Multigenic condition risk assessment in direct-to-consumer genomic services

Melanie Swan, MBA

Purpose: Gene carrier status and pharmacogenomic data may be detectable from single nucleotide polymorphisms (SNPs), but SNP-based research concerning multigenic common disease such as diabetes, cancers, and cardiovascular disease is an emerging field. The many SNPs and loci that may relate to common disease have not yet been comprehensively identified and understood scientifically. In the interim, direct-to-consumer (DTC) genomic companies have forged ahead in developing composite risk interpretations for multigenic conditions. It is useful to understand how variance in risk interpretation may arise. **Methods:** A comprehensive study was conducted to analyze the 213 conditions covered by the 5 identifiable genome-wide DTC genomic companies, and the total SNPs (401) and loci (224) assessed in the 20 common disease conditions with the greatest overlapping coverage. **Results:** Variance in multigenic condition risk interpretation can be explained by differences in the average lifetime risk assigned to similar underlying populations, the loci and SNPs selected for analysis, and the quantitative risk assignment methodologies used by DTC genomic companies. **Conclusion:** At present, multigenic condition analysis is a complicated process. DTC genomic companies have made laudable efforts to deliver risk predictions, but greater consistency is needed for the long-term validity, utility, and credibility of the sector. *Genet Med* 2010:12(5):279–288.

Key Words: *personal genomes, genetic testing, direct-to-consumer genomics, multigenic risk assessment, personalized medicine*

There are three kinds of direct-to-consumer (DTC) genomic services currently available directly to individuals (Fig. 1): tests for one or a few specific conditions, multiple single nucleotide polymorphism (SNP) risk assessment tests mapping many SNPs to dozens of disease conditions, and whole human genome sequencing assessing hundreds of disease risks. Several companies provide analysis for single diseases or traits: for example, DNA Direct and Matrix Genomics offer tests for conditions such as specific genetic disorders and drug response ranging from $200 to $3500. Paternity, zygosity, nutrigenomics, and matchmaking are other examples of DTC genetic tests ranging from $10 to $1995.

At least 5 companies offer multigenic condition analysis by genotyping 600,000 to 1 million SNPs and analyzing their association with several common conditions. 23andMe, deCODEme, and Navigenics have been on the market since late 2007, and as of November 2009, cover 127, 46, and 28 conditions for $429, $985, and $999, respectively. These 3 services allow consumers to download their raw genotyping data, which can then be reviewed in other genome browsers such as the wiki-based SNPedia,[1] where there are 35 publicly available whole and partial human genomes as of November 2009. Pathway Genomics and Gene Essence have launched more recently, in mid-2009, and cover 71 and 84 conditions for $299 and $1,195 respectively. SeqWright also provides a basic service, genotyping 1 million SNPs for $998. 23andMe, deCODEme, and Pathway Genomics provide 2 services, health and ancestry testing. Not a lot is known about the overall DTC genomic testing market size; however, 23andMe reported having 30,000 active genomes as of October 2009.[2]

So far, there are two companies offering whole human genome sequencing to the public, Knome and Illumina. Knome launched in 2007, initially charging $350,000, and as of November 2009, it is charging $68,500 (or $24,500 for exome sequencing, genotyping the 1–2% of the genome that is estimated to be responsible for protein coding). Illumina launched in June 2009 and unlike the other DTC genomic services, requires a physician's prescription. Illumina charges $48,000 and solely delivers the sequenced data. The Knome service also maps hundreds of SNPs to various conditions via a browser interface and focuses on six conditions in a detailed report that is reviewed with consumers: Alzheimer disease, heart disease, immune system disease, liver disease, macular degeneration, and prostate cancer or breast cancer.

The market for DTC genomic testing has been volatile with varied pricing for similar services and dramatic price drops in the course of a few months or years as indicated in Figure 1. The trend of performance improvements for lower prices is emblematic of a technology industry experiencing continuous cost decreases, measured in life sciences by Carlson Curves,[3] a faster analog of Moore's Law.

In addition to delivering consumer-focused services, DTC genomic companies have been establishing partnerships with physicians and researchers. For example, DNA Direct is partnering with El Camino Hospital, Navigenics is partnering with Scripps and other physician groups, 23andMe is partnering with Palomar Pomerado Health, and deCODEme is partnering with Johns Hopkins and Duke Medicine to conduct studies with 3000 and 1000 patients, respectively.

MATERIALS AND METHODS

Many research efforts regarding genome-wide DTC services have thus far primarily focused on the genotyping results received by one or a few individuals trying some of the available testing services.[4,5] This analysis provides a comprehensive look at all the available genome-wide DTC services, all the conditions assessed in these services, and all the quantitative data provided by DTC genomic companies for multigenic condition risk assessment. The goal is to identify the reasons for variance in risk interpretation.

Publicly available information from DTC genomic company websites[6–10] was examined from August to November 2009.
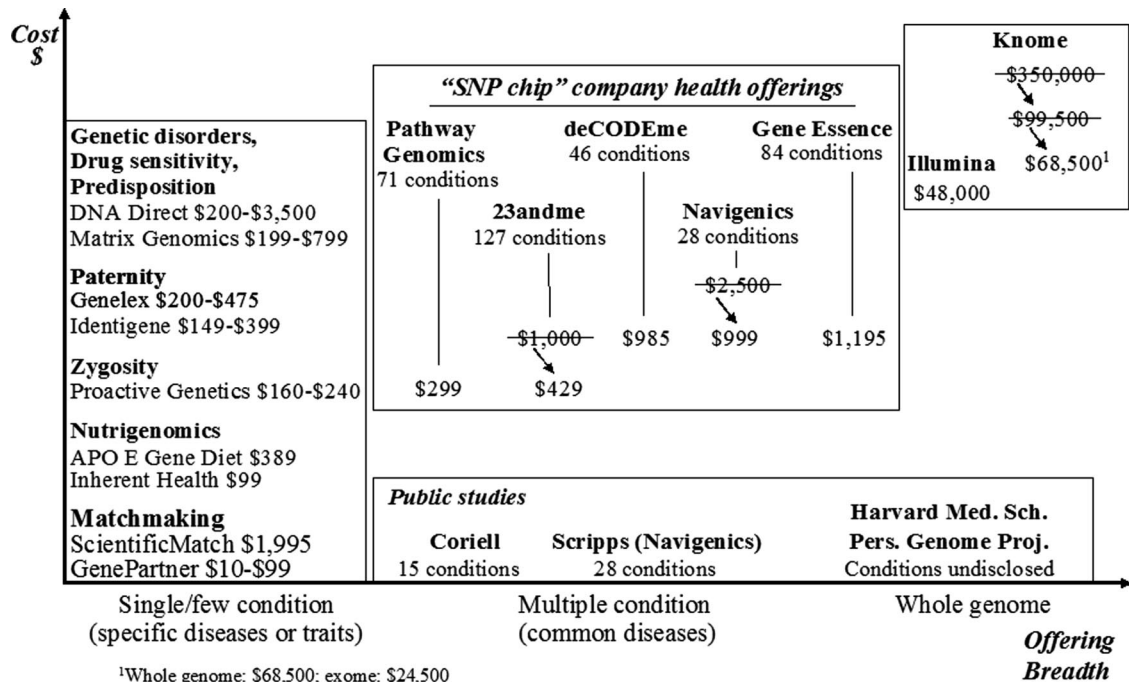
Fig. 1. Direct-to-consumer genomic services as of November 2009. Genomic tests available directly to consumers are organized along the dimensions of cost and breadth of coverage. The types of tests include those for one or a few specific conditions ranging from $10 to $3,500, multiple condition genotyping services ranging from $299 to $1,195, whole human genome sequencing ranging from $48,000 to $68,500, and public research studies.

Supplementary information was gathered at technical conferences and from informal interviews with genomics researchers and industry members from June to October 2009. Analysis was conducted at two levels. First, there was a review of the services and conditions covered by the five identifiable companies in the DTC genomic sector who cover multiple conditions in genome-wide services: 23andMe, deCODEme, Navigenics, Gene Essence, and Pathway Genomics.

Second, a more detailed multigenic condition analysis was completed for all the conditions (20) for which there was overlapping coverage as of November 2009 between three to four DTC genomic companies (23andMe, deCODEme, Navigenics, and Gene Essence; Pathway Genomics does not provide SNP data) and for which quantitative risk assignment data were available. The 20 conditions are all multigenic common diseases: Alzheimer disease, atrial fibrillation, breast cancer, celiac disease, colorectal cancer, Crohn disease, type 1 diabetes, type 2 diabetes, glaucoma, heart attack, lung cancer, lupus, macular degeneration, multiple sclerosis, obesity, prostate cancer, psoriasis, restless legs syndrome, rheumatoid arthritis, and ulcerative colitis.

For each of the 20 conditions, DTC genomic company website information was reviewed, including the average lifetime risk for the disease in a representative population, all the chromosomes, loci, genes, and SNPs associated with a condition, the type and quality of reference studies cited, and the quantitative risk assessment values and methodology. When loci (genomic locations) were not identified definitively by DTC genomic companies, an attempt was made to complete this information using public databases including the National Human Genome Research Institute's online genome-wide association study (GWAS) catalog,[11] the National Center for Biotechnology Information SNP database, dbSNP,[12] and PharmGKB.[13]

Methodologically, specific genes and chromosomal locations were considered to be different loci. For example, the MSMB gene at 10q11.23 was considered to be different from the CTBP2 gene at 10q26.13. Similarly, 12q13.13 was considered to be different from 12q21.1. Regions in the same area were treated as one locus, for example, 8q24 regions 1–3. The locus analysis should be regarded as representative rather than as having the same precision as the SNP analysis. The complete locus analysis is available in Supplemental Digital Content 1, http://links.lww.com/GIM/A95.

RESULTS

Analysis 1: 213 conditions covered by DTC genomic companies

The results from the first analysis concern the services and conditions covered by the five genome-wide DTC genomic companies, 23andMe, deCODEme, Navigenics, Gene Essence, and Pathway Genomics. As of November 2009, the companies review 213 total conditions in the areas of multigenic common disease, single-gene conditions, drug response, and traits. The genotyping platforms used are generally different versions and customizations of chips from Illumina and Affymetrix, and the accuracy of the raw genotyping data has been reported as being consistent between services.[4,14]

Of the 213 total conditions, only 9 are covered by all 5 companies: colorectal cancer, type 2 diabetes, glaucoma, heart attack, lung cancer, multiple sclerosis, obesity, prostate cancer, and rheumatoid arthritis. Fifteen additional conditions are covered by four of the five companies: Alzheimer disease, asthma, atrial fibrillation, breast cancer, celiac disease, Crohn disease,

type 1 diabetes, hemochromatosis, lupus, macular degeneration, osteoarthritis, peripheral arterial disease, psoriasis, restless legs syndrome, and ulcerative colitis. 23andMe includes some less medically oriented conditions such as earwax consistency, odor detection, and hair thickness. At the other end of the spectrum, Gene Essence assesses more detailed medical conditions than the others, for example, restenosis after coronary stenting and renal allograft rejection. A complete list of conditions covered by company is available in Supplemental Digital Content 1, http://links.lww.com/GIM/A95.

Analysis 2: Variance in risk interpretation between DTC genomic services

The results from the second analysis concern multigenic condition risk assessment. DTC genomic companies provide different levels of information. 23andMe, deCODEme, and Navigenics include detailed lists of the SNPs assessed for each condition, research references for the studies that link SNPs to conditions, lifetime risk averages for an underlying population, quantitative risk values assigned to the risk allele, composite risk values for conditions, and information about SNP selection and risk quantification methodologies. Of these three, only Navigenics provides the risk allele (e.g., the "typo" or genotype letter thought to confer higher risk). DeCODEme is the only company readily providing the total number of cases and controls from the underlying reference studies. Gene Essence provides the SNPs assessed for each condition, the risk allele, and the quantitative risk value attributed to the allele, but it does not provide risk composites by condition or research references. For registered customers, but not publicly, Pathway Genomics provides the gene, SNP (one SNP is reviewed per gene), genotype data, risk allele, population frequency, odds ratio, and underlying research study citation. Pathway Genomics does not calculate a composite risk by condition. Knome does not make any of the above information available publicly or upon research request.

Variance in risk predictions for the same conditions among DTC genomic services has been pointed out by those who have had experience with multiple tests. The disparity may have surfaced initially with author David Duncan.[5] National Insti-

tutes of Health Director Francis Collins similarly reported being assessed at high, low, and average risk for the same condition from three DTC genomic companies.[14] A recent study by Ng et al.[4] found that "for seven diseases, 50% or less of the predictions of two companies [23andMe and Navigenics] agreed across five individuals."

On analysis, three reasons for differing risk assessments become clear. First, DTC genomic companies assign different average lifetime risks to the same representative populations for the same conditions. Second, despite some degree of overlap in the assessment of strong-effect markers (markers that have a strong effect on risk), there is considerable variance in the total loci and SNPs selected for analysis. Third, different quantitative outcomes result from differences in the process of assigning values to risk alleles and deriving composite risk values for conditions.

Different average lifetime risk values attributed to the same populations

The companies that cite lifetime risk values, 23andMe, deCODEme, and Navigenics, do not always use the same figures for the average lifetime risk of the overall population. Table 1 lists 11 conditions for which lifetime risk data are available. Some of the biggest differences are for heart attack (ranging 21–42% in men and 7–25% in women), obesity (ranging 34–64% in men and 32–59% in women), type 2 diabetes (ranging 18–30% in women), and Alzheimer disease (ranging 6–9% in men, and 12–17% in women). For some conditions there is similar alignment, although not full agreement, for example, Navigenics, deCODEme, and 23andMe list 17%, 16%, and 18% risk, respectively, for prostate cancer, and 13%, 12%, and 16% for breast cancer. A complete list of the average population risks used in DTC genomic services is available in Supplemental Digital Content 1, http://links.lww.com/GIM/A95.

23andMe does not provide references for lifetime risk averages. DeCODEme and Navigenics both provide their reference sources, and deCODEme has direct links to the citations. The underlying studies cited by deCODEme and Navigenics were reviewed, finding that for 13 of 26 conditions, the numbers in

Table 1 Average lifetime risk for conditions cited by DTC genomic companies as of November 2009

Condition	Male (%)			Female (%)		
	Navigenics	deCODEme	23andMe	Navigenics	deCODEme	23andMe
Alzheimer disease	9	6	n/a	17	12	n/a
Breast cancer	n/a	n/a	n/a	13	12	16
Diabetes, type 2	25	25	24	30	28	18
Glaucoma	1	15	1	2	15	1
Heart attack	42	42	21	25	25	7
Lung cancer	8	17	9	6	12	6
Obesity	34	40	64	32	40	59
Prostate cancer	17	16	18	n/a	n/a	n/a
Psoriasis	4	2	11	4	2	10
Restless legs syndrome	4	7	2	4	13	4
Venous thromboembolism	3	10	12	4	8	10

Sources: http://demo.decodeme.com/summary-report.pdf, http://www.navigenics.com/demo, https://www.23andme.com/you/journal/heartattack/overview/, etc.

the underlying study differed from those cited by deCODEme or Navigenics. This could be partially due to the epidemiological notation in the underlying studies. More detail is available in Supplemental Digital Content 1, http://links.lww.com/GIM/A95. There is less variance between the deCODEme and Navigenics average lifetime risk percentages than with either company and 23andMe. DeCODEme and Navigenics cite the same population studies for 7 of 20 (35%) shared conditions but do not always use the same numbers. For example, the two companies have the same reference sources but different statistics for breast cancer, colon cancer, and type 2 diabetes in women.

Heterogeneous SNPs and loci assessed

The second reason that DTC genomic companies have different risk assessments for the same conditions is that they are evaluating different SNPs and different loci. Table 2 illustrates the number of SNPs and loci assessed in the 20 conditions covered by 3–4 companies as of November 2009. A total of 401 SNPs in ~224 loci are reviewed by DTC genomic companies. Of the 224 loci reviewed by DTC genomic companies, 115 are investigated by only one company and 63 are reviewed by two companies.

For the 12 conditions covered by 4 companies, only 9 SNPs were reviewed by all: for colorectal cancer and prostate cancer, POU5F1P1 rs6983267; for celiac disease, HLA-DQA1 rs2187668; for Crohn disease, PTGER4 rs17234657; for type 2 diabetes, HHEX rs1111875, PPARG rs1801282, and IGF2BP2 rs4402960; for glaucoma, LOXL1 rs2165241; and for multiple sclerosis, IL7R rs6897932. The nine SNPs comprise only 3% of the total SNPs (298) reviewed by all four companies, but 18% of the total loci (28 of 160), supporting the notion that different SNPs in the same location are used in analysis.

The three companies with the most overlap in SNPs and loci are deCODEme, Navigenics, and 23andMe. These three companies review 22 of the same SNPs (8% of 264 total SNPs for the 20 conditions), but 42 of the same loci (22% of 195 total loci). DeCODEme and Navigenics have the most overlap, sharing 38 SNPs and 39 loci in addition to the ones shared with 23andMe, and also have a high degree of concurrence in research references cited. However, despite some overlap be-

Table 2 Number of SNPs and loci assessed in 20 conditions by DTC genomic companies as of November 2009

	1 company		2 companies		3 companies		4 companies		Total	
	SNPs	Loci	SNPs	Loci	SNPs	Loci	SNPs	Loci	SNPs	Loci
Conditions covered by four companies										
Celiac disease	4	2	7	6		1	1	1	12	10
Colorectal cancer	4	4	2	2	2	2	1	1	9	9
Crohn disease	43	12	21	16	4		1	9	69	37
Diabetes, type 2	39	9	10	6	2	1	3	8	54	24
Glaucoma	1						1	1	2	1
Heart attack	16	9	1					1	17	10
Lung cancer	8	3			1			1	9	4
Multiple sclerosis	15	6	2	1		1	1	1	18	9
Obesity	16	11	2					1	18	12
Prostate cancer	35	17	4	3	3	2	1	1	43	23
Restless legs syndrome	9	2	5	1		1		1	14	5
Rheumatoid arthritis	28	11	2	1	3	2		2	33	16
Subtotal	218	86	56	36	15	10	9	28	298	160
Conditions covered by three companies										
Alzheimer disease	1	1	2	1					3	2
Atrial fibrillation	1	1			2	1			3	2
Breast cancer	13	6	3	4	1	2			17	12
Diabetes, type 1	20	5	8	9		1			28	15
Lupus	15	6	3	5					18	11
Macular degeneration	3		6	2		2			9	4
Psoriasis	7	4	1	2	2	2			10	8
Ulcerative colitis	11	6	4	4					15	10
Subtotal	71	29	27	27	5	8			103	64
Total	289	115	83	63	20	18	9	28	401	224

tween each pairing of companies, most SNPs and loci are assessed by only one company.

DTC genomic services may differ among themselves and also with public studies. The Coriell Personalized Medicine Collaborative, for example, has provided data for 6 of the 15 conditions that has been approved by the institute review board for the study: coronary artery disease, type 2 diabetes, iron overload, macular degeneration, melanoma, and prostate cancer. As of November 2009, Coriell only reviews one SNP for each condition. Four Coriell conditions overlap with the 20 analyzed here, and each SNP reviewed by Coriell is also reviewed by at least one DTC genomic company. Specifically, rs16901979 reviewed by Coriell for prostate cancer is also reviewed by Navigenics and Gene Essence, rs10490924 for macular degeneration is reviewed by Navigenics, and rs57754840 for type 2 diabetes and rs1333049 for coronary artery disease are reviewed by Gene Essence. Rather than generating quantitative risk scores, Coriell, similar to Pathway Genomics, provides a qualitative estimate of individual risk based on self-reported health and lifestyle factors along with genetic results.

Heterogeneous quantitative risk assignment methodologies

The third reason that DTC genomic companies have different risk interpretations for the same conditions is that they have different risk assessment methodologies, both in generating the quantitative value assigned to the risk allele and in the composite scoring for overall conditions.

The first step is assigning quantitative values to the risk allele. Table 3 depicts the quantitative values assigned to genotypes for the nine SNPs that are investigated by all four DTC genomic companies. Data are displayed for similar genotype

values as available. Values may differ, often for the same genotype and cited research study. For example, the "CC" genotype values for type 2 diabetes SNP rs1801282 vary from 1.02 to 1.53, and the "CC" genotype values for multiple sclerosis SNP rs6897932 vary from 1.08 to 1.46. Risk assessments can vary in magnitude and sometimes directionally (i.e., for the same genotype, some services assign a value >1.0 and others a value <1.0).

One reason the numbers are not strictly comparable is methodology. DeCODEme is distinct from the others in converting the odds ratios for each marker to risk compared with the general population before taking the product of the markers.[15] The study of Ng et al.,[4] which looked at Navigenics and 23andMe, found odds ratios to be similar for identical and correlated markers; however, a broader comparison of companies and markers in this analysis shows that relative risk values may vary for the same marker. Ng et al.[4] found that 80% of reported relative risk in the DTC test results of their study was between 0.5 and 1.5, indicating only a modest risk when compared with the average population. This could suggest that relative risk values such as 1.08 and 1.46 (Table 3) are not meaningfully different, or meaningfully useful, because they are within the range of average risk.

The genotype at a particular SNP consists of two letters, one inherited from each parent. However, only two possible alleles (variants) are possible at any SNP, for example, different permutations of "A" and "G," or "C" and "T." One of the possible letters or nucleotide bases at each SNP is the normal allele, and the other is the risk allele.

For most SNPs, DTC genomic companies are consistent in the stated or implied risk allele. However, there are two cases of inconsistency noted in Table 3. First, in Celiac disease, four companies cite the SNP rs2187668 with the same supporting

Table 3 Risk allele quantification by DTC genomic companies as of November 2009

Condition and SNP	Risk Allele		Genotype example and assigned risk value							
	N	GE	D		N		2		GE	
Celiac disease										
rs2187668 (HLA-DQA1)	T	A	GG	0.30	CC	1.00	CC	0.48	GG	7.00
Colorectal and prostate cancer										
rs6983267 (POU5F1P1)—colorectal	G	G	GT	0.99	GT	1.04	GG	1.16	GT	1.30
rs6983267 (POU5F1P1)—prostate	G	G	GT	0.99	GT	1.26	GG	1.26	GT	1.40
Crohn disease										
rs17234657 (PTGER4)	G	G	TT	0.96	TT	1.00	TT	0.95	GG	2.30
Type 2 diabetes										
rs1111875 (HHEX)	C	G	GG	1.14	CC	1.20	CT	0.98	AA	1.10
rs1801282 (PPARG)	C	C	CG	0.90	CC	1.53	CC	1.02	CC	1.10
rs4402960 (IGF2BP2)	T	T	GG	0.92	GT	1.16	GT	1.05	GG	1.20
Glaucoma										
rs2165241 (LOXL1)	T	T	CC	0.03	CC	1.00	TT	2.94	TT	2.00
Multiple sclerosis										
rs6897932 (IL7R)	C	C	CC	1.08	CT	1.46	CC	1.08	CC	1.20

D, deCODEme; N, Navigenics; 2, 23andMe; GE, Gene Essence.

reference study. Navigenics cites "T" and Gene Essence cites "A." The underlying study cites "A" as the risk allele[16] and another study listed in the online GWAS catalog confirms this finding. In demonstration examples, deCODEme and Gene Essence suggest "GG" as a possible genotype value for this SNP, and Navigenics and 23andMe, "CC." Presumably not all four bases would be acceptable genotype values, and only one base would be the risk allele.

A similar situation exists for the type 2 diabetes SNP rs1111875. Navigenics cites "C" as the risk allele and Gene Essence cites "G." The underlying reference cites "C" as the risk allele and "C" and "T" as the possible genotype values.[17] The online GWAS catalog cites this study and two others with "C" as the risk allele, and one with "G" as the risk allele. In demonstration examples, deCODEme suggests a possible genotype value as "GG," and Gene Essence, "AA." Again, not all four bases should be acceptable genotype values, and there should be only one risk allele.

These examples of inconsistency underline the challenges of obtaining the correct risk allele and genotype data. These discrepancies could be due to inverse coding in the DNA genotyping process, different coding in DTC genotyping platforms, SNP surrogates, or other issues. The point is that the quantitative detail is not readily comparable, even for the nine SNPs that are analyzed by four DTC genomic companies.

The second step of risk quantification is generating the calculations of composite risk by condition. Three companies, 23andMe, deCODEme, and Navigenics, provide a composite risk score for each condition using a multiplicative technique but have different ways of arriving at the final output. 23andMe takes the product of all SNP relative risk values and multiplies this value by the average population risk for the condition to generate an estimate of the individual's absolute lifetime risk. As mentioned above, deCODEme applies each SNPs risk to the population and then takes the product. Navigenics generates an interim "genetic composite index" number,[18] which is not possible to reproduce because of a proprietary analysis method. Knome uses a weighted average methodology to assign composite risks for overall conditions.

DISCUSSION

This analysis found that multigenic condition risk interpretation may vary between DTC genomic services due to differences in three factors: the average lifetime risk values assigned to the same underlying populations, the SNPs and loci selected for analysis, and the quantitative risk assignment methodologies used.

Average lifetime risk

DTC genomic companies may assign different average lifetime risks to the same underlying populations, because they have selected different research studies as being the most representative and reliable. There are considerable differences in the supporting epidemiologic literature due to the general challenge of predicting lifetime risk for conditions. Specifically, there are differences in phenotype definitions, and the application of the same phenotype definitions may result in different estimates of lifetime risk. Widely agreed upon figures do not exist for most conditions. DTC genomic companies are aware of this issue, and ongoing attempts are being made to increase consistency in identifying and selecting the most accurate references for average lifetime risk.[15,19]

For cancers, deCODEme and Navigenics cite the SEER Cancer Statistics Review, 1975–2003 of the U.S. National Cancer Institute, which provides risk percentages generalized to populations as a whole. For other conditions, the cited studies may apply to certain age groups or populations which DTC genomic companies have often extrapolated to apply to the population as a whole. DeCODEme and Navigenics apply one average lifetime risk value; 23andMe provides different risk assessments by age tier, which can be another source of variance in lifetime risk values among DTC genomic services. Ng et al.[4] found that different lifetime risk values explained one third of the difference in results for five individuals using the 23andMe and Navigenics services. A further challenge in applying genomics to individuals is that, to date, most of the underlying research studies have been carried out in populations of Western European descent. However, researchers are starting to expand genome-wide association studies to other populations. For example, a study in Asian Indian Sikhs found a significant association for four of nine GWAS SNPs for type 2 diabetes, including a link between increased low-density lipoprotein (LDL) cholesterol and an infrequently documented SNP, rs10885409, in the TCF7L2 gene.[20]

SNP selection

Just as DTC genomic companies may choose different studies to obtain lifetime risk averages for the underlying population, they may choose different studies to select SNPs. SNP selection methodology is provided by 23andMe,[21] deCODEme,[22] and Navigenics.[23] Methodologies are rigorous but slightly different for each firm and contain discretionary elements. All require genome-wide significance that is validated in large replicated studies and reported in peer-reviewed journals. Replicated studies are the most important criterion for each company. DeCODEme and 23andMe look for studies that are replicated in at least one independent population, and Navigenics requires replication in at least two other independent populations. Well-powered statistical studies are another important factor: for example, 23andMe requires sample sizes of 1,000 cases and 1,000 controls. The third most important criterion is that studies must be published in respected scientific journals. Looking at the underlying studies cited by DTC genomic companies confirms that the methodologies seem to be followed.

In SNP selection, Navigenics is the most conservative, citing particularly the low replicability of early GWAS,[24] followed by deCODEme and then 23andMe. For the 20 conditions in this analysis, on average, 75% of Navigenics' references are also cited by deCODEme. On average, 18% of the references cited by 23andMe are cited by all three companies, and 35% of the references cited by 23andMe are cited by either Navigenics or deCODEme. So there is some agreement as to the most important reference studies; 19 of 253 (8%) unique reference studies are cited by all three companies. The most cited research studies often correspond to the most cited SNPs. One reason that 23andMe covers more conditions (127 unique conditions versus 28 and 46 for Navigenics and deCODEme respectively) and has more references (203 vs. 55 and 76 for Navigenics and deCODEme, respectively, for the 20 conditions analyzed here) is that 23andMe provides two tiers of information with different vetting standards, established and preliminary. Established research is presented as Clinical Reports (47 conditions) and preliminary research as Research Reports (86 conditions) for findings that have not been replicated in large studies.[25]

DTC genomic companies are quick to point out that they are aware that different companies look at different SNPs. For the 20 conditions in this analysis, deCODEme reviews 179 SNPs at 163 loci, Navigenics reviews 120 SNPs at 108 loci, and 23andMe reviews 78 SNPs at 64 loci. DTC genomic companies

explain that SNP variance occurs partially due to differences in the genotyping process but is mostly due to differences in SNP selection. In the genotyping process, different SNPs may be available on different platforms, different chips may assign different rsID numbers to the same SNPs, and there may be other technical issues. In SNP selection, DTC genomic companies are looking at a wide range of underlying studies and applying different selection criteria. By evaluating loci, this analysis found that about half of SNP variation may be explained by surrogate SNPs (different SNPs at the same locus). However, because there is little overlap in the loci reviewed by DTC genomic companies (<20% of loci are evaluated by more than two firms), selection criteria is the main reason that SNPs and loci vary.

Quantitative risk assignment

In both stages of quantitative risk assessment, assigning risk values to individual SNPs and generating composite risk values for conditions, DTC genomic companies vary again because they are looking at different research studies and using different methodologies. The risk value assigned to each SNP is based on the odds ratios in the underlying research studies. Odds ratios are derived from the frequency distribution of the different possible genotype values in a population. Although the same reference study may be used by several companies, the assigned odds ratios may vary, because each company has different risk methodologies and calibrations.

To generate composite risk scores for conditions, DTC genomic companies are using a multiplicative technique from epidemiology, taking the product of multiple markers. The justification given is that "no other model fits the data better based on the large datasets available for common variants conferring risk to common disease."[15] As the scientific understanding of multigenic common disease matures, it could be that more sophisticated models of risk analysis evolve. Some of the shortcomings of the product model could be that it does not allocate appropriate weight to strong-effect SNPs, overallocates weight to high odds ratio but weak-effect SNPs, and could dilute or change risk over time as more SNPs are found to be associated with conditions. For example, deCODEme doubled the number of SNPs assessed for prostate cancer from 13 to 25 in September 2009. This shift changed the risk in the demonstration example from 12%, lower than the average population risk of 16%, to an above average risk of 17%. At least one study has noticed similar potential shifts in the case of risk prediction updating and explored this together with patient reaction.[26] Even if widely agreed upon sets of core markers for conditions are established, it is possible that subsequently discovered markers could change condition interpretations. Later-discovered markers could be weaker signals that do not dilute or reverse the main markers; they could also be rarer but stronger or protective.

Another potential shortcoming of the product model is that perhaps many of the SNP risk values are too small to be useful or statistically valid and should not be included.[27] Threshold levels could be used, only including SNPs with relative risk above 1.5 in risk prediction, for example, as Ng et al. have contemplated. Scientific norms could be developed over time to investigate and validate different risk assessment methodologies such as the product method, weighted averages, cluster analysis, exclusion filters, and other techniques. It has also been suggested that odds ratios and *P* values may not be the best measures for calculating risk prediction from genetic profiles.[28]

Regulatory reasons may influence the degree of publicly disclosed information and risk methodologies used in DTC

genomic services. The U.S. Food and Drug Administration considers DTC genomic tests to be laboratory-developed tests that are medical devices.[29] The U.S. Food and Drug Administration regulates medical devices based on risk and so far has applied enforcement discretion to most laboratory-developed tests, including DTC genomic tests, not requiring premarket clearance or approval.[29]

Suggestions to improve DTC genomic services

The Ng et al.[4] opinion article in *Nature* provides an analysis of the DTC genomic testing results for five individuals from 23andMe and Navigenics and includes nine recommendations. The authors suggest that DTC genomic companies report the genetic contribution of the markers tested (similar to the environmental versus genetic contribution of disease, this would be the proportion of the genetic contribution that can be explained by known markers), focus on high risk predictions (relative risks over 1.5 for example), refrain from using surrogate markers, test a broader range of pharmacogenomic markers, and agree on strong-effect markers. The authors further suggest that the DTC genomics research community investigate consumer behavioral outcomes, conduct prospective studies to gauge clinical validity, replicate disease associated markers in other ethnicities, and extend genotyping to whole human sequencing when possible.

23andMe and Navigenics responded to the opinion article in an open letter,[30] generally agreeing with the recommendations. The companies discuss how they are implementing some of the suggestions (for example, reducing the use of surrogates and monitoring consumer behavior with regard to health outcomes), and what they found challenging or problematic about other suggestions (for example, applying population level measures to individuals and establishing guidance for meaningful levels of risk and the genetic contribution of known markers). 23andMe and Navigenics claim that for the most part they are including the same strong-effect markers, which this analysis confirms. However, in addition to the core markers, both companies review many additional markers, which together with different risk quantification methodologies and different average lifetime risks attributed to the underlying population can still lead to different overall interpretations for the same conditions. Using strong-effect markers together with other markers is not unique to DTC genomic tests. For example, 18 SNPs were included in a recent type 2 diabetes study,[26] 11 of which were reviewed by at least one DTC genomic company. Three SNPs—HHEX rs1111875, PPARG rs1801282, and IGF2BP2 rs4402960—were analyzed by four companies, indicating some degree of consensus on strong-effect markers.

All the above are helpful recommendations. Development of industry standards is an obvious step for improving consistency in DTC genomic testing, and there have been some steps toward this end.[15,19] However, there is still room for improvement and implementation may be challenging.[31] Specifically, industry standards efforts could work toward establishing a list of third-party curated strong-effect markers that could be expected in any high-quality genome-wide service, a list of the agreed upon key studies by condition, and agreement regarding odds ratios, risk alleles, lifetime risk values, and quantitative risk assessment methodologies. Consistency in DTC genomic testing does not necessarily mean that the tests are accurate, and these tests should be validated.

One of the biggest reasons for variance in DTC genomic test results is the dynamic nature of the underlying data and lack of standards in applying these data. For example, lifetime population risk averages and biomarkers associated with health conditions can shift with emerging research. DTC genomic com-

panies and other research community members are starting to work on these issues. The role of government agencies and other third-party validators in providing leadership and guidance in genomic medicine has been suggested.[30,32] To address the broad issue of the clinical integration of genomics, the Centers for Disease Control and Prevention and the National Institutes of Health created a new initiative in July 2009, the Genomic Applications in Practice and Prevention Network (GAPPNet).[33] The first meeting of GAPPNet was held in October 2009 to focus on accelerating translational genomics research.[34]

There are other efforts to standardize biomarkers used in DTC genomic tests. One project to establish a centralized list of high-quality GWAS citations is the online GWAS catalog. It is a National Human Genome Research Institute curated database of selected genome-wide association studies with SNP-trait associations. As of November 2009, the database contained 440 publications and 2,065 SNPs. The catalog can be searched by journal, author, disease, chromosomal region, gene, SNP, and *P* value threshold. However, information such as a list of core markers by condition is not immediately clear from search results. For example, 52 "strongest SNPs" are listed for type 2 diabetes. Standards for genetic association studies are also under development. There are some proposals to address topics such as increasing data volumes (there are at least 34,000 published genetic association articles at present), the appropriateness of epidemiological models, and the interaction of genetic and environmental factors in disease.[35,36]

Establishing standards is a contemporary issue in genomics outside of the United States as well. The Human Genetics Commission of the United Kingdom has proposed "A Common Framework of Principles for direct-to-consumer genetic testing services."[37] Relevant to the issues in this article, there is a general recommendation that only clinically validated genetic variants published in peer-reviewed journals and standard statistical methodologies should be used in genetic testing, and that the above should be made available for scrutiny.[37] 23andMe, Navigenics, and deCODEme already provide this information, and more detailed standards would need to be established for less discrepancy in risk results interpretations.

In addition to working together with research community-wide initiatives, DTC genomic companies could consider some of the following direct actions. A key point regarding DTC genomic tests and test results is that they are not accurately and usefully comparable by consumers. There are obvious differences between the conditions covered and the SNPs and research references that are used by each company. In addition, tests that seem to be comparable between companies may not be because of the use of surrogate SNPs and different relative risk numbers (deCODEme's being different from those of the others). Even seemingly straightforward conditions like warfarin sensitivity may yield different results.[38] DeCODEme and 23andMe review the same three SNPs but apply different metabolism assessment methodologies.

In other business sectors, service providers or third parties often have web-based side-by-side comparisons explaining how services differ and this approach could be helpful in DTC genomic services too. DeCODEme provides product comparisons,[39] but the information could be more comprehensive and directly address the different potential reasons for variance in risk interpretation between services. As of December 15, 2009, deCODEme is also providing the ability to upload genotyping data from 23andMe to view through deCODEme's annotation interface.[39]

The role of trained health professionals such as physicians, medical geneticists, genetic counselors, private clinicians, and health advisors is likely to be critical to a wider implementation

of personalized genomics. DTC genomic companies could increase their partnerships with the traditional health care system to facilitate the use of personalized genomics. These companies could also form partnerships with other preventive medicine providers, for example, personal health management companies like Keas and Omicia, health data integration portals like Google Health and Microsoft HealthVault, electronic medical record providers like ActiveHealth and Patients Know Best, and health social networking communities like PatientsLikeMe and CureTogether. DTC genomic companies could also consider providing application programming interfaces and encouraging application development to integrate genomic data with other health data (i.e., phenotypic biomarkers from blood tests and other analyses, family history, etc.) into a unified health picture for both consumer and institutional audiences.

A final suggestion for consumer genomic companies would be to help accelerate the public biobank efforts under development. DTC genomic companies could provide interpretative interfaces and help with data structuring, management, and security for biobanks and other large studies. Biobanks could be large scale, on the order of millions of participants for multiple populations, with genotyping and later sequencing data collected together with ongoing phenotypic biomarker measures for longitudinal study. The rapid pace of data creation at multiple levels, including raw genotyping and sequencing data, together with GWAS and other studies, could lead to new classes of medical discovery based on large datasets. One example of this has already occurred in the artificial intelligence field as Google has been making continuous progress in natural-language-related machine learning by applying statistical methods to the large datasets that have arisen on the web.[40]

Customer-contributed databases are starting to be useful to DTC genomic companies for conducting research and could provide an important and complementary nexus of discovery to traditional academic and public health research laboratories. With 30,000 active genomes, an order of magnitude greater than the number of participants in traditional genetic variation studies such as the 1,000 Genomes Project, 23andMe has already been able to generate some interesting findings.[41–43] DeCODE Genetics, affiliated with deCODEme, has 450,000 genomes in their database, mostly from well-phenotyped disease collections in Iceland and other countries, and has used these data as the basis for numerous scientific publications. Data-rich preaggregated genomic communities searchable by pharmaceutical companies, researchers, entrepreneurs, patients, and other interested parties could facilitate medical research. At last, there could be a focus on the ultra-specific cohorts and orphan diseases that were unfeasible to reach previously and will likely be required in an era of personalized medicine.

Concerns of DTC genomic testing

The validity of medical genomics in common disease prediction, diagnosis, and prevention is in the early stages of development. Even if disease risks are found to be accurate, they may not be clinically useful at present and may comprise only a small portion of the overall health picture because of the influence of unidentified genetic factors, and gene-gene and/or gene-environment interactions. Also, DTC genomic tests may omit testing for variants in patented genes that are central to disease risk analysis. Another concern is that individuals reviewing test results without assistance from a genetic counselor may misunderstand or misinterpret the findings. Additionally, consumers may worry about the misuse of their genetic information. Also debated is whether risk prediction and prevention screening are helpful and cost effective.

Benefits of DTC genomic testing

Despite the risks and challenges of DTC genomic testing, the nascent field is demonstrating some promising medical benefits, such as pharmacogenomic testing to customize drug selection (clopidogrel efficacy, for example)[44] and dosing (warfarin, for example).[45] DTC genomic testing may also motivate consumers to self-manage their risk by improving behaviors (e.g., exercising to reduce the risk of heart disease) or seeking medical advice that could lead to the appropriate use of more aggressive medical testing (e.g., mammograms or colonoscopies beginning at an early age). Studies indicate that consumers have reported actual or intended behavioral changes as a result of DTC genomic testing,[46,47] even in cases such as Alzheimer disease where there is no clinical treatment at present.[48,49] Physicians too have reported changing some aspect of patient care after reviewing consumer-provided DTC genomic test results.[50]

CONCLUSION

DTC genomics is an emerging field with benefits, limitations, and potential risks. This analysis reviewed five companies offering genome-wide DTC testing. Risk results interpretations were found to vary for three reasons. First, different average lifetime risks for the same underlying populations are used. Second, different criteria are used in the selection of research studies, which leads to the use of different SNPs and loci by each company. Third, different quantitative risk assignment methodologies are used.

At present, multigenic condition risk assessment is a complicated endeavor with different discretionary details and multiple steps. Key factors such as lifetime population risk averages, genetic association study methodologies, core lists of strong markers, and translational research definitively linking genetic variants to clinical validity are not standardized. It is no surprise that there is variance in DTC genomic test interpretations. In the short-term, DTC genomic companies could create greater alignment in their analysis and interpretation and provide more explanation as to why and how services differ. Specifically, DTC genomic companies, possibly together with third parties, could establish a list of strong markers that could be expected in any high-quality genome-wide service, a list of agreed upon key studies by condition and agreement on lifetime risk values, odds ratios, risk alleles, and quantitative risk assessment methodologies. In the long-term, greater scientific agreement could occur naturally in multigenic condition analysis, especially with regard to core lists of strong markers, clinical validity data, and risk interpretation methodologies.

Simply because there is diversity in multigenic condition risk assessment does not mean that DTC genomics is not useful. The efforts of DTC genomic companies are laudable. Firms have undertaken a tremendous effort to collect, interpret, synthesize, and display large volumes of genomic information in a comprehensive and useful manner to generate overall predictions for disease conditions. It is hoped that refinements by DTC genomic companies and the genomics research community will continue over time and that greater validity, utility, consistency, and credibility in personalized genomics may be realized.

REFERENCES

1. SNPedia. Available at: http://www.snpedia.com/index.php/Genomes. Accessed November 30, 2009.
2. Dolan B. TEDMED: 23andMe has 30,000 "active" genomes, launching "Relative Finder." The mobihealthnews blog. October 28, 2009.
3. Carslon R, 2009. Cost per base of DNA sequencing and synthesis. Available at: http://www.synthesis.cc/carlson_DNA_cost_curves_Sept_09.jpg. Accessed November 30, 2009.
4. Ng PC, Murray SS, Levy S, Venter JC. An agenda for personalized medicine. *Nature* 2009;461:724–726.
5. Duncan DE. Experimental man: What one man's body reveals about his future, your health, and our toxic world. New York: John Wiley & Sons, Inc., 2009:45–48.
6. deCODEme. Available at: http://www.decodeme.com. Accessed November 30, 2009.
7. Navigenics. Available at: http://www.navigenics.com. Accessed November 30, 2009.
8. 23andMe. Available at: http://www.23andme.com. Accessed November 30, 2009.
9. Pathway Genomics. Available at: http://www.pathwaygenomics.com. Accessed November 30, 2009.
10. Gene Essence. Available at: http://www.geneessence.com. Accessed November 30, 2009.
11. Hindorff LA, Sethupathy P, Junkins HA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA* 2009;106:9362–9367. Online GWAS catalog available at: http://www.genome.gov/gwastudies. Accessed November 30, 2009.
12. Sherry ST, Ward MH, Kholodov M, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 2001;29:308–311. Online dbSNP database available at: http://www.ncbi.nlm.nih.gov/SNP. Accessed November 30, 2009.
13. Sangkuhl K, Berlin DS, Altman RB, et al. PharmGKB: understanding the effects of individual genetic variants. *Drug Metab Rev* 2008;40:539–551. Online PharmGKB database available at: http://www.pharmgkb.org. Accessed November 30, 2009.
14. Singer E. A renowned geneticist analyzes consumer tests. The MIT Technology Review blog. June 15, 2009.
15. Personalized Medicine Coalition, 2008. Personal genomics and industry standards: scientific validity. Available at: http://www.personalizedmedicinecoalition.org/objects/pdfs/PMC%20personalgenomicsSci%20Valid15dec08.pdf. Accessed November 30, 2009.
16. van Heel DA, Franke L, Hunt KA, et al. A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. *Nat Genet* 2007;39:827–829.
17. Scott LJ, Mohlke KL, Bonnycastle LL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 2007;316:1341–1345. Supporting Online Material: 31.
18. The science behind the Navigenics service. Available at: http://www.navigenics.com/static/pdf/Navigenics-TheScience.pdf. Accessed November 30, 2009.
19. Khoury MJ, McBride CM, Schully SD, et al. The Scientific Foundation for personal genomics: recommendations from a National Institutes of Health–Centers for Disease Control and Prevention Multidisciplinary Workshop. *Genet Med* 2009;11:559–567.
20. Sanghera DK, Ortega L, Han S, et al. Impact of nine common type 2 diabetes risk polymorphisms in Asian Indian Sikhs: PPARG2 (Pro12Ala), IGF2BP2, TCF7L2 and FTO variants confer a significant risk. *BMC Med Genet* 2008;9:59.
21. Hsu A, Naughton B, 2008. 23andMe white paper 23–03 guidelines on vetting genetic associations. Available at: https://www.23andme.com/res/pdf/23–03_Vetting_Genetic_Associations.pdf. Accessed November 30, 2009.
22. deCODEme. Selection of risk associated genetic variants (SNPs). Available at: http://demo.decodeme.com/health-watch/snp-selection. Accessed November 30, 2009.
23. Navigenics. Applying preventive genomic medicine in clinical practice. Available at: http://www.navigenics.com/static/pdf/physician/physician-whitepaper.pdf. Accessed November 30, 2009.
24. Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. A comprehensive review of genetic association studies. *Genet Med* 2002;4:45–61.
25. 23andMe. Health reports: complete list. Available at: https://www.23andme.com/health/all. Accessed November 30, 2009.
26. Mihaescu R, van Hoek M, Sijbrands EJ, et al. Evaluation of risk prediction updates from commercial genome-wide scans. *Genet Med* 2009;11:588–594.
27. Janssens AC, Gwinn M, Bradley LA, Oostra BA, van Duijn CM, Khoury MJ. A critical appraisal of the scientific basis of commercial genomic profiles used to assess health risks and personalize health interventions. *Am J Hum Genet* 2008;82:593–599.
28. Kraft P, Wacholder S, Cornelis MC, et al. Beyond odds ratios—communicating disease risk based on genetic profiles. *Nat Rev Genet* 2009;10:264–269.
29. Harper C. The FDA and the regulation of direct-to-consumer genetic testing. Presented at the direct-to-consumer genetic testing: a cross-academies workshop, National Academies' Keck Center, Washington, DC, August 31-September 1, 2009.

30. Hsu A. 23andMe & Navigenics' open letter to Nature. The Spittoon blog. November 18, 2009.

31. Wechsler P. Google-backed DNA testers don't always agree on genetic profile. Available at: Bloomberg.com. Accessed October 7, 2009.

32. Avey L. Untitled. The Life and Times of Lilly Mendel blog. November 6, 2009.

33. Khoury MJ, Feero WG, Reyes M, et al. The genomic applications in practice and prevention network. *Genet Med* 2009;11:488–494. GAPPNet website available at: http://www.cdc.gov/genomics/translation/GAPPNet. Accessed November 30, 2009.

34. Genomic Applications in Practice and Prevention Network (GAPPNet). Inaugural Meeting of the Genomic Applications in Practice and Prevention Network (GAPPNet): a national collaboration for realizing the promise of genomics in healthcare and disease prevention. University of Michigan School of Public Health, Ann Arbor, MI, October 29–30, 2009. Meeting agenda and presentations available at: http://www.cdc.gov/genomics/translation/GAPPNet/meeting/agenda.htm. Accessed November 30, 2009.

35. Little J, Higgins JP, Ioannidis JP, et al. STrengthening the REporting of Genetic Association studies (STREGA)—an extension of the STROBE statement. *Eur J Clin Invest* 2009;39:247–266.

36. Ioannidis JP, Boffetta P, Little J, et al. Assessment of cumulative evidence on genetic associations: interim guidelines. *Int J Epidemiol* 2008;37:120–132.

37. U.K. Human Genetics Commission. A Common Framework of Principles for direct-to-consumer genetic testing services (p 13). Available at: http://www.hgc.gov.uk/Client/document.asp?DocId=214&CAtegoryId=3. Accessed November 30, 2009.

38. Shmigelsky G. dnaSNIPs: a comparison study into consumer genetic services. Available at: http://www.dnasnips.com. Accessed November 30, 2009.

39. DeCODEme. Feel free to compare. Available at: http://www.decodeme.com/product-comparison and This Winter, Migrate to deCODEme! Available at: http://www.decodeme.com/data-upload. Accessed November 30, 2009.

40. Halevy A, Norvig P, Pereira F. The unreasonable effectiveness of data. *IEEE Intelligent Systems* 2009;24:8–12. Available at: http://www.computer.org/portal/cms_docs_intelligent/intelligent/homepage/2009/x2 exp.pdf. Accessed November 30, 2009.

41. Eriksson N, Macpherson JM, Tung J, et al. Web-based, participant-driven studies yield novel genetic associations for common traits. Presented at the American Society of Human Genetics 59th Annual Meeting, Hawaii Convention Center, Honolulu, HI, October 20–24, 2009.

42. Hon L, Henn BM, Macpherson JM, et al. Discovering distant relatives within a diverse set of populations using DNA segments identical by descent. Presented at the American Society of Human Genetics 59th Annual Meeting, Hawaii Convention Center, Honolulu, HI, October 20–24, 2009, Poster: 596/W:254.

43. Hon L, Henn BM, Macpherson JM, et al. Fine-scale population structure in worldwide ethnic populations as revealed by identical by descent segments. Presented at the American Society of Human Genetics 59th Annual Meeting. Hawaii Convention Center, Honolulu, HI, October 20–24, 2009, Poster: 642W:300.

44. Shuldiner AR, O'Connell JR, Bliden KP, et al. Association of cytochrome P450 2C19 genotype with the antiplatelet effect and clinical efficacy of clopidogrel therapy. *JAMA* 2009;302:849–857.

45. International Warfarin Pharmacogenetics Consortium, Klein TE, Altman RB, et al. Estimation of the warfarin dose with clinical and pharmacogenetic data. *N Engl J Med* 2009;360:753–764.

46. McGuire AL, Diaz CM, Wang T, Hilsenbeck SG. Social networkers' attitudes toward direct-to consumer personal genome testing. *Am J Bioeth* 2009;9:3–10.

47. McGowan M, Lambrix M, Fishman J. Motivations and moral imperatives of early adopters of personal genome services. Presented at the Genetics and Ethics Conference, Village at Breckenridge Conference Center, Breckenridge, CO, July 24–25, 2009:13, 19.

48. Chao S, Roberts JS, Marteau TM, Silliman R, Cupples LA, Green RC. Health behavior changes after genetic risk assessment for Alzheimer disease: the REVEAL Study. *Alzheimer Dis Assoc Disord* 2008;22:94–97.

49. Green RC, Roberts JS. The impact and utility of personalized genomic information: insights from the REVEAL Study (p 30). Presented at the Consumer Genetics Show, Hynes Convention Center, Boston, MA, June 9–11, 2009. Available at: http://cancercontrol.cancer.gov/od/phg/presentations/Green.pdf. Accessed November 30, 2009.

50. Kolor K, Liu T, St Pierre J, Khoury MJ. Health care provider and consumer awareness, perceptions, and use of direct-to-consumer personal genomic tests, United States, 2008. *Genet Med* 2009;11:595.