Automatic Drug Side Effect Discovery from Online Patient-Submitted Reviews: Focus on Statin Drugs

Jingjing Liu, Alice Li and Stephanie Seneff
MIT Computer Science & Artificial Intelligence Laboratory
Cambridge, MA 02139, USA
{jingl, aliceli, seneff@csail.mit.edu}

Abstract—**In recent years, consumers have become empowered to share personal experiences regarding prescription drugs via Web page discussion groups. This paper describes our recent research involving automatically identifying adverse reactions from patient-provided drug reviews on health-related web sites. We focus on the statin class of cholesterol-lowering drugs. We extract a complete set of side effect expressions from patient-submitted drug reviews, and construct a hierarchical ontology of side effects. We use log-likely ratio estimation to detect biases in word distributions when comparing reviews of statin drugs with age-matched reviews of a broad spectrum of other drugs. We find a highly significant correlation between statins and a wide range of disorders and conditions, including diabetes, amyotrophic lateral sclerosis (ALS), rhabdomyolysis, neuropathy, Parkinson's disease, arthritis, memory loss, and heart failure. A review of the research literature on statin side effects corroborates many of our findings.**

Keywords- medicine data mining; drug side effect discovery

I. INTRODUCTION

The last few decades have witnessed a steady increase in drug prescriptions for the treatment of biometric markers rather than overt physiological symptoms. Today, people regularly take multiple drugs in order to normalize serum levels of biomarkers such as cholesterol or glucose, or to reduce blood pressure. All drugs have side effects, which are sometimes debilitating or even life-threatening. When a person taking multiple drugs experiences a new symptom, it is not always clear which, if any, of the drugs or drug combinations are responsible.

Increasingly, consumers are turning to the Web to seek information, and, increasingly, this information comes in the form of consumer-provided comments in discussion groups or chat rooms. User reviews of products and services have empowered consumers to obtain valuable data to guide their decision process. Recently, statistical and linguistic methods have been applied to large datasets of reviews to extract summary and/or rating information in various domains ([9] [22]).

Health care and prescription drugs are a growing topic of discussion online, not surprising given that almost half of all Americans take prescription drugs each month, costing over $200 billion in 2008 alone ([5]). Though drugs are subject to clinical trials before reaching market, these trials are often too short, and may involve too few people to give conclusive results. A large study recently conducted on the heart failure drug, nesiritude, invalidated the findings of the smaller study that had led to the drug's approval [11]. While regulatory agencies do attempt to monitor the safety of approved medical treatments, surveillance programs such as the U.S. Food and Drug Administration's (FDA's) and Adverse Event Reporting System (AERS) are often difficult for patients to use.

In addition, the large language gap between medical documents and patient vocabulary can cause confusion and misunderstanding ([23]). We hope to take advantage of the vast amount of information available in patient anecdotes posted online to address the dual problems of insufficient clinical studies and mismatched terminology.

We envision a system that increases patient awareness of drug-related side effects by enabling consumers of prescription drugs to easily browse a large consolidated database of posts from health-related web sites. Beyond aggregating data from drug review and health discussion sites, we plan to support spoken queries, which would be answered via a set of succinctly summarized hits that best match the query, based on sophisticated statistical and linguistic techniques. The user could then click on any one of these displayed summaries to read the associated post.

This paper describes our preliminary efforts to detect associations between a drug class and its side effects. We use statistics and heuristic methods to build up a hierarchical ontology of side effects by aggregating patient-submitted drug reviews. We use log-likelihood ratios to extract summary information derived from biases in word and phrase distributions, and to quantify associations between drugs and symptoms. For the scope of this paper, we focus on statin drugs, which are among the most costly and commonly prescribed drugs in the United States. The methods described are applicable to all drug classes.

In the remainder of this paper, we will first review the research literature reflecting known or suspected side effects associated with statin drugs. After explaining our data collection and side-effect ontology construction, we describe our methodology and verify that many of our extracted associations align with observations from the literature.

II. BRIEF LITERATURE REVIEW

A. Side Effects of Statin drugs

Statins (Hydroxy methyl glutaryl coenzyme A reductase inhibitors) have become increasingly popular as very

effective agents to normalize serum cholesterol levels. The most popular of these, atorvastatin, marketed under the trade name, Lipitor, has been the highest revenue branded pharmaceutical for the past 6 years. The official Lipitor web site lists as potential side effects mainly muscle pain and weakness and digestive problems. However, several practitioners and researchers have identified suspected side effects in other more alarming areas, such as heart failure, cognition and memory problems, and even severe neurological diseases such as Parkinson's disease and ALS (Lou Gehrig's disease). [21] provides compelling arguments for the diverse side effects of statins, attributing them mainly to cholesterol depletion in cell membranes.

It is widely acknowledged that statin drugs cause muscle pain, weakness and damage ([7] [12]), likely due in part to their interference with the synthesis of the potent antioxidant Coenzyme Q10 (CoQ10) ([10]). CoQ10 plays an essential role in mitochondrial function to produce energy. Congestive heart failure is a condition in which the heart can no longer pump enough blood to the rest of the body, essentially because it is too weak. Because the heart is a muscle, it is plausible that heart muscle weakness could arise from long-term statin usage. Indeed, atorvastatin has been shown to impair ventricular diastolic heart performance ([14]). Furthermore, CoQ10 supplementation has been shown to improve cardiac function ([13] [20]).

The research literature provides plausible biological explanations for a possible association between statin drugs and neuropathy ([15] [24]). A recent evidence-based article ([1]) found that statin drug users had a high incidence of neurological disorders, especially neuropathy, parasthesia and neuralgia, and appeared to be at higher risk to the debilitating neurological diseases, ALS and Parkinson's disease. The evidence was based on careful manual labeling of a set of self-reported accounts from 351 patients. A mechanism for such damage could involve interference with the ability of oligodendrocytes, specialized glial cells in the nervous system, to supply sufficient cholesterol to the myelin sheath surrounding nerve axons. Genetically-engineered mice with defective oligodendrocytes exhibit visible pathologies in the myelin sheath which manifest as muscle twitches and tremors ([16]).

Cholesterol depletion in the brain would be expected to lead to pathologies in neuron signal transport, due not only to defective myelin sheath but also to interference with signal transport across synapses ([17]). Cognitive impairment, memory loss, mental confusion, and depression were significantly present in Cable's patient population ([1]). Wagstaff et al. ([19]) conducted a survey of cognitive dysfunction from AERS data, and found evidence of both short-term memory loss and amnesia associated with statin usage. Golomb et al. ([6]) conducted a study to evaluate evidence of statin-induced cognitive, mood or behavioral changes in patients. She concluded with a plea for studies that "more clearly establish the impact of hydrophilic and lipophilic statins on cognition, aggression, and serotonin."

B. Relationship between Cholesterol and Health

ALS and heart failure are both conditions for which published literature suggests an increased risk associated with statin therapy ([1] [10]). Indeed, for both of these conditions, a survival benefit is associated with elevated cholesterol levels. A statistically significant inverse correlation was found in a study on mortality in heart failure. For 181 patients with heart disease and heart failure, half of those whose serum cholesterol was below 200 mg/dl were dead three years after diagnosis, whereas only 28% of the patients whose serum cholesterol was above 200 mg/dl had died. In another study on a group of 488 patients diagnosed with ALS, serum levels of triglycerides and fasting cholesterol were measured at the time of diagnosis ([2]). High values for both lipids were associated with improved survival, with a p-value <0.05.

A very recent study on the relationship between various measures of cholesterol status and health in the elderly came up with some surprising results, strongly suggesting that elevated cholesterol is beneficial for this segment of the population [18]. A study population initially over 75 years old was followed over a 17 year period beginning in 1990. In addition to serum cholesterol, a biometric associated with the ability to synthesize cholesterol (lathosterol) and a biometric associated with the ability to absorb cholesterol through the gut (sitosterol) were measured. For all three measures of cholesterol, low values were associated with a poorer prognosis for frailty, mental decline and early death. A reduced ability to *synthesize* cholesterol showed the *strongest* correlation with poor outcome. Individuals with high measures of all three biometrics enjoyed a 4.3 year extension in life span, compared to those for whom all measures were low.

III. SIDE-EFFECT DISCOVERY

A. Data Collection

To learn the underlying associations between side effects and drug usage from patient-provided reviews, we collected drug reviews from three drug discussion forums ("AskPatient.com," "Medications.com" and "WebDB.com") which allow users to post reviews on specific drugs and share their experiences. Table 1 gives the statistics on the review data collection. A total of 8,515 statin reviews were collected from the three data sources. We also collected 105K drug reviews from the AskPatient.com, on drugs to treat a broad range of problems such as depression, acid reflux disease, high blood pressure, diabetes, etc. This set includes reviews for non-statin cholesterol lowering drugs.

Table 1. Statistics on drug review data collection.

Data source	Number of Statin reviews
AskPatient.com	2,647
Medications.com	4,162
WebMD.com	1,706
Total	8,515

A typical review entry contains the personal information of the user (e.g., gender, age), the dosage and duration of the drug treatment, the reason for taking the drug, the side effects that the user has experienced, as well as a free-style text comment. An example of a review is shown in Figure 1.

:drug "Lipitor"
:dosage "40mg 1X D"
:sex "Male"
:age "47"
:duration "4 years"
:reason "high cholesterol"
:side_effects "*Body aches, joint pain, decreased mobility, decreased testosterone and libido, difficulty getting out of bed in the morning, tingling and itchy hands, and decrease in overall strength.*"
:comment "I have been taking lipitor for many years. I started out on 10mgs and now I am on 40mgs. I have had hip replacement, back surgery, and shoulder surgery while on this drug. I have seen my strength decrease dramatically …"

Figure 1. Example of a review from AskPatient.com.

B. *Side-Effect Extraction*

Most previous medical natural language processing research relies on medical lexicons such as those provided by Unified Medical Language System (UMLS) or the Food and Drug Administration's (FDA) COSTART corpus. However, these official lexicons often have low coverage of colloquial side effect expressions, which are very common in patient-submitted reviews. Thus, in this study we extract side effect expressions from the reviews themselves, instead of using these restrictive lexicons.

As shown in the example in Figure 1, the input of "side effects" often contains a list of short phrases describing the major side effects the reviewer has, and the input of "comment" contains free-style texts which tell the story about the reviewer's experience with the drug. To obtain a clean set of side effect expressions, we first automatically extract short phrases from the input of "side_effects" in each review entry. From the 107K reviews collected from AskPatient.com (including both statin and non-statin drugs), we extracted 7,500 words and phrases that describe common side effects on various drugs.

These phrases extracted from general users' input contain a lot of noise. For example, some users may type in long sentences describing their conditions instead of using short phrases. Also, many phrases may describe the same type of side effect (e.g., "joint pain," "pain in joint" and "severe pain in the arm and leg joint"). To eliminate the noise and redundancy, the extracted phrases were first subjected to a stop-word filter, eliminating 377 common stop words. A phrase which contains only stop words is filtered out as noise. We also filter out the phrases which have frequency counts less than five in the whole review dataset. We further filter the phrases by grouping phrases containing the same set of non-stop-word (e.g., "joint pain" and "pain in joint"). With this filtering process, the number of the side effect phrases shrinks to 2,314.

C. *Side-Effect Onology*

To organize the set of side effect phrases, we asked an annotator who is knowledgeable in medical terminology to classify the phrases into a hierarchical ontology. First, synonyms are identified and grouped (as shown in Table 2). For example, "elevated blood pressure," "increase in blood pressure" and "higher blood pressure" are clustered into the same group. Then, these synonym groups are further organized into broad classes. For example, the synonym groups of "achy legs," "muscle pain" and "joint pain" are all clustered into the generic class of "pain."

Table 2. Examples of the synonyms of side effects.

Group	Synonyms
loss of mental clarity	mental slowness, slow brain, fuzzy thinking, foggy head, cloudy head, muddled thinking
all body aches	achy body, achy feeling, achy bones, achy all over, overall aches, body ache, aches and pains
forgetting words	difficulty finding the right word, mixing up words, can't find words, difficulty finding words
diabetes	diabetic, high blood sugar, elevated blood sugar

As a result of this manual process, the 2,314 side effect phrases are clustered into 307 synonym groups, which are furthered grouped into 30 classes. Table 3 shows the 30 classes as well as the number of synonym groups in each class, and Table 4 gives examples of the synonym groups in some classes. Note that this classification schema encompasses side effects for all drugs, and can be used for other drug classes besides statins.

Table 3. Classes of side effects.

Class	#Syn. groups	Class	#Syn. groups	Class	#Syn. groups
aches	11	eyes	6	mouth	11
appetite	6	hair	4	muscle	11
arthritis	4	heart	10	nerve	24
blood	9	infection	10	pain	29
breasts	4	kidney	12	skin	15
breathing	5	libido	9	sleep	13
cognition	13	liver	4	swelling	12
conditions	10	menstrual	8	taste	3
digestion	20	mobility	5	temperature	3
ears	4	mood	27	weight	5

Table 4. Example groupings of side effects into classes.

Class	Synonym groups
cognition	brain shocks, clearer thinking, cognitive problems, dementia, loss for words, loss of mental clarity, memory problems, mental instability, problems concentrating, short attention span
heart	atrial fibrillation, heart attack, heart failure, heart valve, heart palpitations, high heart rate, high pulse, low heart rate, tightness in chest, potassium
mood	aggressive behavior, anxiety, bipolar, bizarre thoughts, blunted emotions, crying easily, depression, despair, disoriented, euphoria
muscle	fatigue, loss of muscle mass, loss of muscle tone, muscle cramps, muscle pain, muscle spasms, muscle tightness, muscle weakness, rhabdomyolysis

D. Association of Drug Class with Side-Effects

Given the hierarchical ontology of side effects, we can now discover which side effects are strongly associated with statin drugs. For this, we make use of log-likelihood ratio ([3]). In statistics, a likelihood ratio test is used to compare the fit of two models, one of which (the null model) is a special case of the other (the alternative model). The test is based on the likelihood ratio, which expresses how many times more likely the data are under one model than the other. This likelihood ratio, or equivalently its logarithm, can then be used to compute a *p*-value to decide whether to reject the null model in favor of the alternative model.

To apply the log likelihood ratio algorithm, we treat the side effect association problem as a coin toss model. The set of reviews on statin drugs (R_1) is analogous to a coin *A*. The set of reviews on non-statin drugs (R_2) is analogous to a coin *B*. Each review (in R_1 or R_2) is a coin toss instance. For a certain side effect phrase t, if a review contains the phrase, the "Head" of the coin shows up; otherwise, the "Tail" shows up. Thus, the null hypothesis is that coin *A* and coin *B* have the same probability of showing "Head" or "Tail," i.e., the review set R_1 and R_2 have the same probability of containing the phrase. The alternative model is that coin *A* and coin *B* have different probabilities of showing up "Head" or "Tail," i.e., one review set (R_1 or R_2) has a higher probability of containing the phrase than the other. The measurement of the hypothesis that the phrase t is more likely to occur in the set of statin reviews (R_1) is calculated by:

$$L_1 = k_1 log \frac{p_1}{p} + (n_2 - k_2) log \frac{1-p_2}{1-p} \quad (1)$$

where k_1 is the counts of statin reviews that contain the side effect phrase t, k_2 is the counts of non-statin reviews that contain the phrase t, p_1 is the probability of the phrase t occuring in R_1, p_2 is the probability of the phrase t occuring in R_2, p is the probability of the phrase t occurring in the whole document set ($R_1 \cup R_2$), n_1 is the size of R_1, and n_2 is the size of R_2.

Maximum likelihood is achieved by:

$$p_1 = \frac{k_1}{n_1} \qquad p_2 = \frac{k_2}{n_2} \quad (2)$$

$$p = \frac{k_1+k_2}{n_1+n_2} \quad (3)$$

A symmetric equation of (1) can be derived for L_2, the hypothesis that the phrase t occurs more frequently in R_2. Whether the alternative model fits significantly better and should thus be preferred can be determined by deriving the probability or *p*-value of the obtained difference $L_1 - L_2$. The probability distribution of the difference can be approximated by a chi-square distribution. *P*-values are computed under the assumption that there is one degree of freedom between the null model and the alternative model.

IV. EXPERIMENTS

In our dataset of reviews, the size of non-statin reviews (105K) is much larger than that of statin reviews (8,515). To make the two document sets equivalent for comparison, we randomly select the same size of reviews (8,515) from the non-statin reviews as R_2. An important consideration is to correct for a possible age bias of review-providers in the data selection process. Figure 2 gives the age distribution of statin drug reviewers. We follow the same distribution of reviewers' age for the random selection of non-statin reviews. We observed that, in the statin review set, most of the reviews (83.6%) are published by users aged from 40 to 70.

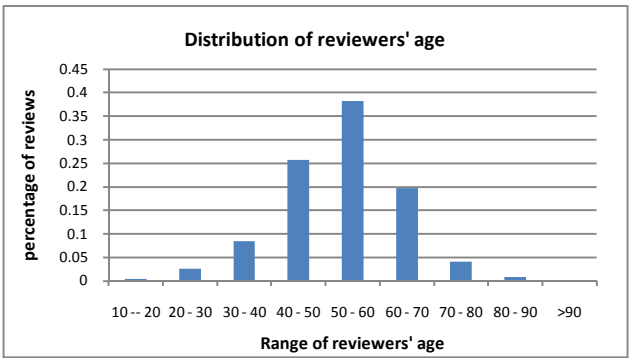

Figure 2. Distribution of reviewers' age in statin reviews.

For each side effect in our hierarchy ontology, we calculate the log likelihood ratio of the statin review set and non-statin review set as explained in the algorithm section. We treat all the synonyms for each side-effect equally, i.e., the occurrences of alternatives in the same group count for the same phrase t.

Table 5 lists all side effect clusters related to that yielded a *p*-value less than 0.05. Pain in essentially all parts of the body -- arms, legs, neck, shoulder, and back, all occurred more frequently in the statin reviews by a substantial margin. "Muscle pain" in particular is overwhelmingly associated with statins, with a *p*-value of 1.4E-06.

Table 5: Pain in various parts of the body pain (sorted by *p*-value).

Side effect	k_1	k_2	$L_1 - L_2$	*p*-value
muscle pain	1029	221	1419.26	1.4E-06
pain	2499	1557	1444.31	0.00004
pain in legs	570	265	514.16	0.00114
shoulder pain	142	34	189.65	0.00485
back pain	323	163	265.20	0.00738
neck pain	111	36	130.77	0.01417
pain in arms	76	21	96.38	0.02009

Table 6 provides the counts and *p*-values for a number of side effect clusters associated with muscle frailty and pathology. Highly disturbing is the very low *p*-value for "difficulty walking" (0.0004). Rhabdomyolysis is a frequently fatal condition involving kidney failure due to toxic exposure to myoglobin debris released into the blood stream following muscle breakdown. Since it is generally rare, it did not occur at all in the non-statin reviews, but appeared 31 times in the statin reviews. "Loss of muscle

mass" is ten times as frequent in the statin reviews. "Muscle cramps," "general weakness" and "muscle weakness," highly associated with "frailty," have extremely low *p*-values. In addition, "general numbness" and "muscle spasms" are also significantly associated with statins.

Table 6: Muscle frailty and pathology pain (sorted by p-value).

Side effect	k_1	k_2	$L_1 - L_2$	*p*-value
muscle cramps	678	193	850.12	0.00005
general weakness	687	210	834.24	0.00006
muscle weakness	302	45	448.73	0.00023
difficulty walking	419	128	508.96	0.00044
loss of muscle mass	54	5	84.75	0.01332
general numbness	293	166	203.34	0.01552
muscle spasms	136	57	135.03	0.01849
rhabdomyolysis	31	0	51.52	0.02177
tendonitis	42	8	59.68	0.03193
balance problems	71	32	65.91	0.05371

Table 7 shows the frequency distributions for several often debilitating conditions associated with pathologies in the brain and nervous system. Most alarming to us is the 10:1 ratio of incidence of ALS, a debilitating disease associated with damaged motor neurons in the spinal cord that is nearly always fatal. The associated *p*-value of 0.008 makes this result highly significant. The ratio is even higher for Parkinson's disease (18:1). Parkinson's disease involves damage to dopamine-secreting cells in the substantia nigra. The *p*-value for memory problems is also very low (0.01), providing powerful evidence that statins cause memory problems. An extreme form of memory problems, dementia, comes in with a *p*-value just above the significance cutoff at 0.056. Neuropathy, due to nerve damage in the peripheral nervous system, is generally associated with muscle weakness, cramps, and spasms, other side effects that occur very frequently in statin drug reviews.

Table 7: Issues related to brain and nervous system.

Side effect	k_1	k_2	$L_1 - L_2$	*p*-value
ALS	71	7	110.75	0.00819
memory problems	545	353	286.76	0.01118
Parkinson's disease	53	3	85.38	0.01135
neuropathy	133	73	97.03	0.04333
dementia	41	13	48.80	0.05598

Table 8 shows other major health issues for which the word frequencies are highly skewed towards the statin reviews. Most of these distributions are highly significant, with a p-value < 0.01. Diabetes is especially striking, with three times the frequency of occurrence in the statin reviews as in the other reviews, despite the fact that diabetes medications are included in the other class. The highly significant results for diabetes are in line with recent concern about the possibility that statins may increase risk to diabetes ([4] [8]).

"Heart attack" has an extremely low *p*-value, but in this case strong compounding from a precondition is undeniable. A similar issue arises with "stroke." The strongest correlation among the remaining conditions is found for liver damage ($p<0.003$), a potential side effect that is acknowledged by the statin manufacturers. It is interesting that arthritis associates strongly with statins, as arthritis has not been identified as a known side effect. Also, "heart failure" and "raised liver enzymes" are under the cutoff of 0.05, and "kidney failure" is six times as frequent in the statin reviews with a *p*-value just above 0.051.

Table 8: Various other conditions.

Side effect	k_1	k_2	$L_1 - L_2$	*p*-value
heart attack	299	73	396.87	0.00068
liver damage	326	133	331.15	0.00285
diabetes	185	62	214.2	0.00565
stroke	147	44	180.18	0.00700
arthritis	245	120	208.51	0.01117
raised liver enzymes	61	22	67.52	0.04204
heart failure	36	8	49.17	0.04473
kidney failure	26	4	38.56	0.05145
kidney damage	87	45	69.05	0.05949

To learn the high level association between side effects and statin drugs, we further aggregate the side effects into classes, and calculate the log likelihood ratio as well as the *p*-values for each class. Table 9 gives the top-ranked classes with *p*-values below 0.05 for statin reviews. These categories are considered as most strongly associated with statin drugs. In particular, "muscle problems" is overwhelmingly associated with statins, with a *p*-value of 2.0E-07.

Table 9. Side effect classes associated with statin drugs.

Class of side effect	k_1	k_2	$L_1 - L_2$	*p*-value
muscle problems	4188	2060	3549.73	2.0E-07
mobility	535	199	581.47	0.00049
liver problems	404	163	413.99	0.00166
pain	4735	3908	731.07	0.00308
nerve problems	1196	894	380.06	0.01108
arthritis	456	317	194.72	0.02690

V. DISCUSSION

In this paper, we have described our vision of a Web-based database providing potential users with a rich facility for exploring the association of prescription drugs with possible side effects. We used the basic strategy of comparing word frequency distributions between two databases as a means to uncover statistically salient phrase patterns. Our efforts focused on statin drugs, as these are a widely prescribed medication with diverse side effects. Through standard statistical log likelihood ratio estimation, we have shown that statin drugs are very strongly associated with muscle pain and weakness, and that there is as well a statistically significant association between statin drugs and several debilitating diseases, such as ALS, Parkinson's disease, rhabdomyolysis, and heart failure. Many of our findings are supported by the research literature on statins.

Our research was inspired by the study conducted by Jeff Cable ([1]). While he looked at only 350 reviews, he used careful manual analysis to deduce associated side effects. We looked at a much larger set of reviews (over 8,000), and used

statistical techniques for analysis. On the one hand, it is gratifying that both methods uncovered similar side-effect profiles on different data. On the other hand, it is disturbing that a drug class as widely prescribed as the statin drugs has such severe and sometimes life-threatening adverse reactions.

One limitation of the method is the compounding effects of preconditions. This clearly influences the bias for statins with regard to "heart attack" and "stoke," but may also contribute to other terms such as diabetes and heart failure. In addition, users occasionally post comments that discuss potential side effects that they did not personally experience.

VI. Future Work

In the future, we will focus on incorporating the results of our statistical analyses into the user database. We will also develop techniques to summarize individual reviews and provide associated index terms. An ambitious goal is to use parsing techniques to extract a story line that captures cause-and-effect relationships. For instance, by commenting, "It's only been 2 days without the medication and cramping is improving," a user clearly implied that the drug had caused the cramping. We also plan to expand the database to other drug classes, such as psychopharmaceuticals and acid reflux therapies. Finally, we would like to provide a speech-based interface for querying the database.

Acknowledgment

This research is supported by Quanta Computers, Inc. through the T-Party project.

References

[1] J. Cable. 2009. Adverse Events of Statins - An Informal Internet-based Study. JOIMR, 7(1).

[2] J. Dorstand, P. K¨uhnlein, C. Hendrich, J. Kassubek, A.D. Sperfeld, and A.C. Ludolph. 2010. Patients with elevated triglyceride and cholesterol serum levels have a prolonged survival in amyotrophic lateral sclerosis. J Neurol, in Press:Published online Dec. 3 2010.

[3] T. Dunning. 1993. Accurate methods for the statistics of surprise and coincidence. Computational linguistics, 19(1):61–74.

[4] M.R. Goldstein and L. Mascitelli. 2010. Statin-induced diabetes: perhaps, its the tip of the iceberg. QJM, Published online, Nov 30.

[5] Q. Gu, CF Dillon, and VL Burt. 2010. Prescription drug use continues to increase: U.S. Prescription drug data for 2007-2008. NCHS data brief, (42):1.

[6] B.A. Golomb, M.H. Criqui, H. White, and J.E. Dimsdale. 2004. Conceptual foundations of the UCSD Statin Study: a randomized controlled trial assessing the impact of statins on cognition, behavior, and biochemistry. Archives of internal medicine, 164(2):153.

[7] J. Hanai, P. Cao, P. Tanksale, S. Imamura, E. Koshimizu, J. Zhao, S. Kishi, M. Yamashita, P.S. Phillips, V.P. Sukhatme, et al. 2007. The muscle-specific ubiquitin ligase atrogin-1/MAFbx mediates statin-induced muscle toxicity. Journal of Clinical Investigation, 117(12):3940–3951.

[8] J. Hagedorn and R. Arora. 2010. Association of Statins and Diabetes Mellitus. American journal of therapeutics, 17(2):e52.

[9] J. Liu and S. Seneff. 2009. Review sentiment scoring via a parse-and-paraphrase paradigm. In Proc. EMNLP, pages 161–169. Association for Computational Linguistics.

[10] P.H. Langsjoen and A.M. Langsjoen. 2003. The clinical use of HMG CoA-reductase inhibitors and the associated depletion of coenzyme Q10. A review of animal and human publications. Biofactors, 18(1):101–111.

[11] Califf R. LBCT I, Abstract 21828. Presented at: American Heart Association Scientific Sessions 2010; Nov. 13-17; Chicago.

[12] M.G. Mohaupt, R.H. Karas, E.B. Babiychuk, V. Sanchez-Freire, K. Monastyrskaya, L. Iyer, H. Hoppeler, F. Breil, and A. Draeger. 2009. Association between statin-associated myopathy and skeletal muscle damage. Canadian Medical Association Journal, 181(1-2):E11.

[13] S.L. Molyneux, C.M. Florkowski, A.M. Richards, M. Lever, J.M. Young, and P.M. George. 2009. Coenzyme Q10; an adjunctive therapy for congestive heart failure? Journal of the New Zealand Medical Association, 122:1305.

[14] M.A. Silver, P.H. Langsjoen, S. Szabo, H. Patil, and A. Zelinger. 2004. Effect of atorvastatin on left ventricular diastolic function and ability of coenzyme Q10 to reverse that dysfunction. The American journal of cardiology, 94(10):1306–1310.

[15] C. Silverberg. 2003. Atorvastatin-induced polyneuropathy. Annals of Internal Medicine, 139(9):792.

[16] G. Saher, B. Br¨ugger, C. Lappe-Siefke, W. M¨obius, R. Tozawa, M.C. Wehr, F. Wieland, S. Ishibashi, and K.A. Nave. 2005. High cholesterol level is essential for myelin membrane growth. Nature neuroscience, 8(4):468–475.

[17] J. Tong, P.P. Borbat, J.H. Freed, and Y.K. Shin. 2009. A scissors mechanism for stimulation of SNAREmediated lipid mixing by cholesterol. Proceedings of the National Academy of Sciences, 106(13):5141.

[18] R.S. Tilvis, J.N. Valvanne, T.E. Strandberg and T.A. Miettinen (2011) Prognostic significance of serum cholesterol, lathosterol, and sitosterol in old age; a 17-year population study. Annals of Medicine, Early Online, 1-10.

[19] L.R.Wagstaff, M.W. Mitton, B.M. ARVIK, and P.M. Doraiswamy. 2003. Statin-associated memory loss: analysis of 60 case reports and review of the literature. Pharmacotherapy, 23(7):871–880.

[20] K.A. Weant and K.M. Smith (2005) The Role of Coenzyme Q10 in Heart Failure. Ann Pharmacother. Sep; 39(9), 1522-6.

[21] G. Wainwright, L. Mascitelli, and M.R. Goldstein. 2009. Cholesterol-lowering therapy and cell membranes: stable plaque at the expense of unstable membranes? Arch Med Sci, 5:3.

[22] L. Zhuang, F. Jing, and X.Y. Zhu. 2006. Movie review mining and summarization. In Proceedings of CIKM, pages 43–50. ACM.

[23] Q. Zeng, S. Kogan, N. Ash, RA Greenes, and AA Boxwala. 2002. Characteristics of consumer terminology for health information retrieval. Methods of information in medicine, 41(4):289–298.

[24] P.E. Ziajka and T. Wehmeier. 1998. Peripheral neuropathy and lipid-lowering therapy. Southern medical journal, 91(7):667.