022.3233844

Neri, Ferrante, Toivanen, Jari, Cascella, G. L., and Ong, Yew-Soon (2007). An adaptive multimeme algorithm for designing HIV multidrug therapies. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* 4 (2), 264–278. doi:10.1109/tcbb.2007.070202

Olivier, M., Asmis, R., Hawkins, G. A., Howard, T. D., and Cox, L. A. (2019). The need for multi-omics biomarker signatures in precision medicine. *Int. J. Mol. Sci.* 20 (19), 4781. doi:10.3390/ijms20194781

Padma, V. V. (2015). An overview of targeted cancer therapy. *BioMedicine* 5, 1–6. doi:10.7603/s40681-015-0019-4

Pazarentzos, E., and Bivona, T. G. (2015). Adaptive stress signaling in targeted cancer therapy resistance. *Oncogene* 34 (45), 5599–5606. doi:10.1038/onc.2015.26

Peng, C., Zheng, Y., and Huang, D. (2020). Capsule network based modeling of multi-omics data for discovery of breast cancer-related genes. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* 17 (5), 1605–1612. doi:10.1109/tcbb.2019.2909905

Saha, S. K., Sarkar, S., and Mitra, P. (2009). Feature selection techniques for maximum entropy based biomedical named entity recognition. *J. Biomed. Inf.* 42 (5), 905–911. doi:10.1016/j.jbi.2008.12.012

San Miguel, M. (2023). Frontiers in complex systems. *Front. Complex Syst.* 1 (2). doi:10.3389/fcpxs.2022.1080801

Sato, K., and Akimoto, K. (2017). Expression levels of KMT2C and SLC20A1 identified by information-theoretical analysis are powerful prognostic biomarkers in estrogen receptor-positive breast cancer. *Clin. Breast Cancer* 17 (3), e135–e142. doi:10.1016/j.clbc.2016.11.005

Sawyers, C. (2004). Targeted cancer therapy. *Nature* 432 (7015), 294–297. doi:10.1038/nature03095

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27 (3), 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x

Sheard, S. A., and Mostashari, A. (2009). Principles of complex systems for systems engineering. *Syst. Eng.* 12 (4), 295–311. doi:10.1002/sys.20124

Shi, K., Lin, W., and Zhao, X. (2021). Identifying molecular biomarkers for diseases with machine learning based on integrative omics. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* 18 (6), 2514–2525. doi:10.1109/tcbb.2020.2986387

Shin, H., Sheu, B., Joseph, M., and Markey, M. K. (2008). Guilt-by-association feature selection: Identifying biomarkers from proteomic profiles. *J. Biomed. Inf.* 41 (1), 124–136. doi:10.1016/j.jbi.2007.04.003

Stone, J. V. (2015). *Information theory, A tutorial introduction.* Sheffield, UK: Sebtel Press.

Strimbu, K., and Tavel, J. A. (2010). What are biomarkers? *Curr. Opin. HIV AIDS* 5 (6), 463–466. doi:10.1097/coh.0b013e32833ed177

Urbanowicz, R. J., Olson, R. S., Schmitt, P., Meeker, M., and Moore, J. H. (2018). Benchmarking relief-based feature selection methods for bioinformatics data mining. *J. Biomed. Inf.* 85, 168–188. doi:10.1016/j.jbi.2018.07.015

Vargas, A. J., and Harris, C. C. (2016). Biomarker development in the precision medicine era: Lung cancer as a case study. *Nat. Rev. Cancer* 16 (8), 525–537. doi:10.1038/nrc.2016.56

Wang, H. Q., Wong, H. S., Zhu, H., and Yip, T. T. (2009). A neural network-based biomarker association information extraction approach for cancer classification. *J. Biomed. Inf.* 42 (4), 654–666. doi:10.1016/j.jbi.2008.12.010

Wang, M. D., Shin, D. M., Simons, J. W., and Nie, S. (2007). Nanotechnology for targeted cancer therapy. *Expert Rev. Anticancer Ther.* 7 (6), 833–837. doi:10.1586/14737140.7.6.833

Weaver, W. (1953). Recent contributions to the mathematical theory of communication. *ETC Rev. Gen. Semant.* 10, 261–281.

Yaghoobi, H., Babaei, E., Hussen, B. M., and Emami, A. (2021). Ebst: An evolutionary multi-objective optimization based tool for discovering potential biomarkers in ovarian cancer. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* 18 (6), 2384–2393. doi:10.1109/tcbb.2020.2993150

Yang, L., Shi, P., Zhao, G., Xu, J., Peng, W., Zhang, J., et al. Targeting cancer stem cell pathways for cancer therapy. *Signal Transduct. Target. Ther.* 5 (8). doi:10.1038/s41392-020-0110-5

Zhong, L., Li, Y., Xiong, L., Wang, W., Wu, M., Yuan, T., et al. (2021). Small molecules in targeted cancer therapy: Advances, challenges, and future perspectives. *Signal Transduct. Target. Ther.* 6 (1), 201. doi:10.1038/s41392-021-00572-w