Self-Tracking for Distinguishing Evidence-Based Protocols in Optimizing Human Performance and Treating Chronic Illness

Janet Chang
Study Manager, DIYgenomics
Community Chair, Ancestral Health Society
Mountain View, CA 94043
janetlaichang@gmail.com

Abstract

Self-tracking technologies used by healthy self-experimenters and chronic illness patients are relatively new but offer potential to accelerate the discovery of evidence-based protocols in the fields of human biology and medicine. Among both academic researchers and real-world practitioners in these fields there is an ever-present body of misinformation, leading to the proliferation of myth-based protocols in health-promoting lifestyles and treatment. This collection of four case studies spanning seven years' worth of observations in a self-experimenting endurance athlete and, later, chronically ill individual, aims to bring to attention the most common incorrect assumptions regarding: nutrition, athletic performance, sleep, and treatment of hypothyroidism. We hope that, with these insights about misleading scientific conclusions, artificial intelligence researchers and anyone interested in developing technological solutions for public health purposes, will explore ways to bridge the gap between academic research and real-world practice of optimizing human biology, and rid the misinformation on both sides.

Introduction

About Self Tracking

Self-tracking experimenters found in the small but growing Quantified Self (QS) community play an increasingly large role in public health research, as seen by QS community presence at industry events such as Health 2.0, TEDxMed, and Open Science Summit. The origin and growth of the QS movement may indicate the high prevalence of untested and incorrect assumptions, "myths", about the human body and how an individual may respond to lifestyle protocols and medical treatments commonly recommended by medical professionals, scientists, popular media, and fellow laypersons.

Reasons for Study and Purpose

Through data tracking and self-experimentation, these incorrect assumptions can be either validated or discredited, and prevailing claims thereby become truly "evidence-based". Unfortunately, it is often the layperson, rather than the researcher or healthcare professional, who wonders, "Why is 'evidence-based' even a term? Isn't all science and medicine supposed to be based on evidence?" The answer is that all science and medicine is indeed based on evidence, but often the wrong kind. The case studies presented aim to demonstrate the appropriate kind of evidence: testing in real-world contexts, not in controlled-environment laboratories, of the supposed mechanisms of biology, and its effects in unearthing unchallenged assumptions.

Case Study Subject

The particular individual described in these case studies was chosen for observation due to the subject's rare combination of having been an athlete of national level caliber and a patient with chronic illness within one decade. As such, the subject is at the two frontiers of greatest interest to researchers: the areas of (i) optimized physical performance, and (ii) treatment of chronic illness. Also, it is in these two predicaments, especially the latter, that subjects are most incentivized to invest the time, financial cost, and risk associated with adopting new AI technologies that may bring about knowledge to enhance their biological functioning.

Assumptions in Human Performance, Health, and Medicine

This study tests existing assumptions about the following four areas of human performance, health, and medicine (Table 1).

Area of Study	Assumption	Conclusion
Exercise	High volume weekly training is most effective for endurance athletes.	Further investigation needed.
Diet	Low calorie diet is most effective for health.	Assumption is false.
Hypothyroidism	Combination T3/T4 drug is not an effective option for treating hypothyroidism.	Assumption is false.
Sleep	Sleep quality can be increased by increasing sleep length or medicating. Other interventions are rarely considered.	Further investigation needed.

Table 1. Overview of assumptions tested and conclusions.

Significance to the AI Community

The implications of this paper aim to encourage researchers in AI to find solutions in technology that sift information into "tested" and "untested", or increase the adoption rate of self-tracking technologies by making them easier to use. The significance lies in the possibility that, in seeking solutions to reduce misinformed assumptions in human biology, AI researchers could simultaneously change the nature of HCI and how public health research is conducted.

Case Studies: Assumptions about Human Biology Tested by Female Patient-Athlete

The following document four experiments of the "subject", a female athlete at ages 13 through 20 who later became a hypothyroid patient. In each of the following cases, there will be a presentation of the common assumptions held by researchers and/or practitioners in the domain, then the experimental method taken to test the assumptions, any relevant background information and numerical calculations, and finally, the results and conclusion.

Exercise for Endurance Athletes

Existing Assumptions

Practitioners believe long distance training is best for long distance athletes, so the majority of the endurance athlete community trains, on average, 20 hours per week in preparation for a full Ironman distance triathlon race (Murphy, 2011), (Grasky Endurance, 2008), (Herrick, 2005). This is despite the fact that researchers have found that muscle fibers are the same or even better in endurance athletes who trained for short periods, than those who trained for long periods (Tabata et. Al, 1997), (Cherami, 2004).

Experiment Method

Subject exercised at 49% weekly training time of typical Ironman triathletes (9.29 hours per week on average) for 14 weeks in preparation for an Ironman distance race, from April 27, 2009 to July 31, 2009. The Ironman race was the following day, on August 1, 2009. Female subject, at 18 years of age, competed in the "F19 and under" age group (Table 2).

Week	Training Time (hrs)	Week	Training Time (Hrs)
1	8	8	12
2	9.5	9	8.5
3	8.5	10	6.75
4	11.5	11	6.5
5	11.25	12	12
6	11.75	13	8.75
7	12.5	14	2.5

Table 2. Hours trained per week in preparation for Ironman Race.

Background Information

An Ironman race entails: swimming 2.4 miles, biking 112 miles, and running 26.2 miles. Although "Ironman" connotes a standard distance as aforementioned, finish times of an Ironman race varies widely due to differing terrain and conditions at race venues worldwide (Table 3).

Age Group	Average Finish Time
all females	14:13 hrs
"19 and under" males & females	15:13 hrs
20-24 females	14:07 hrs

Table 3. Average Ironman finish times of past races at Guerneville, CA by age group (J-ChipUSA, 2006), (J-ChipUSA, 2010).

Results

Subject completed Ironman triathlon as first of two finishers in "19 and under" age group, in 15:30:28 hours, as youngest finisher (males and females combined) of the year at the international venue (Table 4).

Section	Time (hrs)
Swim	1:27
T1 (swim-to-bike)	0:05
Bike	7:22
T2 (bike-to-run)	0:12
Run	6:25

Table 4. Subject's finish times per section of race.

At the 8.5th mile of run section, projected finish time was 13:12 hours:

1:27 hrs swim + 5 min T1 + 7:22 hrs bike + 12 min T2 + 4:06 hrs run = 13:12 hrs

Projected 4:06 hrs run based on average running pace over first 8.5 mile section:

8.5mi/1.33hr = 6.39mph

26.2mi/6.39mph = 4.10 hr

However, actual run finish time was 6:25 hours, due to injury in both foot arches after first 9 miles, which dramatically decreased run pace.

26.2mi/6.42hrs = 4.08mph average

Conclusion

Due to the extraneous factor of foot injury, run pace slowed, and consequently average running speed was reduced to 4.08mph, thereby increasing total run time to 6:25 hrs. This caused actual race finish time to be approximately 2.5 hrs longer than projected finish time of 13:12 hrs.

However, if the subject were to finish within 1 hour of the projected time (12:12-14:12) at any Vineman race in the next 4 years following the original 2009 event (years of 2010-2013), subject would finish in the 50th percentile, at least, of her age group (females aged 20-24). If this were the case, the common assumption of endurance athletes needing 20 hours/week of training in preparation for an Ironman triathlon to perform competitively is false.

In summary, there is some supporting evidence for the assumptions of researchers over the assumptions of practitioners. More investigation is needed in order to discover more conclusive evidence of efficacy of low versus high weekly training volume for endurance athletes.

Diet and Nutrition

Existing Assumptions

Conventional science claims that low fat, low calorie diets to be effective in improving various aspects of health (Sacks et. Al, 2009), (Dhindsa, Scott, and Donnelly, 2003), (Nelson, 2009). Healthcare practitioners often recommend low fat, low calorie diets. However, they fail to recognize the low adherence rates of these diets, which may render them ineffective.

Experiment Method

Subject attempted 3 diets within a 7 year period, for minimum of 1 year per diet. Adherence was tracked daily in written diet journal. Daily caloric and carbohydrate intake was measured using NutritionData.com. Caloric expenditure was measured with standard calorie counter[1] (Table 5).

Diet Name	Description	Dates	Length
Low Cal	low fat, low calorie, moderate protein, high carb	2004-2007	3 yrs
Low Carb	low carb, moderate protein, moderate fat, no calorie restriction	2007-2008	1 yrs
Paleolithic	low carb, moderate protein, high/mod fat, no calorie restriction; abstinence from grains, legumes, dairy (except butter), potatoes, refined oils + sugars	2008-2011	3.5 yrs

Table 5. Description and length of diet experiments.

Results

Subject showed highest adherence on “Paleo” diet. Daily adherence for each diet was tracked on a binary, “yes” or “no” basis according to criteria listed below. Adherence rates were calculated as number of “yes” answers, divided by number of total days of diet attempt (Table 6).

Diet Name	Adherence	Adherence Y/N Criteria
Low Cal	35%	Caloric expenditure more than caloric intake?
Low Carb	61%	Ate less than 100 grams of carbohydrate?
Paleo	93%	Avoided eating all grains, legumes, dairy (except butter), potatoes, refined oils + sugars?

Table 6. Adherence to diet experiments.

Observations for each diet are as follows:

- “Low Cal” – constant hunger in between meals, cravings for high calorie foods, steady weight gain despite regular exercise, and facial eczema

1http://nutritiondata.self.com/tools/calories-burned/

- "Low Carb" – quick weight loss in first month, then slow progress; decreased hunger in between meals and cravings for food
- "Paleo" – satiety lasting up to 12 hours, few cravings, increased energy; rapid and easily sustained weight loss; reduced eczema, menstrual cramping, bloating in face, and morning phlegm accumulation in throat and nasal cavities; eliminated indigestion and abdominal cramps

Conclusion

Assumptions of majority of practitioners and researchers are false, both when adherence rates and qualitative observations are considered.

Medical Treatment for Hypothyroidism

Existing Assumptions

Combined thyroxine/liothyromine treatment does not improve well-being compared to thyroxine alone for patients with hypothyroidism (Walsh et. Al, 2003).

Background Information

Hypothyroidism symptoms include: weight gain, swollen face, dry skin, indigestion, low energy, and sensitivity to changes in outside temperature. Subject experienced symptoms, primarily weight gain, since June 1st, 2010, and was diagnosed with hypothyroidism on May of 2011.

Experiment Method

- Subject measured adipose tissue percentage of total body weight, on May 7th, 2011.
- Subject started thyroid treatment (pill taken orally, with 5mcg T3 and 20mcg T4) on May 25th, 2011.
- Subject measured body composition again 4 months later, on Sept 7th, 2011.

Body composition was measured with DEXA scan procedure by Body Composition Center in Redwood City, CA. Subject sustained sedentary lifestyle prior to and throughout experiment period. Three laboratory tests are performed in same period by Palo Alto Medical Foundation Laboratory to measure subject's blood levels of thyroid-stimulating hormone (TSH), the "gold standard" metric for monitoring treatment of hypothyroidism (Hershman and Berg, 1997).

Results

Subject lost 4% total adipose fatty tissue, from 27.6% to 23.6%, over 4 month period. Subject showed improvement in TSH levels (Table 7).

	May	June	July	Aug	Sept	Oct	Trend
TBFC	27.6	-	-	-	23.6	-	dec
TSH	1.86	-	1.46	-	-	0.8	dec

Table 7. Changes in total body fat composition % (TBFC) and thyroid-stimulating hormone (TSH) levels (in mIU/mL) during T3/T4 combination for hypothyroid patient ("dec" = decrease).

Conclusion

After treatment with combination T3/T4 drug, hypothyroid patient showed improvement in condition over 6 month period, based on laboratory results and symptoms. This fails to support existing assumptions, and warrants further investigation by medical practitioners and pharmaceutical researchers.

Sleep Quality

Existing Assumptions

To increase sleep quality in patients, healthcare practitioners rely on one behavioral approach (sleeping longer) and several medicinal approaches. Rarely are other behavioral approaches explored, with the assumption that behavioral changes are not as effective as medication (Block, 2007), (JAMA and Archives Journals, 2006).

Experiment Method

Subject used different sleep times, bedtimes, and intermittent introduction of a pre-bed dietary intervention to observe changes in sleep quality score. Dietary intervention consisted of: chicken broth with animal fat (beef tallow, butter, and marrow bone fat).

Subject tracked sleep quality using Zeo Sleep Coach device, which measures brain activity indicating the four different stages of sleep (REM, deep sleep, light sleep, and waking) through a head strap worn during sleep, then provides a quantitative sleep quality value called "Z score".

Background Information

Average Z score for subject's age group (20-29 years) is 86[2].

Results

Length of sleep is slightly positively correlated with sleep quality. Increased sleep quality on nights with pre-bed dietary intervention of chicken broth with animal fat. Average with dietary intervention was 107, significantly higher than subject's average baseline Z score of 82, adjusted for sleep length factor (Figure 1) – assuming linear correlation between Z score and sleep length:

baseline average length = 7.97 hrs
intervention average length = 8.38 hrs
expected Z average = 86
actual Z average = 107

2http://whatiszeo3.myzeo.com/hp/3/whats-your-zq/

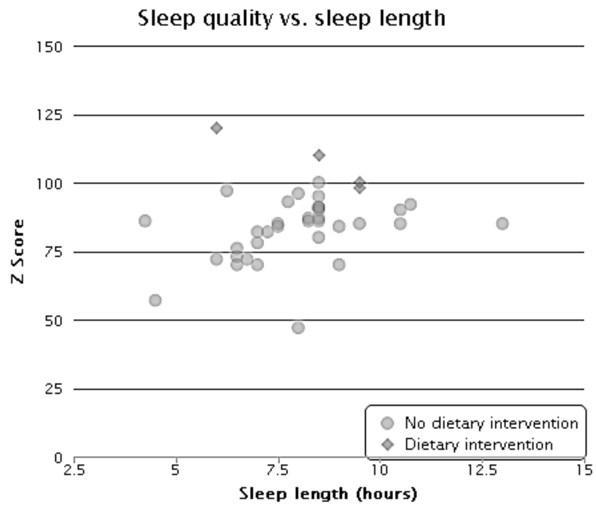

Figure 1. Baseline and intervention sleep quality vs. sleep length.

Earlier bedtimes are somewhat correlated with sleep quality (Figure 2).

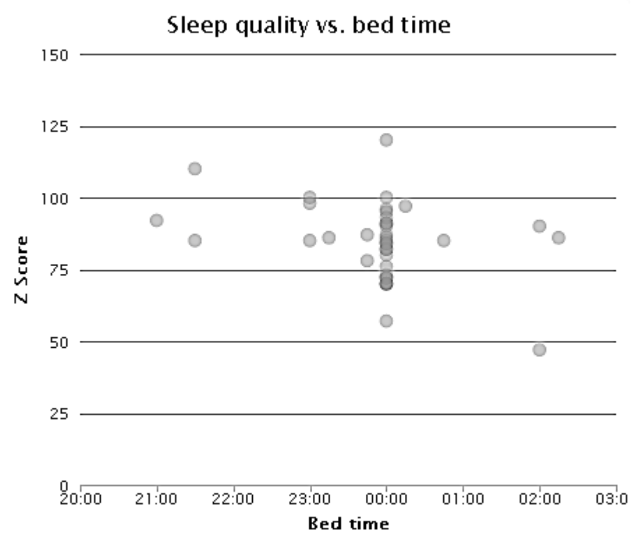

Figure 2. Sleep quality vs. bed time.

Conclusion

The assumption by medical practitioners that we need pharmaceuticals to effectively improve sleep quality may be false. Self-tracking and experimentation with behavioral factors such as pre-bed foods, bed times, and sleep length can lead to increased sleep quality. Further investigation is needed to identify the most effective behavioral changes in sleep factors.

Discussion

This study tests existing assumptions about the following four areas of human performance, health, and medicine: (i) exercise, (ii) diet, (iii) medical treatment of hypothyroid, and (iv) sleep. Results indicate two areas in which existing assumptions are not supported, and two other areas in which further investigation is warranted.

Although there are various options for the research community at large to apply our findings, we suggest one suite of options in particular for the AI community. We need technology that promotes self-tracking and experimentation by individuals, in real-world contexts, especially in the areas of healthcare treatment and physical performance enhancement. Gathering of real-world data by individuals would accelerate innovation and increase focus on evidence-based protocols rather than on hypothetical possibilities available only in highly controlled laboratory environments.

Therefore, the AI community can enhance public health in the following ways:

- promoting experimentation (eg. crowd-sourced clinical trials with Genomera and DIYgenomics)
- developing machine learning algorithms that meaningfully organize data obtained from aforementioned crowd-sourced clinical trials
- promoting engagement in online patient communities via anonymization of user data (eg. Patientslikeme, Curetogether)
- using HCI research insights without the pitfalls of failing to test effectiveness in real-world contexts (Pierce et. al, 2010) in order to engage members of the general population who are, by definition, less motivated than early adopter users of self-tracking technologies

Acknowledgments

The author would like to thank Takashi Kido, Melanie Swan, Stanford University, and AAAI for allowing the opportunity for the findings of this study to be published.

References

Block, A. R. 2007. Using Medication to Manage Pain and Reduce Sleep Problems. *Spine-Health.com*. Retrieved Jan 18, 2012 from http://www.spine-health.com/treatment/pain-medication/using-medication-manage-pain-and-reduce-sleep-problems

Cheramie, W. M. 1999. Effects of Aerobic and Anaerobic Training Protocols on 4000m Track Cycling Time Trial. M.Sc. thesis, Department of Kinesiology, Louisiana State University, Eunice, LA.

Dhindsa, P., Scott, A. R., and Donnelly, R. 2003. Metabolic and Cardiovascular Effects of Very-Low-Calorie Diet Therapy in Obese Patients with Type 2 Diabetes in Secondary Failure: Outcomes After 1 Year. *Diabet Med* 20(4):319-324.

Grasky Endurance. 2008. Triathlon Training Plans. *Phoenix Triathlon Club*. Retrieved Jan 18, 2012 from http://www.phoenix-triathlonclub.org/Training.html

Herrick, Scott. 2005. Beginner Full Ironman Training Program. *BeginnerTriathlete.com*. Retrieved Jan 18, 2012 from http://beginnertriathlete.com/cms/article-detail.asp?articleid=441

Hershman, J. M., and Berg, L. 1997. Rapid Qualitative TSH Test to Screen for Primary Hypothyroidism. *American Association for Clinical Chemistry* 43:1097-1098.

JAMA and Archives Journals. 2006. Effects of New Sleep Medication Appear Unlikely to have Potential for Abuse or Cognitive Impairment. *ScienceDaily.com*. Retrieved Jan 18, 2012 from http://www.sciencedaily.com/releases/2006/10/061002215138.htm

J-ChipUSA. 2006. Full Vineman Triathlon and Barb's Race 17th Annual Results. *J-ChipUSA.com*. Retrieved Jan 18, 2012 from http://www.j-chipusa.com/cgi-bin/pdf_resage.pl?eventid=1943

J-ChipUSA. 2010. Full Vineman Triathlon and Barb's Race 21st Annual Results. *J-ChipUSA.com*. Retrieved Jan 18, 2012 from http://www.j-chipusa.com/cgi-bin/pdf_resage.pl?eventid=3093

Murphy, P. 2011. The Daily Schedule of Ironman Training with a Full-Time Job. *LiveStrong.com*. Retrieved Jan 18, 2012 from http://www.livestrong.com/article/351462-the-daily-schedule-of-ironman-training-with-a-full-time-job/

Nelson, L. 2009. Eating Too Much Red Meat Increases Heart Disease Risk. *HealthCentral.com*. Retrieved Jan 16, 2012 from http://www.healthcentral.com/heart-disease/c/45112/74344/increases

Pierce, J., Fan, C., Lomas, D., Marcu, G., and Paulos, E. 2010. *Some Considerations on the (In)effectiveness of Residential Energy Feedback Systems*. New York: ACM Press.

Sacks, F. M., Bray, G. A., Carey, V. J., Smith, S. R., Ryan, D. H., Anton, S. D., McManus, K., Champagne, C. M., Bishop, L. M., Laranjo, N., Leboff, M. S., Rood, J. C., de Jonge, L., Greenway, F. L., Loria, C. M., Obarzanek, E., and Williamson, D. A. 2009. Comparison of Weight-Loss Diets with Different Compositions of Fat, Protein, and Carbohydrates. *N Engl J Med* 360:859-873.

Tabata, I., Irisawa, K., Kouzaki, M., Nishimura, K., Ogita, F., and Miyachi, M. 1997. Metabolic Profile of High Intensity Intermittent Exercises. *Med. Sci. Sports Exerc.* 29(3):390-395.

Walsh, J.P., Shiels, L,. Lim, E. M., Bhagat, C. I., Ward L. C., Stuckey, B. G. A., Dhaliwal, S. S., Chew, G. T., Bhagat, M. C., and Cussons, A. J. 2003. Combined Thyroxine/Liothyromine Treatment Does Not Improve Well-Being, Quality of Life, or Cognitive Function Compared to Thyroxine Alone: A Randomized Controlled Trial in Patients With Primary Hypothyroidism. *J Clin Endocrinol Metab* 88:4543-4550.