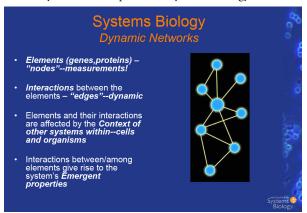
CURTIS A. BAGNE, PH.D. DATASPEAKS, INC.

International Symposium, Institute for Systems Biology, April 2008

1 Background

Dr. Leroy Hood has represented systems biology as shown:



DataSpeaks Interactions° Software embodies a computational algorithm that measures "interactions" or "edges" as interactions over time when the "elements" or "nodes" are action variables. Furthermore, DataSpeaks Interactions° measures coordination of action – the harmonious functioning of parts for effective results – as an "Emergent" system property.

Scientists and engineers have developed great tools to identify and measure parts of cells, organs, organisms, ecosystems and their environments. Now DataSpeaks Interactions® can help users describe and predict how such parts interact to form systems that work:

- **Function** internally (Demonstration 1),
- **Respond** to their environments including treatments (Demonstration 2), and
- Act as agents on their environments.

New measures of interaction over time will help users build systems science on more reductionistic approaches to basic and applied science. DataSpeaks Interactions® can help take us from "Genomes to Life." As such, DataSpeaks Interactions® deserves to be evaluated as a tool and potential killer app that can help expedite 21st Century science and P4 medicine.

2 Technical Overview

DataSpeaks Interactions® applies to **time series data** with at least two but preferably many more repeated "measurements" to measure the amount, strength and positive or negative direction of evidence for *interactions over time* between at least one action variable functioning as an independent variable and at least one action variable functioning as a dependent variable.

Action variables are variables that can change and fluctuate in level over time for individual systems.

Independent action variables include levels, concentrations and doses of "elements" that form "nodes." Endogenous independent action variables include levels of gene expression and concentrations of proteins, metabolites and other biological constituents. Exogenous or environmental independent action variables that sustain and can **perturb**

systems include levels of nutrients, allergens, pollutants and other stimuli as well as doses of drugs.

Most laboratory variables in medicine and measures of vital signs, symptoms, mental and motor performance, and quality of life can be investigated as dependent action variables. Although death rates can be investigated as action variables for whole populations, vital status and survival time are *not* action variables for individual patients. Electrophysiological measures are action variables.

Genome and genotype are categorical or classification variables – not action variables – for individual organisms. DataSpeaks Interactions® does not apply to classification variables. However, DataSpeaks Interactions® can help elucidate what genetic differences mean in terms of how different organisms work and how genetically similar organisms

(e.g., man, mouse) can work quite differently.

DataSpeaks Interactions® can help users transcend the limits of data snapshots, which include change scores, to the advantages of *multiple time series data* or *data movies*. Compared to data snapshots, *data movies* can account for time, time dependent system dynamics and can include orders of magnitude more information to understand how individual systems *work*. Furthermore, *data movies* might be the best way to investigate **coordination of action** as a time dependent "emergent" system property. Many data collection technologies already collect or are capable of collecting *data movies*.

Demonstrations 1 and 2 illustrate fundamental innovations in scientific methodology using some of the most rudimentary features and capabilities of DataSpeaks Interactions® Software.

CURTIS A. BAGNE, PH.D. DATASPEAKS, INC. International Symposium, Institute for Systems Biology, April 2008

3 Demonstration 1

Measuring Coordinated Biological Function as a Time Dependent Emergent System Property

The time series in **Figure 1** are for two "elements" or "nodes." These data were collected every five minutes for about 12 hours from hypophyseal portal blood from one ewe. Our task is to demonstrate the capabilities of DataSpeaks Interactions° and help validate this software by measuring a well known *interaction over time* between two hormones.

Figure 1 is a *data movie* with two pixels or time series variables and 143 repeated measurements (movie frames) for one individual. See the *coordination of action*. Now we will measure this coordination as interactions over time so that the coordinated action can be investigated more scientifically.

Luteinizing hormone (LH), ng/mL

1000

100

Gonadotropin releasing hormone (GnRH), pg/mL

Figure 1. Data for Demonstration 1

3.1 Developing "Data Movies," Distribution of Potential Interaction Scores and Results

DataSpeaks Interactions® was used to "develop" the data movie in Figure 1 by measuring both pairwise directional interactions (GnRH to LH, LH to GnRH) with a user specified scoring protocol that yielded a detailed but easy to summarize 4-dimensional array of 1,008 (6 x 6 x 7 x 4) standardized interaction scores that accounted for:

- Level of the independent variable (6 levels of dimensional resolution),
- Level of the dependent variable (6 levels of dimensional resolution),
- Any time delay of apparent effect of independent events on dependent events (7 levels, 0 thru 6), and
- Any time persistence of effects of independent events on dependent events (4 levels, 1 thru 4).

This array of 1,008 standardized interaction scores was summarized by selecting the most extreme positive or negative value, which is 76.028. On

30 of 32 repeated measurement times in Figure 1 when a specific type of independent GnRH event was present, a specific type of dependent LH event also was present. In addition, the same specific type of dependent LH event never was present on any of 111 times when the same specific type of independent GnRH event was not present.

400

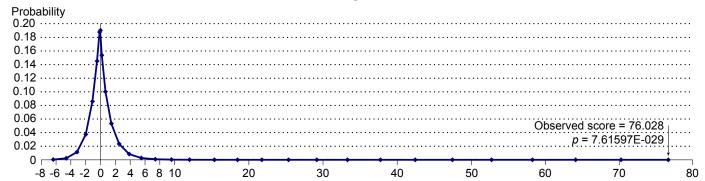
Time, Minutes

500

600

This summary interaction score (76.028) is one score from a distribution of potential scores, defined by the data in Figure 1 in combination with the scoring protocol. This distribution, shown in **Figure 2**, has a mean of zero and a standard deviation of 1. The summary score of 76.028 indicates substantial evidence for a GnRH to LH interaction. This confirms what most people are apt to think after looking at Figure 1. However, unlike subjective impressions, the software yields a reproducible quantitative result and applies to problems beyond unaided human capabilities.

Figure 2



CURTIS A. BAGNE, PH.D. DATASPEAKS, INC.

International Symposium, Institute for Systems Biology, April 2008

Figure 3 illustrates performance characteristics of two different and complementary types of measures that quantify:

- Amount and direction of evidence versus,
- Strength and direction of evidence for the interaction over time for the GnRH (independent variable) to LH (dependent variable) interaction.

These interaction scores were computed iteratively across the 143 repeated measurement times (715 minutes) that are shown in Figure 1.

Figure 4 summarizes the 4-dimensional arrays of 1,008 standardized interaction scores as functions of their four dimensions or **analysis parameters**. Summary scores are the most extreme positive or negative values in a row, a column or an overall array.

Parts A and B of **Figure 4** summarize the GnRH to LH interaction as functions of the levels of the interactants expressed as residuals from a linear regression line that was used to de-trend the time series before measuring the interactions over time. Taken together, Parts A and B show that at least moderate level LH events are most apt to be associated with high level GnRH events. Temporal asymmetries in Parts C and D suggest that high levels of GnRH are more predictive of high levels of LH than vice versa, especially at time delay level 1. More generally, DataSpeaks Interactions® can be used to explore the **temporal criterion of causal relationships.**

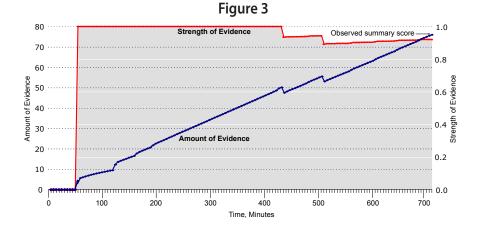
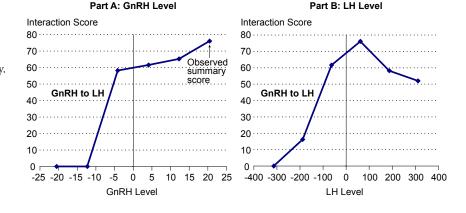
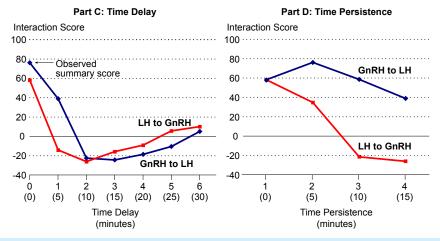


Figure 4





3.2 Prospectus – Demonstration 1

Measurement of *coordination of action* as interactions over time as illustrated by **Demonstration 1**:

- Can be extended to hundreds or thousands of time series variables for highthroughput investigations as with functional brain imaging, which is nondestructive;
- Enables both global and detailed assessments of coordinated system function;
- · Has high potential value for:
 - Elucidating and visualizing normal coordination,
 - Objective diagnoses of

disorders of coordinated function,

- Elucidating mechanisms of action,
- Functional biomarkers and phenotypes.

CURTIS A. BAGNE, PH.D. DATASPEAKS, INC.

International Symposium, Institute for Systems Biology, April 2008

4 Demonstration 2

A Placebo Controlled Single Group Randomized Controlled Trial

(Clue - Randomize doses to time periods.)

"FDA's evaluation methods have remained largely unchanged over the last half century." (FDA Science Board, FDA Science and Mission at Risk, November 2007)

Randomized controlled trials (RCT) are the best way to assess causality. However, current

first generation RCT designs, which provide much of the **evidentiary basis** of the current public health (not personalized or P4) approach to medicine, predate:

- Modern computers,
- The genomic revolution and systems biology,
- High-throughput and Internet enabled measurement technologies and
- Major companies such as Microsoft and Google, "the two leading candidates for Web supremacy," that now seek to provide software-based services to improve health ("Dr. Google and Dr. Microsoft: Two Giants Have Plans to Change Health Care," New York Times, 8/14/07).

4.1 Confounding in Current First Generation RCT Designs

Not only have RCT designs "remained largely unchanged over the last half century," but such designs also confound.

Current first generation RCT designs that:

- Randomize patients to different groups to assess causality (1) confound individuality with measurement error and (2) confound true responders to active treatment with responders on active treatment that would have responded to placebo,
- Randomize patients to different groups defined by different doses of a particular type of treatment in essentially the same way that patients are randomized to groups defined by different types of treatment (3) confound dose with type of treatment as in current first generation placebo controlled RCT designs and

 Perform statistical tests on health variables or changes in health variables (4) confound treatment effects and how treatment effects are valued.

We can end these four types of confounding to help:

- Revitalize the pharmaceutical industry by:
 - Translating advancements in life sciences into better products and health,
 - Improving productivity and research efficiency and
 - Reducing safety problems;
 - Improve research investment returns;
 - Save lives, avoid harm and avoid legal liability:
 - Avoid loss of products as with Vioxx (Merck, pain) and Bextra (Pfizer, pain) and loss of potential products as with torcetrapib (Pfizer, cholesterol

management);

- Improve the economic competitiveness of states and nations;
- Improve public credibility and support for science.

Table 1 illustrates a **second generation RCT design** that does **not** commit these four types of confounding. Such new designs can be applied to evaluate drug treatments for **chronic health problems of function** such as chronic pain, hypertension, lipid disorders, diabetes, endocrine disorders and neuropsychiatric disorders.

Second generation RCT designs are possible when both independent and dependent variables can be investigated as action variables – variables that vary and fluctuate in level over time for individuals. Table 1 uses mock data for treatment and health.

Table 1

| | Week | | | | | | | | | | | | | | | | | | | | |
|------------------|-------------------|-------------------|------------------|--------|----|----|------|--------|----|----|------|--------|----|--------|----|-------------|--|------------------------|---------------------|------------------|---------------------------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | ĺ | | | | |
| | | Pa | ir 1 | | | Pa | ir 2 | | | Pa | ir 3 | | | Pair 4 | | | 1 | | ارموررا | | Overall |
| Patient | Period | | | Period | | | | Period | | | | Period | | | | Interaction | | Benefit/ Harm | | Benefit/ Harm | |
| Variable | 1 | | - 2 | 2 | 1 | | | 2 | • | 1 | - 1 | 2 | • | l | 1 | 2 | score | Direction ¹ | Weight ² | Score | |
| Patient 1 | | | | | | | | | | | | | | | | | | | | | |
| Dose | 20 | 20 | 40 | 40 | 0 | 0 | 40 | 40 | 40 | 40 | 20 | 20 | 80 | 80 | 40 | 40 | | | | | |
| DBP ³ | 96 | 98 | 85 | 81 | 91 | 96 | 84 | 87 | 80 | 78 | 93 | 98 | 82 | 77 | 81 | 78 | -8.92 | - | 8.92 | 4 | |
| ED ⁴ | 3 | 2 | 2 | 3 | 2 | 1 | 2 | 1 | 3 | 2 | 1 | 2 | 3 | 2 | 2 | 2 | 0.74 | - | -0.74 | 2 | |
| Energy | 4 | 3 | 4 | 5 | 4 | 3 | 4 | 3 | 2 | 4 | 3 | 3 | 4 | 4 | 2 | 3 | 1.46 | + | 1.46 | 2 | 4.64 |
| Patient 2 | | | | | | | | | | | | | | | | | | | | | |
| Dose | 0 | 0 | 20 | 20 | 40 | 40 | 0 | 0 | 20 | 20 | 80 | 80 | 80 | 80 | 40 | 40 | | | | | |
| DBP | 98 | 91 | 89 | 88 | 89 | 84 | 96 | 98 | 89 | 93 | 86 | 80 | 76 | 75 | 80 | 92 | -8.56 | - | 8.56 | 4 | |
| ED | 2 | 1 | 2 | 2 | 3 | 2 | 1 | 2 | 2 | 3 | 2 | 4 | 4 | 3 | 4 | 2 | 5.93 | + | 5.93 | 1 | |
| Energy | 2 | 3 | 2 | 2 | 3 | 1 | 2 | 3 | 2 | 2 | 3 | 3 | 4 | 3 | 2 | 3 | 3.05 | + | 3.05 | 2 | 6.61 |
| Patient 3 | | | | | | | | | | | | | | | | | | | | | |
| Dose | 40 | 40 | 20 | 20 | 0 | 0 | 80 | 80 | 20 | 20 | 40 | 40 | 20 | 20 | 80 | 80 | | | | | |
| DBP | 76 | 79 | 74 | 80 | 88 | 90 | 78 | 68 | 75 | 77 | 81 | 78 | 73 | 79 | 76 | 82 | -7.81 | - | 7.81 | 4 | |
| ED | 1 | 3 | 2 | 2 | 1 | 1 | 2 | 4 | 2 | 2 | 2 | 1 | 3 | 2 | 3 | 5 | 5.04 | - | -5.04 | 1 | |
| Energy | 4 | 3 | 2 | 2 | 3 | 4 | 3 | 2 | 5 | 3 | 4 | 3 | 4 | 4 | 1 | 2 | -2.67 | + | -2.67 | 1 | 3.92 |
| | are us Diastol | ed to s ic Blo | pecify od Pre | the re | | | | | | | | | | | | | or each particular or each particular | | Grou | p Average | 5.06 t=6.29 p=.0242 |

CURTIS A. BAGNE, PH.D. DATASPEAKS, INC.

International Symposium, Institute for Systems Biology, April 2008

Each of the nine health variable specific benefit/harm scores in Table 1 is one score from a distribution of potential scores that has a mean of zero and a standard deviation of 1. **Figure 5** shows all nine of these distributions. Standardization helps make it feasible to compare and combine observed benefit/harm scores from different distributions.

Figure 5: Distributions of potential benefit/harm scores for Demonstration 2

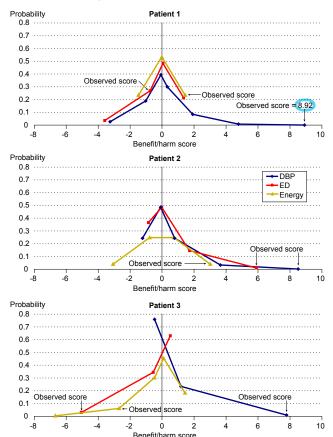


Figure 6 shows health variable specific and overall benefit/harm as a function of dose for each patient. The **optimal minimal doses** for patients 1, 2 and 3 were 40, 80 and 20 respectively. Part D shows the group averages. Current first generation RCT designs do *not* identify optimal minimal doses for individual patients or for groups.

Figure 7 shows how evidence for benefit/harm can be monitored over time (repeated measurement times) for each individual patient and for the group on average. Evidence from such monitoring can help assure that harm does not exceed benefit for any patient. Individual patients in RCTs that evaluate drugs for chronic health problems can be treated as patients – not research subjects. **Research and clinical practice can be more integrated**.

Second generation RCT designs as illustrated by Demonstration 2 are more scientific, informative, efficient and ethical than current first generation RCT designs.

Figure 6: Benefit/harm as a function of drug dose for Demonstration 2

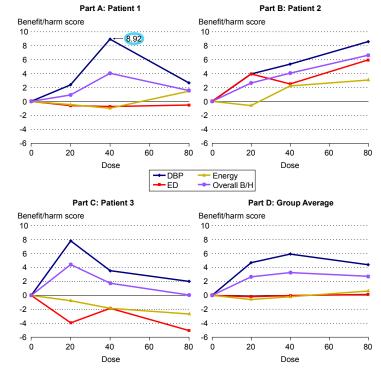
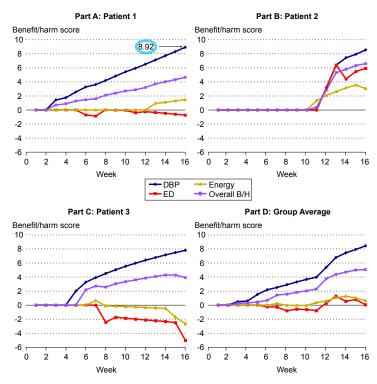


Figure 7: Monitoring evidence over time for apparent benefit/harm for Demonstration 2



CURTIS A. BAGNE, PH.D. DATASPEAKS, INC.

International Symposium, Institute for Systems Biology, April 2008

4.2 DataSpeaks Interactions® enables fundamental innovations in RCT design:

- Causality can be assessed by randomizing various doses to different periods of time and measuring *interactions over time* for **each** individual **before** statistical analyses and hypothesis testing for group data.
 - Dose can be investigated as an independent action variable for each patient. Patients were not classified into groups defined by doses.
 - No patient was randomized to placebo only, only an insufficient or ineffective dose, or only an excessive or harmful dose. This is an important ethical issue.
 - True response to active treatment can be differentiated from placebo response for each individual patient.
- Optimal minimal doses can be ascertained for each individual as illustrated by Figure 6.
- Commensurability. Beneficial treatment effects (effectiveness) and harmful treatment effects (safety) with respect to various health action variables can be measured and used to compute one overall benefit/harm score for each patient using explicit directionalities (positive or negative) and weights (indicating clinical and personal significance) as shown in Table 1.

- Safety and effectiveness can be profiled across health variables and balanced scientifically.
- One hypothesis and one statistical test in one RCT can evaluate benefit/harm with respect to many health action variables.
- Treatment evaluations can be comprehensive of effects on all health action variables that can be measured repeatedly.
- Although hypotheses do need to be tested with predetermined directionalities and weights, ex post facto exploratory investigations can be conducted with different directionalities and weights.
- One way to transform interaction scores into benefit/harm scores is to investigate how and to what extent interaction scores predict survival times and death rates.
- The null hypothesis of no overall benefit/ harm was rejected in the positive or beneficial direction using a two-tailed single group t-test on the mean. No statistical test was performed on health variables or changes

- in health variables.
- Time delays and persistencies in treatment response can be accounted for as indicated by Figure 4, Part C and Part D for Demonstration 1.
- Data from 16 repeated measurements for each patient were used to **overcome measurement error, increase statistical power** and help **achieve statistical significance** with only three patients (see Table 1, page 4).
- All of the above offer promise to help investigators identify genetic and other predictors of differential responses to treatments and to make **genotyping** more valuable.
- Potential new gold standards for clinical research and practice can be integrated.
 The market for DataSpeaks Interactions® in clinical practice is larger than its market as a research tool.

Comparative head-to-head RCTs can be conducted by randomizing patients to different types of treatment before conducting a single group RCT for each type of treatment.

5 DataSpeaks Interactions® Can Help Expedite 21st Century Systems Science

Measurement of **coordinated action** as a time dependent "emergent" system property (**Demonstration 1**) and measurement of the **benefit/harm** of treatments (**Demonstration 2**) makes such concepts amenable to scientific investigations. New measures have a history of advancing science and commerce – sometimes in surprising and dramatic ways (Archimedes, density and his Eureka moment).

Measurement of interactions with DataSpeaks Interactions® can help mathematicians and other investigators understand and model real systems that are:

• Complex. Table 1 illustrates how DataSpeaks Interactions® can help users investigate one aspect of complexity – one type of treatment, an independent variable, affecting many dependent health variables. This one-many

type problem was addressed by dimension reduction. DataSpeaks Interactions® also includes tools for addressing **many-one** and **many-many** type problems. Some of these tools involve applying Boolean operators to digital time series (see Table 2) for two or more independent or dependent variables.

- Adaptive. Iterative processing as illustrated in Figures 3 and 7 can be used to investigate forms of adaptation such as learning, habituation, sensitization, tolerance and neuroplasticity. Adaptation would be indicated by substantial changes in slope or changes in sign.
- Stochastic. Figures 2 and 5 illustrate how DataSpeaks' computational algorithm is fundamentally stochastic. Table 1 differentiates DataSpeaks' algorithm from classical statistics. Assessment of causality

- and measurement of benefit/harm for each individual came *before* statistical analysis of group results. DataSpeaks' algorithm, biostatistics and quantum mechanics all use probability. However, it often is useful to distinguish such different disciplines.
- Nonlinear. Figure 4, parts A and B, and Figure 6 seem to illustrate a form of nonlinearity. Nonlinearity can be investigated with more variables to investigate synergy and antagonism.
- **Hierarchical**. DataSpeaks Interactions® can help foster interdisciplinary investigations by measuring *interactions over time* within and between time series at different levels of investigation such as cell, organ and organ system and at biological, psychological and social levels.

CURTIS A. BAGNE, PH.D. DATASPEAKS, INC.

International Symposium, Institute for Systems Biology, April 2008

6 DataSpeaks Interactions® Can Help Expedite P4 Medicine

Drs. Hood and Zerhouni champion P4 medicine. DataSpeaks Interactions® can help expedite medicine that is:

- **Predictive**. Measures of interaction over time that describe past behavior can help predict future behavior. Furthermore, DataSpeaks' algorithm measures time dependent emergent system properties. Such measures can be used to identify genetic and other predictors, develop better classifications of chronic health problems and increase the value of genotyping.
- Personalized. Two keys to personalized medicine are assessing causality for individual patients (Table 1) and identifying optimal minimal drug doses (Figure 6). Personalized medicine that improves individual health will improve group average or public health.
- Preventive or Preemptive. Twenty-first

Century science that improves scientific understanding can be used to help prevent health problems and control health care costs.

• Participatory. DataSpeaks Interactions® implemented on the Internet would help empower people by themselves to achieve better health for themselves and their loved ones based on the collection and processing of data about themselves as illustrated in Demonstration 2. In addition, this could help create Web-based cumulative repositories of anonymous data and information (as in HealthVault.com?) that could be of great value to decision makers such as patients, clinicians, pharmaceutical companies, health care providers, regulators and payers.

Is it acceptable to prescribe and administer drugs without monitoring drug effects as

illustrated in **Figure** 7 when there is significant uncertainty about safety and cost-effectiveness?

DataSpeaks Interactions® has the potential to make **electronic medical records more attractive** to clinicians and patients as they seek better outcomes. DataSpeaks Interactions® has potential to help create **second generation evidence-based medicine**. New software technology can help **fix health care**.

DataSpeaks Interactions® can raise standards of evidence for much of systems science and medicine.

We are at the cusp of a revolution toward 21st Century systems science and P4 medicine. DataSpeaks Interactions® can help advance this revolution.

7 The Digital Promise of DataSpeaks Interactions® Software

Digitalization of biology and medicine is a "revolution that will transform medicine even more than digitalization transformed information technology and communications"

—Dr. Leroy Hood

DataSpeaks Interactions® embodies a computational algorithm that essentially is digital in a manner that goes beyond the way much other software, such as software for classical statistics, has been made to be digital for digital computers.

This innovative digital nature of DataSpeaks Interactions® is illustrated by showing part of how the DataSpeaks' algorithm was used to compute the benefit/harm score (8.92) with respect to diastolic blood pressure (DBP) for Patient 1 in Table 1. One crucial step is to transform the analog time series for dose, an exogenous independent action variable, and the analog time series for DBP, a dependent action variable, into sets of digital time series as shown in Table 2. There are various ways to transform analog series (a time series with more than two levels) into a set of digital series, often without necessary loss of information.

Table 2
Digitalization of analog time series for Dose and DBP for Patient 1

| | | | | | | | | | We | ek | | | | | | | | | | | |
|------|------|---|----|----|----|----|-------|-------|---------------------|-------|--------|-------|----|----|----|---|----|--|--|--|--|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | | | | |
| | | | | | | Or | ne An | alog | Time | e Ser | ies fo | or Do | se | | | | | | | | |
| | | 20 | 20 | 40 | 40 | 0 | 0 | 40 | 40 | 40 | 40 | 20 | 20 | 80 | 80 | 40 | 40 | | | | |
| D | | Set of Three Digital Time Series for Dose | | | | | | | | | | | | | | | | | | | |
| Dose | ≥ 20 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | | |
| | ≥ 40 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | | | | |
| | = 80 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | | | | |
| | | | | | | Oı | ne Ar | nalog | Time Series for DBP | | | | | | | | | | | | |
| | | 96 | 98 | 85 | 81 | 91 | 96 | 84 | 87 | 80 | 78 | 93 | 98 | 82 | 77 | 81 | 78 | | | | |
| | | Set of Eleven Digital Time Series for DBP | | | | | | | | | | | | | | | | | | | |
| | ≥ 78 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | | | | |
| | ≥ 80 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 | 0 | | | | |
| | ≥81 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | | | | |
| DBP | ≥ 82 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | | | | |
| DBP | ≥ 84 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | | | | |
| | ≥ 85 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | | | | |
| | ≥ 87 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | | | | |
| | ≥ 91 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | | | | |
| | ≥ 93 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | | | | |
| | ≥ 96 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 40 4 1 1 0 81 7 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 | 0 | | | | |
| | ≥ 98 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | | | | |

CURTIS A. BAGNE, PH.D. DATASPEAKS, INC.

International Symposium, Institute for Systems Biology, April 2008

Cross-classification of independent and dependent events: Another crucial step in DataSpeaks' algorithm is to cross-classify each digital series for the independent variable, dose, with each digital series for the dependent variable, DBP. The cell labels for this cross-classification are shown in **Table 3**.

Count the number of repeated measurement times for each cell. Interaction scores and benefit/harm scores are computed from the resulting cell frequencies – **a, b, c** and **d**. Computational details are published and in Patent 6,317,700 (see Patents Issued

Table 4 shows the cross-classification portion of DataSpeak's algorithm that was used to obtain the summary score of 8.92. DBP was lower on all 10 weeks when dose was 40 or more than for any of 6 weeks when dose was less than 40.

(The score with a value of 8.92) summarizes a 3-dimensional array with 66 (3 x 11 x 2) scores – the 3 levels of dose and 11 levels of DPB level shown in Table 2 and 2 levels of persistency. Table 4 shows only 11 of these 66 standardized benefit/harm scores.)

to Bagne). These four cell frequencies can help account for much complexity in how systems *work* in a more digital science much as A, T, C and G of the genetic code underlies much complexity in how living systems *work*. DataSpeaks' computational algorithm is basically simple.

Both spatial order or sequence (snapshots and maps for statics) and temporal order or sequence (movies for dynamics of interaction) are important. DataSpeaks Interactions® finds and measures patterns of *interaction over time* in temporal order.

Table 3 Digital Independent Events Present, 1 Absent, 0 Digital Present, 1 a=0 b=6 Dependent Events Absent, 0 Digital Independent Events Present, 1 Absent, 0 1,1 0,1 0,0 c=10 b=6 1,0 0,0 c=10 d=0 c+d=10 a+c=10 b+d=6

The distributions of potential scores shown in **Figures 2 and 5** show all scores that are possible given the marginal frequencies of the corresponding observed 2 x 2 tables.

Table 4: Digital pattern finding for multiple time series data.

| | Week | | | | | | | | | | | | | 2 x 2 Table | | | Benefit/ | | | | | |
|------|-------|---|---|---|-----|---|---|-----|---|-----|----|-------|---|-------------|-------|-----|----------|---------|---|--------|---|-------|
| | | | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 12 | | 13 | 13 14 | | 16 | Cell C | | Counts | | Harm |
| Dose | ≥ 40 | 0 | 0 | 1 | - 1 | 0 | 0 | - 1 | 1 | - 1 | 1 | 0 | 0 | - 1 | 1 | - 1 | - 1 | a b c d | | | | Score |
| | ≥ 78 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 9 | 6 | 1 | 0 | .77 |
| | ≥ 80 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 7 | 6 | 3 | 0 | 1.36 |
| | ≥ 81 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 6 | 6 | 4 | 0 | 1.86 |
| | ≥ 82 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 4 | 6 | 6 | 0 | 3.24 |
| | ≥ 84 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 3 | 6 | 7 | 0 | 4.15 |
| DBP | ≥ 85 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 2 | 6 | 8 | 0 | 5.35 |
| DDF | ≥ 87 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 6 | 9 | 0 | 6.87 |
| | ≥ 91 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | | | | | |
| | Cells | b | b | c | c | b | b | c | С | С | С | b | b | С | С | c | c | 0 | 6 | 10 | 0 | 8.92 |
| | ≥ 93 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 5 | 10 | 1 | 6.82 |
| | ≥ 96 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 4 | 10 | 2 | 5.06 |
| | = 98 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 2 | 10 | 4 | 2.37 |

8 Value Now

The value of DataSpeaks Interactions® is based on the value of fundamental principles of scientific inquiry such as use of:

- Data,
- Measurement,
- Randomized experimental control and
- Operational definitions.

Demonstrations 1 and 2 show how users of DataSpeaks Interactions® can apply these principles **more extensively** and with less confounding. It is not too early to try DataSpeaks Interactions® Software.

DataSpeaks Interactions® has potential to help establish a new high standard of empirical evidence for much of systems science and medicine. It can drive more extensive use of computing and the World Wide Web for understanding people and other types of systems. This, together with its potential to improve health and welfare, can help make DataSpeaks Interactions® a killer app.

Publication:

Bagne CA, Lewis RF. Evaluating the effects of drugs on behavior and quality of life: an alternative strategy for clinical trials. Journal of Consulting and Clinical Psychology 1992; 60:225-239. (http://www.dataspeaks.com/resources/ APA-JCCP-1992-Vol60-No2-P225-239.pdf)



Patents issued to Bagne:

6,317,700

Computational Method and System to Perform Empirical Induction

6,516,288

Method and System to Construct Action Coordination Profiles Please Contact:

Curtis A. Bagne, Ph.D. 2971 Vineyards Drive Troy, MI 48098 248 952-1968 bagne_curt@msn.com www.dataspeaks.com