Original Paper

Public Health Genomics

Public Health Genomics 2015;18:216–224
DOI: 10.1159/000431250

Received: January 12, 2015
Accepted: May 8, 2015
Published online: June 16, 2015

How Well Do Customers of Direct-to-Consumer Personal Genomic Testing Services Comprehend Genetic Test Results? Findings from the Impact of Personal Genomics Study

Jenny E. Ostergren[a] Michele C. Gornick[a, e] Deanna Alexis Carere[f, g] Sarah S. Kalia[i] Wendy R. Uhlmann[a, c, d] Mack T. Ruffin[a, b] Joanna L. Mountain[j] Robert C. Green[g, h] J. Scott Roberts[a] for the PGen Study Group

[a]School of Public Health, [b]School of Medicine and Departments of [c]Human Genetics and [d]Internal Medicine, University of Michigan, and [e]Department of Veterans Affairs Health Services Research and Development, Ann Arbor, Mich., [f]Harvard T.H. Chan School of Public Health, [g]Brigham and Women's Hospital, and [h]Harvard Medical School, Boston, Mass., [i]Icahn School of Medicine at Mount Sinai, New York, N.Y., and [j]23andMe Inc., Mountain View, Calif., USA

Key Words
Commercial genetics · Direct-to-consumer genetic testing · Personal genomic testing · Public health policy · Risk comprehension

Abstract
Aim: To assess customer comprehension of health-related personal genomic testing (PGT) results. ***Methods:*** We presented sample reports of genetic results and examined responses to comprehension questions in 1,030 PGT customers (mean age: 46.7 years; 59.9% female; 79.0% college graduates; 14.9% non-White; 4.7% of Hispanic/Latino ethnicity). Sample reports presented a genetic risk for Alzheimer's disease and type 2 diabetes, carrier screening summary results for >30 conditions, results for phenylketonuria and cystic fibrosis, and drug response results for a statin drug. Logistic regression was used to identify correlates of participant comprehension. ***Results:*** Participants exhibited high overall comprehension (mean score: 79.1% correct). The highest comprehension (range: 81.1–97.4% correct) was observed in the statin drug response and carrier screening summary results, and lower comprehension (range: 63.6–74.8% correct) on specific carrier screening results. Higher levels of numeracy, genetic knowledge, and education were significantly associated with greater comprehension. Older age (≥60 years) was associated with lower comprehension scores. ***Conclusions:*** Most customers accurately interpreted the health implications of PGT results; however, comprehension varied by demographic characteristics, numeracy and genetic knowledge, and types and format of the genetic information presented. Results suggest a need to tailor the presentation of PGT results by test type and customer characteristics.

Introduction

Direct-to-consumer personal genomic testing (PGT) was introduced in 2007 and allows customers to obtain personalized genetic risk information for a variety of com-

For a list of the PGen Study members, see Appendix.

E-Mail karger@karger.com
www.karger.com/phg

Jenny E. Ostergren, MPH
School of Public Health, University of Michigan
1415 Washington Heights
Ann Arbor, MI 48109-2029 (USA)
E-Mail jeosterg@umich.edu

plex disorders and specific traits without going through a health care provider [1]. The commercialization of genetic testing services has fueled debates among many different stakeholder groups, including researchers, health care professionals, lawyers, ethicists, and policy-makers [2–4]. Proponents of PGT assert that customers should be able to independently access personalized risk information, citing possible health benefits, such as increased awareness about disease risks and motivation to make important lifestyle and treatment choices. However, a number of governmental regulators and professional organizations, including the Food and Drug Administration (FDA) [5], the Government Accountability Office (GAO) [6], the American College of Medical Genetics and Genomics (ACMG) [7], and the American Society of Human Genetics (ASHG) [8], have raised concerns about the potential for customer misunderstandings, noting that misinterpretation of test results could result in psychological harms and misuse of health care system resources. On November 22, 2013, the FDA sent a warning letter to 23andMe, Inc., a leading provider of PGT services, raising concerns about the potential consequences of misunderstood test results, such as patient noncompliance or mismanagement of medications [5]. The company discontinued the provision of its health-related PGT services in the United States, although they have recently released their services in the UK and Canada. The debate about the appropriateness of the action by the FDA and the potential for both benefit and harm from PGT continues [9, 10].

The ability to accurately interpret and understand PGT information requires an understanding of the meaning of risk values associated with test results [11] and is aided by an understanding of genetic principles. However, genetic literacy and numeracy skills among the general public tend to be low [12, 13]. While the average customer of PGT services is likely to have greater awareness of genetic principles than the general population, the lack of a health professional to help interpret PGT results may lead to problems comprehending results and misinterpretation. Some potential risks of test misinterpretation are unnecessary health-related or medical decisions in the case of results that are perceived to be higher risk, and false reassurance in the case of results that are perceived to be lower risk. Accordingly, some have advocated that test results of health-related or medical significance should be delivered via a trained professional such as a physician or genetic counselor [8].

To date, relatively few studies have examined comprehension after PGT [14], and those that exist have presented mixed findings. As part of the Multiplex Initiative, Kaphingst et al. [15] examined patients' recall and interpretation of genetic susceptibility test results for 8 health conditions sent by mail to study participants from a large health maintenance organization. The authors reported that 80% of the 199 participants accurately recalled their results, and that most participants did not interpret the risk information as deterministic. Participants who had a more deterministic interpretation of genetic test results were more likely to be confused about the information, have lower levels of education, and be members of racial or ethnic minority groups [15].

Other studies have suggested that misunderstanding or misinterpretation of results may be relatively common, at least when exploring the understanding of the general public. For instance, Leighton et al. [11] compared the responses of individuals from the general public (n = 145) to genetic counselors' (n = 171) responses to four mock test result scenarios for risk of developing colorectal cancer, heart disease, and skin cancer. While a majority of public responders interpreted the results correctly across scenarios (58–72.4% correct), on average they exhibited lower levels of risk accuracy than the genetic counselors and were more likely to overestimate the benefits of testing [11].

These studies provide some insight into how customers interpret risk information from PGT services; however, few have recruited actual customers of PGT services as study participants, study sample sizes have been small, and scenarios used have been limited in scope, often focusing on risk information for one or two well-known diseases.

Here, we report on customer comprehension of hypothetical PGT results from the NIH-funded Impact of Personal Genomics (PGen) Study [16, 17], a web-based survey of new customers from two PGT companies, 23andMe, Inc., (23andMe) and Pathway Genomics (Pathway). Our primary aims were to: (1) assess participants' comprehension of the implications of PGT results using four hypothetical scenarios, and (2) examine possible demographic correlates of comprehension. We have previously shown that, among customers in the PGen Study, genetic literacy and self-efficacy with genomic information (defined as confidence in one's ability to understand and use genetic information), as captured by stand-alone measures, are high prior to testing [18]. However, it is unclear whether and how performance on these measures translates into comprehension of the actual genetic testing results provided to customers. Based on prior research findings [15, 19, 20], we hypothesized that there would be generally high customer comprehension across scenarios, and participants with higher education, nu-

meracy, genetic knowledge, and self-efficacy with genomic information would have a significantly higher proportion of correct responses.

Materials and Methods

Participants and Procedures

The PGen Study was developed in collaboration with academic researchers at Brigham and Women's Hospital/Harvard Medical School and the University of Michigan School of Public Health, research scientists from 23andMe and Pathway, and experts in survey design and administration from the Survey Sciences Group (SSG). Complete details of the development of this academic-industrial partnership [16] and the design and methods used in the PGen Study [17] have been previously published.

Briefly, new customers of 23andMe and Pathway were sent E-mail invitations between March and July of 2012 to participate in the PGen Study, and a banner was posted on the Pathway website inviting new customers to join the study. The invitation E-mail and banner included a link to a consent form, and participants who consented to complete study surveys and share their de-identified genetic risk information with the study investigators were enrolled in the study.

Web surveys were administered by the SSG, an independent survey research firm, at three time points: baseline (BL; upon study enrollment and after genetic testing was ordered, but before results were returned), approximately 2 weeks (2W) after the return of the results, and approximately 6 months after the return of the results.

This analysis used data from the BL and 2W follow-up surveys. Of the 1,046 2W follow-up survey respondents, 20 were partial completers (i.e., did not reach the end of the survey), leaving 1,026 participants who submitted a full 2W survey. Of the 20 partial completers, 4 completed the scenario questions we assessed and were included in the analysis (n = 1,030).

Survey questions were customized to be consistent with each company's report content and format. Unique identifiers were handled by the SSG to protect participant confidentiality. Additional details regarding recruitment and enrollment, survey customization and administration, data flow and curation, and protection of participant confidentiality have been previously reported [17].

Demographics

Participants' demographic characteristics and levels of numeracy, genetic knowledge, and self-efficacy with genetic information were collected at BL. Responses to genetic test scenarios were collected via the 2W follow-up survey.

Demographic characteristics, including age, gender, race/ethnicity, educational level, and household income, were assessed through self-report.

Numeracy

Five items were included in the baseline survey that assessed numeracy. The items were adapted from the seven-item expanded numeracy scale by Lipkus et al. [21] and assessed concepts such as converting percentages and probabilities to proportions and determining magnitudes of risk. A summed score for each participant was created by totaling the number of correct responses (range: 0–5). In order to minimize participant burden, the full scale was not used.

Genetic Knowledge

Few validated instruments exist to measure genetic literacy or knowledge, and those that do have been validated for use in specific populations, such as undergraduate biology students [22]. Since none of the available instruments were appropriate for a population of customers undergoing PGT, we selected individual items from a number of validated scales measuring genetic knowledge in the lay public [20, 23–25] to build a set of questions that matched both the study participants and the PGT context. Participants responded to nine statements (response options: true or false) concerning genetic and environmental influences on health and disease (table 1). The number of correct responses was summed for each participant (range: 0–9).

Self-Efficacy with Genomic Information

Participants' belief in their confidence and ability to understand and use genetic information was assessed through five items adapted from the six-item measure of genetic self-efficacy by Kaphingst et al. [15] (table 1). Participants were asked to indicate their level of agreement or disagreement (7-point Likert scale from 1 = strongly disagree to 7 = strongly agree). This scale has demonstrated high internal consistency across items (Cronbach's alpha = 0.94) in the PGen Study population [18]. Responses to the items were summed for each participant, with higher scores indicating greater self-efficacy (range: 5–35).

Comprehension of Hypothetical Results Scenarios

Four scenarios were presented to participants. Each displayed risk information for a hypothetical customer in the format of a typical disease risk report provided by the respective companies (for scenarios and answers, see suppl. fig. S1–S6; www.karger.com/doi/10.1159/000431250) followed by several questions to assess comprehension. Some question items were adapted from those used by Kaufman et al. [26] and others were constructed by the PGen Study team. The development of scenarios involved a multidisciplinary team of experts in medical genetics, genetic counseling, health education, primary care, and survey methodology. While the scenarios or questions were created for the purposes of the PGen Study, the result reports were modeled based upon actual 23andMe and Pathway reports. Pilot testing of survey items was conducted prior to launching the survey to ensure clarity of presentation. To determine overall comprehension for each participant, a comprehension score was calculated by summing the number of correct responses across the four scenarios (range: 0–11).

Two scenarios presented reports on disease risk. In the first, participants received Alzheimer's risk information [*Apolipoprotein-E* (*APOE*) genetic results] for Lindsay, a 55-year old woman. In the second, participants were provided with a risk report for type 2 diabetes for Dan, a 35-year-old man who, as part of the scenario, is obese according to his body mass index.

Three other scenarios presented carrier screening reports. In the first, a carrier screening summary report on more than 30 different conditions was presented for Erin (age and health status not specified). Participants were randomized to receive either Erin's detailed phenylketonuria (PKU) carrier result or her detailed cystic fibrosis (CF) carrier result, so that about half received the positive PKU carrier screening result (higher than average risk) and the other half the negative CF carrier screening result (lower than average risk). Randomization was performed to reduce the time bur-

den for participants and allow a comparison of responses across scenarios that would not be contaminated by order effects.

One scenario presented a drug response report on Frank (age unspecified), who takes a statin drug called simvastatin to reduce his cholesterol level. The results indicated that Frank has a genetic marker that increases the risk of statin-induced myopathy.

Data Analyses

Data were analyzed using SPSS version 22 software for Windows. Descriptive statistics were used to characterize customers with regard to demographics, genetic knowledge, numeracy and self-efficacy with genetic information. The percentage of participants who correctly identified the risk level in each hypothetical results scenario was determined, and an overall average score across scenarios was calculated. Cronbach's alpha statistics were computed for the five numeracy items. We assessed multicollinearity among independent variables (e.g., genetic knowledge, numeracy, self-efficacy, education, and income) using the standard errors for the β coefficients. A standard error >2.0 may indicate a problem, such as multicollinearity among independent variables. In SPSS we performed stepwise binary logistic regression using the forward logistic regression method (forward LR) for including variables in the model. Forward LR uses the likelihood ratio test to determine which variables are entered in the model and in what order.

We used a logistic regression analysis to examine the potential impact of several factors on the likelihood that participants would have a high comprehension score on the hypothetical scenarios, with statistical significance assessed at $p < 0.05$. The model contained 9 predictors: the sum scores for numeracy, genetic knowledge, and self-efficacy with genetic information, as well as age (3 categories: 19–39, 40–59, and ≥60 years), gender (male or female), self-reported race (White or other), education (4 categories: some college or less, college graduate, some postgraduate study, and doctorate or professional degree), income (3 categories: USD <40,000, 40,000–100,000, and ≥100,000), and we also controlled for respective company (Pathway or 23andMe). Age was entered into the model as a categorical variable for ease of interpretation. The comprehension sum score was dichotomized into higher and lower comprehension through a median split procedure, with lower comprehension (below or equal to the median of 9; n = 588) or higher comprehension (above the median of 9; n = 442) groups. To explore whether associations were unique to specific test scenarios, we ran three separate regressions with comprehension sum scores that were specific to each of the three types of scenarios: disease risk, carrier status, and drug response.

Table 1. Genetic knowledge and self-efficacy items (n = 1,030)

Genetic knowledge survey items[a]	
Healthy parents can have a child with an inherited disease (true)	99.4
Some genetic disorders occur more often within particular ethnic groups (true)	99.1
A healthy lifestyle can prevent or lessen the negative consequences of having genetic predispositions to some diseases (true)	95.8
If your close relatives have diabetes or heart disease, you are more likely to develop these conditions (true)	95.7
The environment has little or no effect on how genes contribute to disease (false)	93.9
Some of the genetic disorders occur later in adult life (true)	93.2
Once a genetic marker for a disorder is identified in a person, the disorder can usually be prevented or cured (false)	88.9
A disease is only genetically determined if more than one family member is affected (false)	87.5
Most genetic disorders are caused by only a single gene (False).	63.0
Self-efficacy survey items[b]	
I am able to understand information about how my genes can affect my health	94.5 (43.6)
I am confident in my ability to understand information about genetics	91.8 (42.7)
I have a good idea about how genetics may influence risk for disease generally	91.4 (34.1)
I have a good idea about how my own genetic make-up might affect my risk for disease	83.7 (27.6)
I am able to explain to others how genes affect one's health	76.1 (22.8)

[a] The correct answer is shown in parentheses, and the results are presented as % correct.

[b] The results are presented as % agree (including somewhat agree, agree, and strongly agree), with % strongly agree in parentheses.

Results

Sample Characteristics

Demographic characteristics are summarized in table 2. The majority of the 1,030 respondents self-identified as White (85.1%), followed by Asian (4.6%), African-American (3.3%), and American Indian/Native Alaskan (3.0%). No significant demographic differences emerged between the 2W follow-up survey respondents and the full sample (n = 1,648) who had completed baseline survey data [17].

Participants demonstrated high numeracy (mean score: 4.7 on a 5-point scale); on four of the five items, >96% of the participants answered correctly. The numeracy scale had an internal consistency (Cronbach's alpha) of 0.37. Genetic knowledge (mean score: 8.15 on a 9-point scale) was also high, with ≥93% responding correctly to six of the genetic knowledge items (table 1). Only 63% provided a correct answer to the following item: 'Most genetic disorders are caused by only a single gene'. Participants also displayed high self-efficacy (mean score: 29 on a 35-point scale; table 1). Three of the self-efficacy

Table 2. Participant characteristics (n = 1,030)

Mean age ± SD (range), years	46.7±15.7 (19–91)
Age groups	
19–39 years	35.1
40–59 years	37.5
≥60 years	27.4
Gender	
Female	59.9
Race	
Non-White	14.9
Ethnicity	
Hispanic/Latino	4.7
Education	
Less than a college degree	21.1
College degree	30.1
Some graduate school[a]	35.7
Doctoral degree[b]	13.2
Employment status	
Full-time	50.8
Part-time	9.7
Retired	20.9
Unemployed	6.8
Student	8.0
Household income	
USD <40,000	17.1
USD 40,000–99,999	34.9
USD 100,000–199,999	31.7
USD >200,000	12.2
Company	
Pathway	39.5
23andMe	60.5

Values are percentages unless otherwise indicated.

[a] Some graduate school, Master's degree, or some doctoral work.

[b] Doctoral degree (e.g., PhD, DSc, EdD), Doctor of Medicine (MD), or other professional degree equivalent to a doctoral degree (e.g., JD, LLB, DDS, DVM).

Table 3. Responses to PGT scenarios (n = 1,030)

Scenario 1: Alzheimer's disease risk		
(1)	Based on these results, what are Lindsay's chances of developing Alzhiemer's disease compared to the average woman of her age and ethnicity? *[correct answer = much higher]*	66.0
(2)	Based on her results, which of the following statements best describes Lindsay›s chances of developing Alzheimer's disease? *[correct answer = she has a 43% chance of developing Alzheimer's disease by age 79 (23andMe); she has a >13% chance of developing Alzheimer's disease (Pathway)]*	83.2
Scenario 2: Type 2 diabetes risk		
(3)	Based on his GENETIC results, what are Dan's chances of developing diabetes compared to the average man of his age and ethnicity? *[correct answer = somewhat lower]*	59.2
(4)	Based on his GENETIC results, will Dan develop diabetes? *[correct answer = probably not]*	82.3
(5)	Which of the following is a true statement about Dan's risk of diabetes? *[correct answer = Dan's obesity is an important risk factor for diabetes regardless of his genetic results]*	93.1
Scenario 3: Carrier screening results		
(6)	Erin does not carry any variants/mutations for the diseases listed in the report. *[correct answer = false]*	96.8
(7)	Erin herself likely has one of the diseases or conditions listed in the report. *[correct answer = false]*	81.1
(8)	Erin's children could inherit a variant or mutation for one of the conditions listed in the report. *[correct answer = true]*	97.4
Phenylketonuria (PKU) results[a]		
(9a)	Based on these results, what are the chances that Erin has phenylketonuria (PKU)? *[correct answer = most likely does not have PKU]*	71.0
(10a)	The father of Erin's child is a carrier of a PKU mutation. Based on these results, what is the chance for Erin's child to have PKU? *[correct answer = 25%]*	63.6
Cystic fibrosis (CF) results[a]		
(9b)	Based on these results, what are the chances that Erin has cystic fibrosis (CF)? *[correct answer = most likely does not have CF]*	67.2
(10b)	Based on these results, what is the chance that Erin is a carrier of a CF mutation? *[correct answer = most likely is not a carrier]*	74.8
Scenario 4: Statin drug response		
(11)	Based on his statin drug response results, what are Frank's chances of myopathy while taking statin therapy? *[correct answer = higher than average]*	92.7
Average of correct responses across items		79.1

The results are presented as % correct. [a] Participants were randomized to receive PKU or CF results.

statements had over 90% agreement. The lowest item (76.1%) was: 'I am able to explain to others how genetic variants affect one's health'.

Comprehension of Test Scenarios

There was high comprehension across the four scenarios (table 3). The majority of participants chose the correct response for each question, with an average overall score of 8.7 out of 11 (79.1% correct) across scenarios. Participants demonstrated the highest comprehension (81.1–97.4% correct) for the statin drug response result and carrier screening summary report, and lower comprehension (63.6–74.8% correct) for the specific carrier screening results for PKU and CF.

Table 4 presents the results of the logistic regression analysis examining the association of participant characteristics on the comprehension sum score. The full model containing all predictors was statistically significant, [χ^2 (11, n = 1,013) = 326.17, p < 0.001] and explained between 27.5% (Cox and Snell R^2) and 37.0% (Nagelkerke

Table 4. Summary of logistic regression analysis (n = 1,013)

Variables	OR	95% CI	Wald	d.f.	p
Age groups					
19–39 years (ref.)					
40–59 years	0.93	0.64–1.33	0.18	1	0.675
≥60 years	0.57	0.39–0.85	7.61	1	0.006
Gender					
Male (ref.) or female	0.94	0.69–1.27	0.18	1	0.668
Race					
Non-White (ref.) or White	2.08	1.34–3.21	10.81	1	0.001
Education					
Less than college degree (ref.)					
College degree	1.17	0.75–1.82	0.47	1	0.494
Some graduate school	1.60	1.05–2.45	4.72	1	0.030
Doctoral degree	1.39	0.81–2.39	1.40	1	0.237
Income					
USD <40,000 (ref.)					
USD 40,000–99,000	0.93	0.58–1.48	0.10	1	0.755
USD ≥100,000	0.90	0.57–1.43	0.20	1	0.654
Numeracy	1.71	1.29–2.26	13.93	1	<0.001
Genetic knowledge	1.38	1.17–1.63	14.13	1	<0.001
Self-efficacy	1.02	0.99–1.05	2.43	1	0.119

Data were adjusted for the respective company. A median split procedure was used to dichotomize the comprehension sum score (range: 0–11) into lower comprehension (below or equal to the median of 9; n = 588) or higher comprehension (above the median of 9; n = 442).

The reduced sample size in the regression model reflects missing data on income.

R^2) of the variance in comprehension score. In our assessment of multicollinearity, none of the independent variables had standard errors >2.0 (standard error for genetic knowledge: 0.085; numeracy: 0.158; self-efficacy: 0.014; education: 0.258). Therefore, we did not find evidence of the independent variables being highly correlated (Pearson correlations ranged from –0.166 to 0.472). Five of the independent variables made statistically significant contributions to the model (numeracy, genetic knowledge, race, age, and education). Respondents with high numeracy scores were nearly 2 times more likely than the lower numeracy group to have comprehension scores above the median. Higher genetic knowledge was significantly associated with greater comprehension of test scenarios. Self-identified race was a predictor of comprehension score, with White participants being more than twice as likely to score above the median on the scenarios. Older age (≥60 years) was associated with lower comprehension scores. Participants with some graduate education were significantly more likely to have higher comprehension than those without a college degree. Finally, the company from which customers received their interpretation of genetic findings was also significantly correlated with comprehension score (OR = 0.076, p < 0.001). To facilitate companies' participation in the PGen Study, our research team agreed not to publish results highlighting company differences that could be used to gain an advantage in the marketplace. Thus, we do not report the direction of this finding here. Details about the development of this academic-industrial partnership have been previously published [16].

Secondary Analyses

We ran an additional subset of analyses to see if the associations were unique to specific scenarios. The same effects were observed in the disease risk and carrier status scenario analyses. However, in addition to the variables and categories that were found to be significant in the original model, doctoral education was found to be significantly associated with higher comprehension scores in the disease risk analysis (OR = 2.15, p = 0.012), and the age category of 40–59 years was significantly associated with lower comprehension scores in the carrier status analysis (OR = 0.71, p = 0.029). However, in the drug response scenario analysis, only genetic knowledge was found to be marginally significantly associated with comprehension score (OR = 1.26, p = 0.054).

Discussion

Given the concerns raised about marketing and provision of PGT services, this study offers a timely assessment of customer comprehension of PGT information. Overall, most participants were able to correctly interpret each scenario presented to them. This was especially true for the statin drug response and carrier screening summary reports, in which participants identified the correct response 81–97% of the time. These findings are consistent with other recent studies of comprehension of PGT results and suggest that major misunderstandings of health implications of PGT information are not widespread [15, 19]. For instance, Kaufman et al. [19] reported that over 90% of PGT customers correctly interpreted hypothetical disease risk scenarios for type 2 diabetes and colorectal cancer.

Participants demonstrated lower comprehension, however, on the specific carrier screening questions for PKU and CF. One possible explanation is the lack of familiarity with the concept of recessive traits. To correctly interpret the CF and PKU scenarios, customers must grasp the distinction between carrier status and being affected by a condition and recognize that they could have a negative result but still 'carry' a genetic mutation. Although the biological mechanisms underpinning multifactorial disease risk (including genetic and environmental factors)

are more complex than those for autosomal recessive Mendelian conditions, multifactorial disease risks obtained from PGT may in fact be simpler for customers to understand because they are presented as a risk of 0–100%, or from low to high. Moreover, multifactorial disease risk estimates are relevant to the immediate customer, whereas carrier testing rarely has implications for the ordering customer, but must instead be understood in the context of reproductive risk. Indeed, as has been suggested by others, the notion of inheriting a disease that is never seen or expressed may be a difficult concept to grasp [12]. The results for CF and PKU may also have been presented less clearly than the statin drug report, which directly states that the individual has a marker that significantly increases the risk for myopathy. Individuals may also be more prone to think about medication side effects in general (due to advertising), which may make the statin drug question easier to answer.

Responses to one of the items in the type 2 diabetes scenarios suggest that participants may have been focused on information that the test subject was obese (obesity is a well-known risk factor for type 2 diabetes). They therefore were more likely to think that Dan's *genetic* risk was higher than the report indicated, despite the fact that the majority of participants correctly identified obesity as an important risk factor regardless of genetic results. This finding may suggest that customers are not fully appreciating that PGT services often do not factor in behavioral risk factors when quantitatively estimating future disease risks. Behavioral and environmental influences contribute to risk for most common diseases; when these factors are not incorporated into quantitative risk estimates, such limitations should be clearly conveyed to customers.

Several of our findings were consistent with those of other studies indicating that numeracy, genetic knowledge, and education level are likely to predict comprehension of genetic information [24, 27]. It is important to note that both genetic knowledge and numeracy were found to be significant correlates of comprehension. This suggests that genetic knowledge alone may not be sufficient, as correct interpretation also involves numeracy skills independent of genetic knowledge. It is also interesting that objective measures of genetic knowledge and numeracy were predictive of comprehension, but the subjective measure of self-efficacy was not. This may suggest that patients' self-ratings of their genetic knowledge should not be taken at face value. Our finding of group differences with regard to race/ethnicity and comprehension is also consistent with some studies that suggested possible racial group differences in genetic knowledge [28, 29], although this finding should be interpreted with caution given that over 85.1% of the participants in our study self-identified as White. In addition, racial differences in comprehension may be due to educational differences or other social or cultural influences that were insufficiently adjusted for in our study. Lastly, older age has been found to be associated with incorrect interpretations of genetic information [19], which is consistent with our finding that the age group of ≥60 years had significantly lower comprehension than the youngest age group. Some older adults may lack exposure to genetics as it was not a part of their science curriculums. Sources of genetic information, such as news media articles intended for a lay audience, may also be overwhelming for some older adults or even inaccurate [30]. For these reasons, older adults may be more prone to misconceptions about PGT information.

One important implication of these findings is that there may not be a one-size-fits-all approach to communicating genetic test information. Greater tailoring of the presentation of PGT information based on individual characteristics and type of test result may be needed, especially when results are not delivered in a clinical setting or via a trained health care professional. Lautenbach et al. [31] provide several strategies for communicating risk information, including varying the presentation format based on individual preference. For instance, customers who have lower numeracy may prefer formats that present risk figures in qualitative terms. Individuals with low genetic knowledge may require a simpler explanation of certain types of genetic risk information, such as carrier screening results for lesser-known conditions. Presentation of results could also be tailored to specific demographic identifiers, such as older age and education. Lautenbach et al. [31] also suggest that the framing of risk information is an important consideration. As demonstrated in the scenario about Dan's genetic risk for type 2 diabetes, how information is presented may lead to bias in risk perception. Presenting information using absolute versus relative risk estimates can influence individual interpretation [31]. The differences observed by company in participants' comprehension scores may also suggest that the presentation or format of the test results influence customer understanding. Utilizing best practices in genetic risk communication could enhance comprehension of results, particularly in a PGT delivery model where an intermediary may not be available to clarify confusing points or respond to questions about the meaning of test results.

The present study has a number of strengths. There were more than 1,000 participants, making this one of the larger studies on comprehension of PGT information and

of PGT services more generally (a recent review of empirical studies in this area has been published by Roberts and Ostergren [14]). Our scenarios cover more ground than prior efforts, assessing comprehension of not just health risks but carrier screening and drug response. The scenarios were also presented in the format of a typical disease risk report provided by the respective companies, enhancing the authenticity of responses. Study findings may help contribute to ongoing policy debates around potential risks of customer misunderstanding of PGT information. However, it should be noted that the purpose of this study was to examine correlates of comprehension in a large sample of current PGT customers, and PGT customers at this time are not representative of the general population [17]. The customers in this analysis tended to be highly educated, with generally strong numeracy and genetic knowledge skills. On the one hand, the generally high comprehension of hypothetical test results among PGT customers is reassuring. On the other hand, the misinterpretation that did occur may be magnified if PGT expands out to sectors of the population with lower genetic knowledge and numeracy skills. As PGT services become more popular, less expensive, and more mainstream, they are likely to be obtained by a more diverse customer population; future studies in this area may want to examine comprehension in both a less educated and more diverse sample.

There were several limitations to this study. First, as noted earlier, the study participants had higher levels of education and income than the general population and were also predominantly White. Other studies have been similarly limited in terms of study population. Therefore, our sample is not representative of the broader US population, and our findings cannot be generalized beyond patients who are currently using PGT services. More research is needed to determine if study participants are representative of PGT customers as a whole and how a less educated sample would have fared in terms of comprehension. Second, there may have been limitations in the test scenarios themselves. Each scenario had high face validity and was pilot tested, but the comprehension scale itself was not formally validated. Since comprehension score data were highly skewed, a logistic regression analysis using a dichotomous outcome generated by a median split procedure was employed. However, this lack of variability heightened the risk of misclassification of participants. It is also important to note that we chose to apply a more strict definition of accuracy in assessing responses to scenario questions (e.g., in a scenario where the correct response was viewed as 'much higher' risk, we did not count responses of 'somewhat higher' as correct). In addition, fewer items were used to assess comprehension of pharmacogenomic results than disease and carrier status results. As a result, the depth of comprehension of pharmacogenomic results was assessed to a lesser extent than for disease and carrier status results, which may have implications for the interpretation of study findings. Third, this analysis did not examine how participants understand their own personal genetic test results. Because results were presented in the style and format used by the companies, responses to the hypothetical scenarios arguably provide insight into how participants might have interpreted their own personal genetic test information. However, more research is needed that looks at customers' comprehension of their actual genetic test results. Other PGen Study data may allow us to shed light on this question through separate analyses.

In sum, it is important that PGT customers are able to understand their test results and make informed health care decisions. This study examined customers' comprehension of PGT results in several hypothetical scenarios and investigated possible predictors of correct interpretation of genetic information. Study participants generally demonstrated a solid grasp of information presented to them. Factors such as age, numeracy, race, genetic knowledge, and education played an important role in test interpretation. Additional research is needed to further elucidate the role of these factors and to examine the behavioral implications of participants' interpretation of their PGT results.

Appendix

Members of the PGen Study at the time of publication are as follows: Robert C. Green, Joel B. Krier, Caroline M. Weipert, Sarah S. Kalia, Kurt D. Christensen, and Lisa S. Lehmann, Harvard Medical School and Brigham and Women's Hospital; Deanna Alexis Carere and Peter Kraft, Harvard T.H. Chan School of Public Health; J. Scott Roberts, Mack T. Ruffin IV, Lan Q. Le, and Jenny Ostergren, University of Michigan School of Public Health; Wendy R. Uhlmann and Mick P. Couper, University of Michigan; Joanna L. Mountain and Amy K. Kiefer, 23andMe; Tanya A. Moreno and Adrian Vilalta, Pathway Genomics; Scott D. Crawford, Survey Sciences Group; L. Adrienne Cupples, Clara A. Chen, and Catharine Wang, Boston University; Stacy W. Gray, Dana-Farber Cancer Institute; Barbara A. Koenig, University of California San Francisco; Kimberly Kaphingst, University of Utah; Sarah Gollust, University of Minnesota.

Acknowledgments

The PGen Study is supported by the National Institutes of Health (NIH) National Human Genome Research Institute (R01-HG005092). D.A.C. is supported by the Canadian Institutes of

Health Research (Doctoral Foreign Study Award). R.C.G. is supported by NIH U01-HG006500, U19-HD077671 and R01-HG002213. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Center for Research Resources, the National Institutes of Health, or the Canadian Institutes of Health Research.

Disclosure Statement

J.L.M. is a paid employee of 23andMe. The other authors declare no conflicts of interest. Neither 23andMe nor Pathway Genomics provided any funding or other financial contribution to the PGen Study. Final decisions regarding study design, data collection and analysis, and publication were made by academic researchers.

References

1 Caulfield T, McGuire AL: Direct-to-consumer genetic testing: perceptions, problems, and policy responses. Annu Rev Med 2012;63:23–33.
2 Frueh FW, Greely HT, Green RC, Hogarth S, Siegel S: The future of direct-to-consumer clinical genetic tests. Nat Rev Genet 2011;12:511–515.
3 Evans JP, Green RC: Direct to consumer genetic testing: avoiding a culture war. Genet Med 2009;11:568–569.
4 Foster MW, Sharp RR: The contractual genome: how direct-to-consumer genomic services may help patients take ownership of their DNA. Per Med 2008;5:399–404.
5 US Food and Drug Administration: Re: Personal genome service (PGS). 2013. http://www.fda.gov/ICECI/EnforcementActions/WarningLetters/2013/ucm376296.htm (accessed December 18, 2014).
6 United States Government Accountability Office: Misleading Test Results Are Further Complicated by Deceptive Marketing and Other Questionable Practices. 2010. http://www.gao.gov/products/GAO-10-847T (accessed December 18, 2014).
7 American College of Medical Genetics: ACMG Statement on Direct-to-Consumer Genetic Testing. 2008. https://www.acmg.net/StaticContent/StaticPages/DTC_Statement.pdf (accessed December 18, 2014).
8 Hudson K, Javitt G, Burke W, Byers P; American Society of Human Genetics Social Issues Committee: ASHG statement* on direct-to-consumer genetic testing in the United States. Obstet Gynecol 2007;110:1392–1395.
9 Green RC, Farahany NA: Regulation: the FDA is overcautious on consumer genomics. Nature 2014;505:286–287.
10 Janssens C: How FDA and 23andMe dance around evidence that is not there. 2014. http://www.huffingtonpost.com/cecile-janssens/post_6753_b_4671077.html (accessed March 12, 2015).
11 Leighton JW, Valverde K, Bernhardt BA: The general public's understanding and perception of direct-to-consumer genetic test results. Public Health Genomics 2012;15:11–21.
12 Lanie AD, Jayaratne TE, Sheldon JP, Kardia SL, Anderson ES, Feldbaum M, Petty EM: Exploring the public understanding of basic genetic concepts. J Genet Couns 2004;13:305–320.
13 Kelly KM, Graves KD, Harper FW, Schmidt JE, Dickinson SL, Andrykowski MA: Assessing perceptions of cancer risk: does mode of assessment or numeracy matter? Cancer Detect Prev 2007;31:465–473.
14 Roberts JS, Ostergren J: Direct-to-consumer genetic testing and personal genomics services: a review of recent empirical studies. Curr Genet Med Rep 2013;1:182–200.
15 Kaphingst KA, McBride CM, Wade C, Alford SH, Reid R, Larson E, Baxevanis AD, Brody LC: Patients' understanding of and responses to multiplex genetic susceptibility test results. Genet Med 2012;14:681–687.
16 Lehmann LS, Kaufman DJ, Sharp RR, Moreno TA, Mountain JL, Roberts JS, Green RC: Navigating a research partnership between academia and industry to assess the impact of personalized genetic testing. Genet Med 2012;14:268–273.
17 Carere DA, Couper MP, Crawford SD, Kalia SS, Duggan JR, Moreno TA, Mountain JL, Roberts JS, Green RC: Protocol development and initial results from the Impact of Personal Genomics (PGen) Study. Genome Med 2014;6:96.
18 Carere DA, Kraft P, Kaphingst KA, Roberts JS, Green RC; PGen Study Group: Consumers report lower confidence in their genetics knowledge following direct-to-consumer personal genomic testing. Genet Med, Epub ahead of print.
19 Kaufman DJ, Bollinger JM, Dvoskin RL, Scott JA: Risky business: risk perception and the use of medical services among customers of DTC personal genetic testing. J Genet Couns 2012;21:413–422.
20 Molster C, Charles T, Samanek A, O'Leary P: Australian study on public knowledge of human genetics and health. Public Health Genomics 2009;12:84–91.
21 Lipkus IM, Samsa G, Rimer BK: General performance on a numeracy scale among highly educated samples. Med Decis Making 2001;21:37–44.
22 Daack-Hirsch S, Driessnack M, Perkhounkova Y, Furukawa R, Ramirez A: A practical first step to integrating genetics into the curriculum. J Nurs Educ 2012;51:294–298.
23 Furr LA, Kelly SE: The genetic knowledge index: developing a standard measure of genetic knowledge. Genet Test 1999;3:193–199.
24 Smerecnik CM, Mesters I, de Vries NK, de Vries H: Applying a theory-based framework to understand public knowledge of genetic risk factors: a case for the distinction between how-to knowledge and principles knowledge. Public Health Genomics 2011;14:259–270.
25 Bowling BV, Acra EE, Wang L, Myers MF, Dean GE, Markle GC, Moskalik CL, Huether CA: Development and evaluation of a genetics literacy assessment instrument for undergraduates. Genetics 2008;178:15–22.
26 Kaufman D, Murphy J, Scott J, Hudson K: Subjects matter: a survey of public opinions about a large genetic cohort study. Genet Med 2008;10:831–839.
27 Lea DH, Kaphingst KA, Bowen D, Lipkus I, Hadley DW: Communicating genetic and genomic information: health literacy and numeracy considerations. Public Health Genomics 2011;14:279–289.
28 Donovan KA, Tucker DC: Knowledge about genetic risk for breast cancer and perceptions of genetic testing in a sociodemographically diverse sample. J Behav Med 2000;23:15–36.
29 Haga SB, Barry WT, Mills R, Ginsburg GS, Svetkey L, Sullivan J, Willard HF: Public knowledge of and attitudes toward genetics and genetic testing. Genet Test Mol Biomarkers 2013;17:327–335.
30 Frazier L, Calvin AO, Mudd GT, Cohen MZ: Understanding of genetics among older adults. J Nurs Scholarsh 2006;38:126–132.
31 Lautenbach DM, Christensen KD, Sparks JA, Green RC: Communicating genetic risk information for common disorders in the era of genomic medicine. Annu Rev Genomics Hum Genet 2013;14:491–513.