Volume Title: Economic Dimensions of Personalized and Precision Medicine

Volume Authors/Editors: Ernst R. Berndt, Dana P. Goldman, and John W. Rowe, editor

Volume Publisher: University of Chicago Press

Volume ISBNs: 978-0-226-61106-8 (cloth); 978-0-226-61123-5 (electronic)

Volume URL: http://www.nber.org/books/bern-13

Conference Date: September 13–14, 2017

Publication Date: April 2019

Chapter Title: Measuring the Potential Health Impact of Personalized Medicine: Evidence from Multiple Sclerosis Treatments

Chapter Author(s): Kristopher J. Hult

Chapter URL: http://www.nber.org/chapters/c13990

Chapter pages in book: (p. 185 – 216)

7

Measuring the Potential Health Impact of Personalized Medicine
Evidence from Multiple Sclerosis Treatments

Kristopher J. Hult

7.1 Introduction

Pharmaceutical treatments in the United States are typically homogeneous products that are tightly regulated by the Food and Drug Administration (FDA) to ensure that the dosage and delivery are consistent across each prescription. However, differences across patients—including genetics, age, comorbidity, preferences, and environment—and differences across diseases—such as severity and progression—cause the impact of a treatment to vary across patients. Patients respond to the same dosage differently, from how their bodies process and react to the treatment to the side effects that arise.[1]

This heterogeneity is often not apparent when assessing the impact of

Kristopher J. Hult is associate principal of Charles River Associates.

I am grateful for thoughtful comments from seminar participants at Columbia University and the NBER Program on Economic Dimensions of Personalized and Precision Medicine and from comments from an anonymous reviewer and Ernst Berndt. The conclusions set forth herein are based on independent research and publicly available material. The views expressed herein are the views and opinions of the author and do not reflect or represent the views of Charles River Associates or any of the organizations with which the author is affiliated. Any opinion expressed herein shall not amount to any form of guarantee that the author or Charles River Associates have determined or predicted future events or circumstances and no such reliance may be inferred or implied. The author and Charles River Associates accept no duty of care or liability of any kind whatsoever to any party, and no responsibility for damages, if any, suffered by any party as a result of decisions made, or not made, or actions taken, or not taken, based on this chapter. Detailed information about Charles River Associates, a registered trade name of CRA International, Inc., is available at www.crai.com. For acknowledgments, sources of research support, and disclosure of the author's material financial relationships, if any, please see http://www.nber.org/chapters/c13990.ack.

1. See discussions and examples in Basu et al. (2014), Kravitz, Duan, and Braslow (2004), and Segal, Weiss, and Varadhan (2012).

innovations because clinical trials and cost-effectiveness research focus on the average treatment effect, even if this effect varies significantly across patients. To understand the potential impact of heterogeneity across treatments, consider two treatments in the same disease category where the health impact of each treatment is measured in quality-adjusted life years (QALYs) across the patient population follows an independent normal distribution with a mean of one QALY and a standard deviation of one.[2] If each patient matches with the treatment that provides the highest individual health impact, then the total impact across all patients is over 56 percent higher than if patients are randomly assigned a treatment.[3] Since the median innovation in health care increases QALYs by around 1 percent relative to existing innovations (Hult, Jaffe, and Philipson 2016), this result suggests that there is enormous potential in addressing treatment heterogeneity.

Personalized medicine is a growing field that addresses the heterogeneity in treatment effects across patients by targeting or tailoring treatments to individuals based on their characteristics. Personalized medicine has the ability to create novel treatments, such as treatments that target specific genes or proteins, and the ability to guide patients to the most efficacious treatment through diagnostic testing or data-driven analysis.[4]

Understanding the potential impact of incorporating patient heterogeneity in biology, environment, and behavior, the United States announced a $215 million Precision Medicine Initiative in 2015.[5] The purpose of this initiative is to provide funding for research in personalized medicine, including building a research cohort to collect individual-level data to help develop more effective treatments and funding cancer genomics, one of the leading research fields in personalized medicine.[6]

We are just beginning to understand the potential impact of precision medicine. Goldman et al. (2013) present a framework for understanding the value of diagnostic tests. In a case study of rofecoxib, a nonsteroidal anti-inflammatory drug that was withdrawn from the market, they show that diagnostic testing can have a large social value by avoiding unnecessary treatment and identifying patients who would not otherwise be treated. Basu (2013) discusses the difference between passive personalization, which is a form of learning-by-doing where patients and physicians learn about patient-specific treatment effects through a trial-and-error process, and

2. QALYs are a frequently used measure of either disease burden or treatment effect that includes the quality and quantity of life lived by the patient.

3. The maximum of two independently distributed normal distributions with mean μ and a standard deviation of one is distributed as a Gumbel or Extreme Value Type 1 distribution, which has a mean of $\mu + (1/\pi) > \mu + 0.56$ (Nadarajah and Kotz 2008).

4. Examples of targeted treatments include human epidermal growth factor receptor 2 (HER2) in breast cancer, epidermal growth factor receptor (eGFR) in colorectal cancer, and BRAF inhibitors for melanoma. See Hutchinson et al. (2015).

5. See www.whitehouse.gov/precisionmedicine.

6. See Chin, Andersen, and Futreal (2011).

active personalization, which involves biomarker and genetic tests that inform patient-specific treatment effects. Egan and Philipson (2014) discuss the role of passive personalization in measuring adherence. They create a dynamic model to argue that personalized medicine has the capacity to expedite this search process, which reduces overadherence and increases underadherence.

The goal of this chapter is to present a framework for understanding the potential health impact of personalized medicine and to compare it to the health impact of other types of pharmaceutical innovations.[7] I present a theoretical framework for measuring the health impact of personalized medicine by modifying the model of Hult (2014). The model in this chapter measures the value of two types of personalized medicine: allocating patients to treatments based on individual treatment effects and identifying individual risk to serious side effects from a treatment. The health impact of these types of personalized medicine depends on the number of treatments, the variance in the health impact within a treatment, the noise in a patient's signal of their treatment effect, and the correlation of treatment effects across the different treatment options.

Using this model, I measure the relative impact of personalized medicine compared with the introduction of new treatments in a case study of multiple sclerosis (MS). I find that the potential health impact of personalized medicine for MS patients would increase the health impact of existing treatments by 21 percent by improving the ability to match patients with the treatment that provides the largest treatment effect and by 30 percent by properly identifying a patient's risk of serious side effects, which can prevent a patient from having a significant adverse side effect.

To understand the value of personalized medicine, consider two examples. First, consider an MS patient deciding which first-line therapy to take. If that patient chooses a therapy on which they will eventually fail (meaning they have a suboptimal response and switch to a different therapy), that patient experiences a relapse rate five times higher compared with their second therapy (Río et al. 2012). These patients stay on their unsuccessful first treatment for almost as long as they stay on their successful treatment (3.9 years versus 4.2 years). For diseases like MS, where the disease progression is irreversible and failing on a treatment produces similar results to taking no treatment at all, the effect of choosing an ineffective treatment can be significant and permanent.

Second, consider an MS patient deciding which second-line therapy to take. Two second-line options are Tysabri, a treatment with the highest efficacy but the risk of a potentially fatal side effect, and Gilenya, a treatment with lower efficacy but with much less severe side effects. When personalized

7. This chapter focuses on pharmaceuticals, but the implications of the chapter are also relevant for medical devices and other medical treatments.

information is used to inform a patient of their individual side effect risk level (which ranges from less than 1 in 10,000 to 1 in 89), high-risk patients are able to avoid being exposed to the potentially fatal side effect while low-risk patients are able to take more efficacious treatments than they would have without personalized information (Sørensen et al. 2012).

This chapter is organized as follows. Section 7.2 discusses the different types of pharmaceutical innovations, which include novel, follow-on, and personalized innovations. Section 7.3 discusses the theory of how to measure the value of personalized innovations. Section 7.4 is a case study of disease-modifying therapies in MS to illustrate the rate of return to personalized medicine. Section 7.5 discusses the development costs and costs to patients of personalized medicine. Section 7.6 concludes.

7.2 Types of Pharmaceutical Innovations

I consider three main types of pharmaceutical innovations: novel innovations, follow-on innovations, and innovations in personalized medicine. A novel innovation is the approval of a chemical entity that has not already been approved by the FDA and is the part of the pharmaceutical treatment that is responsible for the pharmacological action of the treatment. These approvals are either a new molecular entity (for smaller chemically synthesized molecules) or a new biologic (for larger treatments extracted from biological sources). Novel innovation is a necessary precursor for follow-on innovation and personalized medicine. However, novel innovation in its original form often extracts only part of the potential health impact of the new molecule because it generally provides only one treatment that has not been adapted to the heterogeneity of the treatment population or to the learning that takes place from treatments being on the market. Follow-on innovations and personalized medicine develop the molecule into a more efficacious or desirable treatment for patients.

Follow-on innovations take already FDA-approved molecules and create new treatments by changing the dosage, formulation, indication, active ingredient, or by combining two molecules.[8] Roughly 70 percent of all FDA-approved innovations and over half of all prescriptions use follow-on innovations (Hult 2014). Follow-on innovations make three main types of improvements. First, they can create a new treatment by combining existing molecules. Second, they can make existing treatments either more effective or better tolerated. Third, they can expand the number of treatment options available and expand the availability of treatment to subgroups of

8. A follow-on innovation that contains a new active ingredient means that it contains the same active moiety but includes a different enantiomer, racemate, salt, ester, complex, chelate, or clathrate.

the population.[9] The main focus of these innovations is to expand the treatment population, reduce treatment burden, or increase efficacy for a group of patients. For instance, HIV/AIDS treatments in their original form were unable to be taken by pediatric, elderly, and pregnant patients. With follow-on innovation, all of these patient groups now have a variety of treatment options, including oral pellets that can be mixed into children's food or intravenous treatments for patients that cannot take the pill regimen.[10]

The third type of innovations are innovations in personalized medicine, which take follow-on innovations a step further by creating directed treatments or diagnostics tests from the characteristics of an individual patient. These innovations can create new treatments, create data sets or diagnostic tests to determine the best treatment considering individual treatment effects, and identify individual treatment burdens for patients.

Improve Matching and Reduce Searching. One way in which personalized medicine improves the health outcomes of patients is that it can inform a patient about which treatment will either be more efficacious or have a lower burden of treatment through diagnostic testing or patient databases.[11] If patients learn about their individual treatment effect, it can direct patients toward a treatment that makes them better off than if they do not have any individual specific information.

Consider the two treatment options shown in figure 7.1. This figure plots the distribution of patient outcomes for the treatment efficacy and burden of two treatments: Treatment 1 (represented with a dashed line) with average outcome μ_1 and Treatment 2 (represented with a solid line) with average outcome μ_2. The indifference curves (IC) show the efficacy and burden combinations for which patients are equally well off, so a patient is indifferent between receiving the treatment effect of any points along the same IC. Patients are better off with higher efficacy and lower treatment burdens, so they are better off on indifference curves closer to the upper left of the graph. A patient with no information about his individual treatment effect would be indifferent between these two treatments because the average treatment effects (μ_1 and μ_2) lie on the same indifference curve.

However, if a patient learns from diagnostic testing that they receive the efficacy and treatment burden at point h_1 for Treatment 1 and the efficacy

9. Examples of follow-on innovations include the creation of CART treatments used in HIV/AIDS, which combine three different molecules in a treatment that reduces pill burden and potential drug interactions; Fetzima, an SNRI drug used to treat major depressive disorder, was approved as a new active ingredient using a different orientation of the molecule in milnacipran HCI (Savella), which is used to treat fibromyalgia; and Norvir, an HIV/AIDS treatment, received a new formulation that eliminated the need for refrigeration, reduced the number of drug and food interactions, and provided extended release for drugs.

10. See UNAIDS (2015).

11. As a simplification, throughout this chapter I treat the patient as the person who decides what treatment to take even when this decision is heavily influenced by the physician.

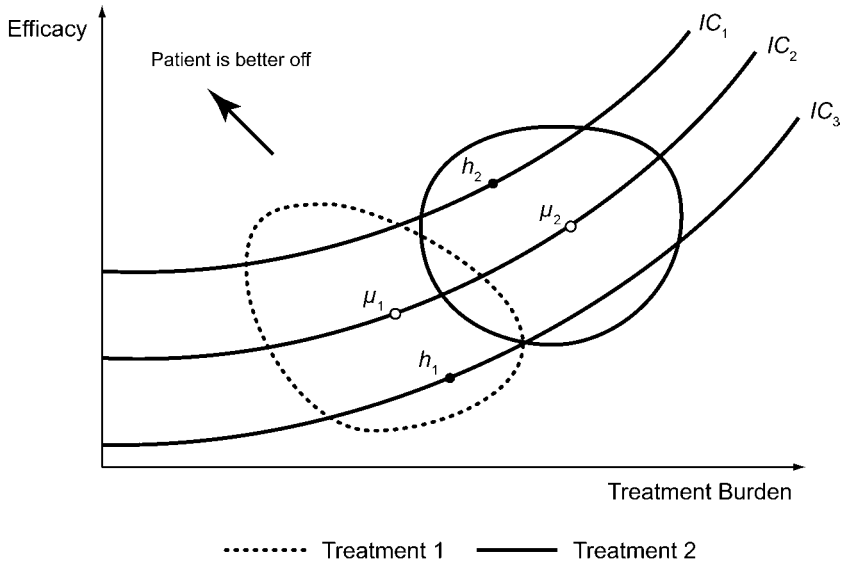

Fig. 7.1 Improve matching and reduce searching

and treatment burden at point h_2 for Treatment 2, then the patient is better off taking Treatment 1 than Treatment 2.

There are different ways a patient can learn about their individual treatment effect, which can broadly be categorized as passive and active personalized medicine.[12] In passive personalized medicine, treatments are experience goods, which means that the patient or his physician learns about a patient's individual treatment effect through learning by doing.[13] Experience goods have several costs including opportunity costs, side effects, and financial costs. For example, if a patient has an aggressive form of MS, taking a treatment that the patient does not respond to can cause irreversible damage and disability and allow the disease to progress to a form of MS that is less responsive to therapy (see Rush, MacLean, and Freedman 2015). In addition, taking less efficacious but milder treatments at an early stage of the disease can increase the risk of serious side effects for a patient who takes more efficacious treatments at a later stage of treatment. Therefore, when patients learn by experiencing the treatment outcome, a patient with a more aggressive disease will be more prone to serious side effects than a patient who can be matched to the more efficacious treatment earlier (see Zaheer and Berger 2012). Finally, MS patients may develop neutralizing

12. See a discussion of passive and active personalization in Basu (2013).
13. See Nelson (1970, 1974) for a discussion of search goods versus experience goods.

antibodies taking one treatment that makes other treatments ineffective. For example, if a patient takes either interferon beta-1b (Betaseron), interferon beta-1b (Extavia), or interferon beta-1a (Rebif), that patient may develop neutralizing antibodies that will block the biological activity of the other two treatments (Malucchi et al. 2004).

In active personalized medicine the treatment is a search good, which means that a patient can learn about the individual treatment effect without having to take the treatment. The patients learn through diagnostic testing or using patient databases to inform how an individual patient may respond to a treatment. In the MS examples above, this would include a test that determines the aggressiveness of a patient's MS or susceptibility of developing neutralizing antibodies to determine which course of treatment is best for that patient.

Part of the potential of personalized medicine is that it has the ability to convert experience goods to search goods, thereby eliminating the potential costs of learning by doing (see chapter 1, this volume). However, it is also important to understand that there is still potential value in personalized medicine, even if the treatment cannot be converted to a search good. For example, even if a patient has to experience the treatment to understand the individual treatment effect, personalized medicine can improve that patient's understanding of whether to switch to a different treatment or adjust taking the current treatment.

Risk Assessment. Another way that personalized medicine impacts health is through risk assessment. Some treatments have very serious side effects for a fraction of the patient population. For example, Tysabri, an MS treatment, has a side effect of progressive multifocal leukoencephalopathy (PML) for up to 0.013 percent of patients. PML is a devastating disease that has a mortality rate up to 50 percent within two years and potentially severe neurological disabilities for those who survive (Pavlovic et al. 2015).[14] PML is an infection caused by the John Cunningham virus (JCV).[15] The risk of PML was enough to get Tysabri pulled from the market within four months of FDA approval. However, Tysabri is the most efficacious treatment for MS patients. The ability to identify patients with higher risk factors can reduce the odds of getting PML from 0.013 to less than 0.001 percent.

For example, figure 7.2 shows the treatment distributions of two treatments, Treatment 1 in a dashed line and Treatment 2 in a solid line. Treatment 2 has a potentially serious side effect (which increases the treatment burden) as shown by the two Treatment 2 distributions. The Treatment 2 distribution on the left (with mean μ_{2A}) is for those patients who do not have the side effect, and the Treatment 2 distribution on the right (with mean μ_{2B})

14. See https://www.tysabri.com/en_us/home/about/safety-side-effects.html.
15. See https://www.tysabrihcp.com/en_us/home/safety/risk-pml-jcv.html.

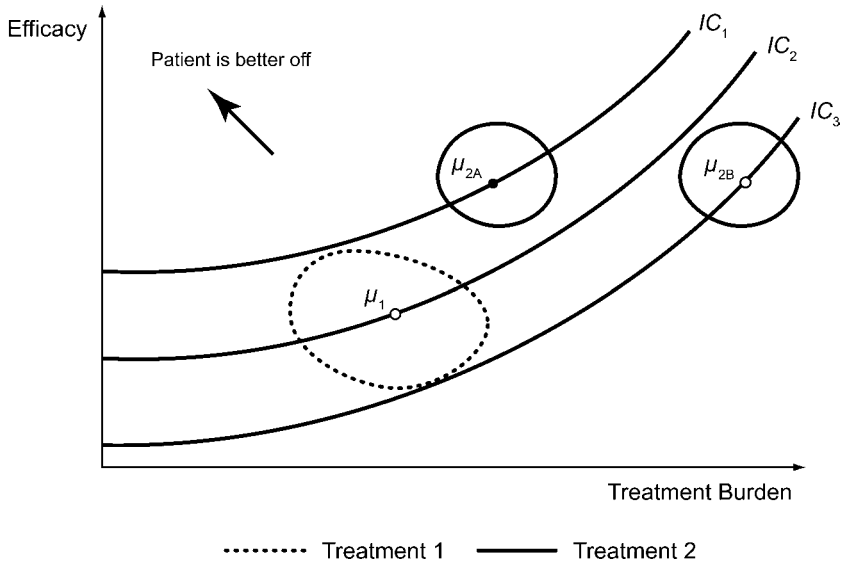

Fig. 7.2 Risk assessment

is for those patients who have the side effect. Without any information about their side-effect risk factor, patients may choose Treatment 1. However, if patients can identify whether they would have the serious side effect from Treatment 2, then patients could incorporate this information into their treatment decision. In this bifurcated outcome, the health impact of treatment increases on average.

7.3 Measuring Health Impact

In this section, I describe a model from Hult (2014) that describes how to measure the health impact of pharmaceutical treatments, and I discuss an extension of the model to incorporate patient heterogeneity and the potential health impact of personalized medicine.

7.3.1 Health Impact of Novel and Incremental Innovation

The health impact of a novel or incremental innovation is how much it increases the patient population's length and quality of life. Innovations affect health through three channels: adherence, quantity measured as the number of users, and efficacy measured in QALYs.

The health impact of treatment t on individual $i(h_{it})$ is

$$h_{it} = a_{it}e_{it},$$

where a_{it} is the adherence and e_{it} is the efficacy conditional on being fully adherent for patient i with treatment t.[16] Health impact is a one-dimensional measure of the total impact of a treatment incorporating efficacy, as well as the treatment burden such as side effects or burden of administration. A negative value for h_{it} means that a patient is worse off taking the treatment relative to not taking the treatment, and the more positive the value for h_{it} the better off the patient is taking the treatment.

Summing across all patients who take treatment $t(i \in T)$, the aggregate health impact of treatment t, H_t, is

$$H_t = \sum_{i \in T} h_{it} = \sum_{i \in T} a_{it} e_{it} = q_t \overline{h}_t = q_t \overline{a}_t e_t,$$

where q_t is the quantity measured as the number of users, whether or not the users are adherent to the treatment.[17] If 100 people take a drug with a 60 percent adherence rate that adds one QALY on average, then the health impact of the drug is 60 QALYs.[18]

To measure the increase in health impact produced by treatment t, which is how treatment t increases health impact relative to the standard of care (SOC) that existed before the innovation, I construct $\Delta H_t^{\text{innovation}}$:

$$\begin{aligned} \Delta H_t^{\text{innovation}} &= \frac{\partial h_t}{\partial q} \Delta q + \frac{\partial h_t}{\partial a} \Delta a + \frac{\partial h_t}{\partial h} \Delta e \\ &= \Delta q_t a_t e_t + \Delta a_t q_t e_t + \Delta e_t q_t a_t \\ &= \Delta q_t h_t + \Delta h_t q_t, \end{aligned}$$

where q_t is the average quantity of treatment t per year, Δq_t is how treatment t changes the quantity relative to the SOC, Δa_t is how treatment t changes the adherence rate relative to the SOC, and Δe_t is how treatment t changes efficacy relative to the SOC. Hence, the health impact of treatment t is the effect of the change in the quantity, adherence, and efficacy relative to what would be used instead of that treatment. For instance, if a treatment with 100 users and an efficacy of one QALY increases the adherence rate relative to the previous SOC by 5 percentage points, then the health impact of that innovation is $0.05 * 100 * 1 = 5$ QALYs. If that drug innovation had an adherence rate of 60 percent and also increased efficacy by 5 percent, then the health impact would be $5 + 0.05 * 100 * 0.6 = 5.3$ QALYs.

16. I treat adherence as a dichotomous variable, rather than a continuous variable, where a patient either adheres to a treatment regimen or does not adhere.

17. Quantity can be either defined as the number of users or the number of patients for a particular disease. This distinction does not matter for the framework because patients who are not adherent would have an adherence of zero, and therefore a health impact of zero.

18. $100 * 0.6 * 1$ QALY = 60 QALYs.

7.3.2 Potential Health Impact of Personalized Medicine

This section describes how to modify the health impact framework to measure the potential health impact of personalized medicine by incorporating patient heterogeneity and individualized treatments options.

Expected Value of Individualized Care

The framework in this section is similar to the expected value of individualized care (EVIC) used in the cost-effectiveness literature (Basu and Meltzer 2007). The EVIC measures the value of individualized care by calculating the potential value that society is willing to pay for individualized care. This framework has been expanded to incorporate the identification and selection of population subgroups (Espinoza et al. 2014), determining which sources of patient heterogeneity to consider (Grutters et al. 2013), and the value of genomic information (chapter 3, this volume).

The framework I discuss is similar to the EVIC framework in that it attempts to quantify the value of individualized care, but it differs from the existing literature in two main ways. First, it measures the benefit of individualized treatment in terms of health impact instead of societal value. Second, it focuses on understanding the form and distribution of health impact using a model from Hult (2014).

Theoretical Framework

To understand the effect of innovations in personalized medicine, consider patient i who receives health impact (h) measured in QALYs and has the choice between two treatments, Treatment A and Treatment B. The health impact for the two treatments is distributed as a bivariate normal

$$h \sim N\left(\begin{bmatrix} \mu_A \\ \mu_B \end{bmatrix}, \begin{bmatrix} \sigma_A^2 & \sigma_{AB} \\ \sigma_{AB} & \sigma_B^2 \end{bmatrix}\right),$$

where μ_t and σ_t^2 are the mean and variance of each treatment $t \in \{A, B\}$ and σ_{AB} is the covariance between A and B. The covariance between treatment effects is important because the more correlated effects are across treatments, the lower the value of identifying individual treatment effects.

Impact of Individual Treatment Effect on Searching. Information about the patient's individual treatment effect can come from numerous sources, including disease severity and progression, genetics, environment, and comorbid conditions. In this section I do not distinguish learning through passive or active learning, as they have the same effect. For simplicity in this section assume $\mu_A > \mu_B$.[19]

19. In this framework, generics and biosimilars can be thought of as treatments with the same distribution and perfect correlation with the branded treatment. Therefore, having a generic

If a patient has no information about the individual treatment effect, then the patient chooses the treatment t with the highest μ_t because the patient's expected health impact for each treatment is $E[h_t] = \mu_t$. In this scenario, each patient chooses Treatment A, and the average treatment effect across all patients is $\bar{h}_1 = \max_{t\in\{A,B\}}(\mu_t) = \mu_A$.[20]

With perfect information, a patient knows his exact h for each treatment so he simply chooses the highest h_t. In this scenario, the average treatment effect across all patients is

$$\bar{h}_2 = \mu_A\Phi(\eta) + \mu_B[1 - \Phi(\eta)] + \theta\phi(\eta),$$

where $\theta = \sqrt{\sigma_A^2 + \sigma_B^2 - 2\sigma_{AB}}$, $\eta = (\mu_A - \mu_B)/\theta$, and $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of a standard normal distribution respectively (Nadarajah and Kotz 2008). Note that $\bar{h}_2 \geq \bar{h}_1$, so patients are not worse off having perfect information about their individual treatment effect.[21]

The third scenario is patients receive a noisy signal of their individual treatment effect from t. For a patient who has treatment effect h_t from treatment t, that individual gets a signal $s_t \sim N(h_t, \sigma_s)$. In this scenario, patients choose the treatment with the highest signal s_t, and the average treatment effect across all patients is $\bar{h}_3 = \max_{t\in\{A,B\}}(s_t)$. Receiving an inaccurate signal means that a patient can choose a treatment with a lower treatment effect (h_t). Across the population, patients get a greater health impact with perfect information compared to either a noisy signal or no information ($\bar{h}_2 \geq \bar{h}_3$ and $\bar{h}_2 \geq \bar{h}_1$), but having a noisy signal does not necessarily make the patient better off than having no signal ($\bar{h}_1$ can be greater than, less than, or equal to $\bar{h}_3$).

As a result, the maximum potential health impact of personalized medicine in this market is

$$\Delta H^p = q\Delta h^p = q(\bar{h}_2 - \bar{h}_1)$$

With a noisy signal the maximum potential health impact of personalized medicine is

$$\Delta H^p = q\Delta h^p = q(\bar{h}_3 - \bar{h}_1)$$

where $(\partial\bar{h}_3/\partial\sigma_{AB}) \leq 0$ and $(\partial\bar{h}_3/\partial\sigma_s) \leq 0$. Therefore, the less correlated the different treatment outcomes and the less noise that a patient has about his treatment effect, the larger the health impact of personalized medicine for improving the matching of patients to treatments.

option does not provide an increase in health impact. If a generic uses a different formulation or delivery mechanism, then it would not necessarily be perfectly correlated with the branded version. But in this case they would not be AB-rated generics.

20. Throughout this section, I assume patients are risk neutral.

21. The max of two or more independently distributed normals generalizes to the Gumbel distribution, or Type 1 extreme value distribution, which for two standard normals has a mean of $1/\sqrt{\pi} \approx 0.56$ (Nadarajah and Kotz 2008).

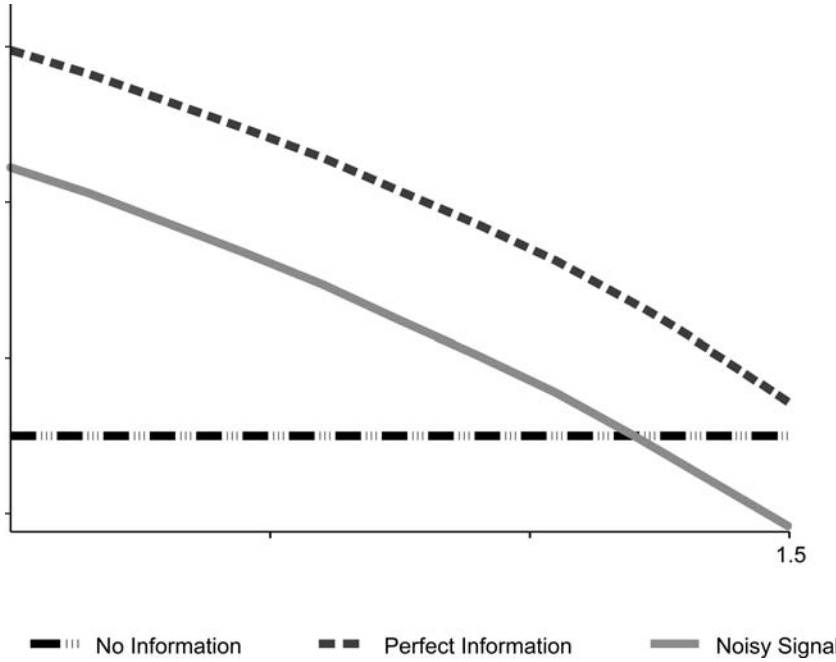

Fig. 7.3 Average treatment effect by σ_{AB} and patient signal
Source: Author calculations.

For this chapter the relevant comparison is how much personalized medicine can increase the total health impact compared to how patients and physicians choose treatments in the real world

$$\Delta H^{p} = q\left(\bar{h}_{3} - \bar{h}_{\text{actual}}\right). \tag{1}$$

To understand the effect of a patient's knowledge of his individual treatment effect, consider an example where the distribution of h across two treatments A and B is

$$(h_A, h_B) \sim N\left(\begin{bmatrix} 1.5 \\ 1 \end{bmatrix}, \begin{bmatrix} 1.5 & \sigma_{AB} \\ \sigma_{AB} & 1 \end{bmatrix}\right)$$

and patients receive a noisy signal of their individual treatment effect observe

$$s_t \sim N(h_t, 1)\ \forall t \in \{A, B\}.$$

Figure 7.3 illustrates the average treatment effect for different covariances between the two treatments (σ_{AB}) for the three scenarios discussed: patients have no information about their individual treatment effect, patients have full information about their individual treatment effect, and patients have a noisy signal of their individual treatment effect.

In this example, perfect knowledge of an individual's treatment effect increases the health impact by up to 33 percent relative to patients choosing the treatment with the highest average health impact. The largest increase in health impact comes when the treatments are uncorrelated and there is no health impact when the treatments are perfectly correlated. With a noisy signal, the increase in health impact with uncorrelated treatment effects drops to 23 percent and is negative with perfectly correlated treatment effects.

Impact of Individual Treatment Effect on Risk Assessment. The impact of risk assessment is similar to treatment effect of searching except the health impact of Treatment $A(h_A)$ comes from a multimodal normal distribution. This distribution represents the two possible outcomes that occur depending on whether the patient does not get the serious side effect (State 1) or the patient does get the serious side effect (State 2). Therefore

$$\mu_A = p\mu_{A1} + (1-p)\mu_{A2}$$

and

$$\sigma_A^2 = p\sigma_{A1}^2 + (1-p)\sigma_{A2}^2 + \gamma$$

where $\lambda = p(1-p)(\mu_{A1} - \mu_{A2})^2$ and $A1$ represents Treatment A in State 1 and $A2$ represents Treatment A in State 2.

Consider a patient choosing between Treatment A and Treatment B where a patient knows his individual treatment effect for each treatment in each state in the world such that $h_{A1} > h_B > h_{A2}$.[22] With no individual information about p, the probability a patient is in State 1 (no serious side effect) versus State 2 (serious side effect), the patient may have information about $\overline{p}$, the average share of patients in State 1 in the patient population. A patient then chooses Treatment A if

$$\overline{p} > \frac{h_B - h_{A2}}{h_{A1} - h_{A2}},$$

and Treatment B if the inequality holds in the other direction.[23] As a result, all patients choose either Treatment A or Treatment B based on $\overline{p}$ and the average health impact of the different treatments.

If a patient has information about his individual probability of getting the serious side effect, p_i, then he chooses Treatment A if

$$p_i > \frac{h_B - h_{A2}}{h_{A1} - h_{A2}},$$

and Treatment B if the inequality holds in the other direction. As a result, for the case when $\overline{p} > (h_B - h_{A2})/(h_{A1} - h_{A2})$, patients with p_i such that

22. It is straightforward to adapt this example to the case where the patient has no information or noisy information about his health impact and makes his choice based on either the treatment average across the population (μ) with no information or the treatment signal (s) with noisy information.

23. The patient chooses Treatment A if $\overline{p}h_{A1} + (1-\overline{p})h_{A2} > h_B$.

$$\overline{p} > \frac{h_B - h_{A2}}{h_{A1} - h_{A2}} > p_i$$

would choose Treatment A in the case of no information and Treatment B in the case of full information about p.[24] This patient is better off in the case of full information by $p_i h_{A1} + (1 - p_i)h_{A2} - h_B > 0$. Summing over all patients the increase in health effect is

$$\Delta H^p = \sum_{i \in \left\{p_i < \frac{h_B - h_{A2}}{h_{A1} - h_{A2}}\right\}} p_i h_{A1} + (1 - p_i)h_{A2} - h_B.$$

7.4 Case Study in MS

Multiple sclerosis is a good case study for understanding the value of innovations in personalized medicine because there is profound heterogeneity in the MS population, disease course, and treatment response (Lucchinetti et al. 2000). Multiple sclerosis is a chronic condition that occurs when the body's immune system attacks the central nervous system and damages or destroys the nerves' protective covering, causing flare-ups that range from dizziness to paralysis and cognitive loss.[25] There are currently more than 400,000 patients with MS in the United States with almost $14 billion in annual spending on MS treatments, which makes it the fourth largest specialty pharmacy class in the United States.[26]

Currently, physicians can rely on clinical trials data, biomarkers, and passive searching to determine the best course of treatment. Clinical trials data in MS is useful at the group level, but it is viewed as insufficient to influence individual treatment decisions and "few biomarkers have made their way into clinical practice" in MS (Derfuss 2012). As a result, there is very little predictive power about how a patient will respond to an individual treatment (Derfuss 2012). Passive searching, while frequently used, is costly because, as previously discussed, it can cause irreversible damage and disability, increase disease progression, increase future side effects, and increase the probability that a patient will be unresponsive to alternative treatments.

There are twelve disease-modifying therapies (DMTs) available in the United States to treat MS, which are listed in appendix table 7A.1. The purpose of these treatments is to reduce the number of flare-ups that patients suffer, but they do not cure the underlying disease. These treatments can broadly be categorized in two ways: by line of treatment and mode of admin-

24. The case when $\overline{p} < (h_B - h_{A2})/(h_{A1} - h_{A2})$ and $\overline{p} < (h_B - h_{A2})/(h_{A1} - h_{A2}) < p_i$ is symmetric.

25. There are four disease courses for MS patients, clinically isolated syndrome (CIS), relapsing-remitting MS (RRMS), primary progressive MS (PPMS), and progressive relapsing MS (RPMS). Eighty-five percent of MS patients have RRMS, which is the focus of this case study (Trapp and Nave 2008).

26. See Pietrangelo and Higuera (2016) and IMS (2016).

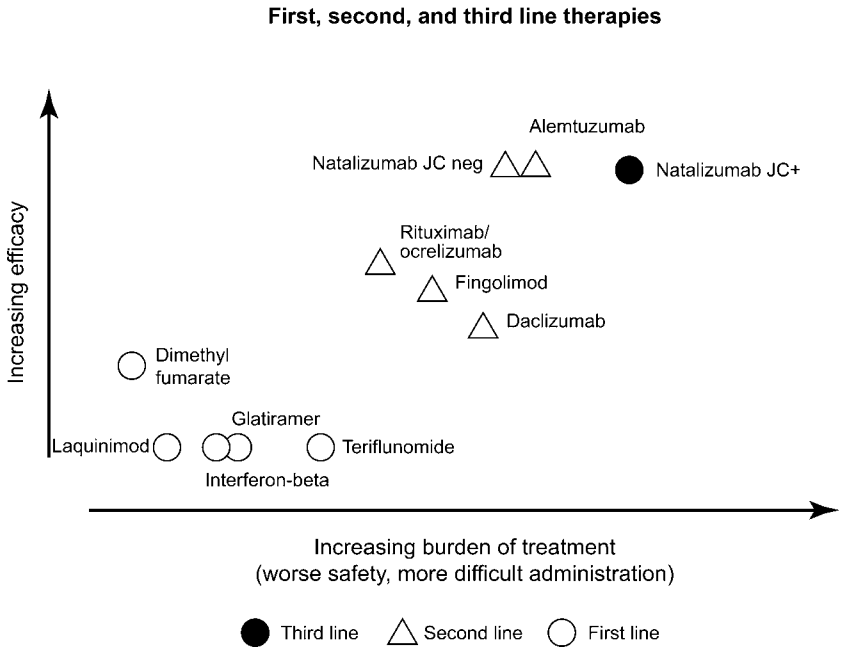

Fig. 7.4 Market for MS treatments
Source: Coles (2015) with author edits.

istration. Figure 7.4 shows the typical line of treatment for each DMT, where first-line treatments are safer treatments with lower efficacy and lower treatment burden, second-line treatments are more aggressive treatments that feature higher efficacy but also higher treatment burden, and third-line treatments feature the highest efficacy but also have potentially life-threatening side effects. Most of the first-line treatments are referred to as ABCRE treatments, which represent Avonex, Betaseron, Copaxone, Rebif, and Extavia.

The other way the market is divided is by the mode of administration. There are three modes of administration: injection, infusion, and oral. The ABCRE treatments are all injectable (either with intramuscular or subcutaneous injection), and injectables were the only option from 1993 to 2004.

In 2004 Tysabri, a more efficacious treatment that is administered through infusion, was introduced. Tysabri plays an important role in the MS market because it is not only the most efficacious treatment, but it has been linked to a rare and highly fatal brain disease, PML. Tysabri was approved by the FDA in 2004 as the first infusion treatment and was almost six times more efficacious than any existing treatment. By February 2005, the treatment was withdrawn from the market after three patients developed PML. In February 2006, Tysabri returned to the market with conditions including mandatory patient registration in a database, follow-ups every six months,

and magnetic resonance imaging (MRI) evaluation prior to initiation. In 2010, oral treatments were introduced that reduced the burden of treatment administration.

7.4.1 Data

There are three main types of data necessary to estimate the health impact of MS treatments: the distribution of QALYs for each treatment t (previously denoted as μ_t and σ_t), the covariance in treatment outcomes between treatments ($\sigma_{t_1 t_2}$), and the patient count estimates for each treatment (q_t). For my analysis, I measure the market size as the number of patients on treatment. The data appendix provides additional details about the data.

The QALYs estimates are taken from published clinical studies, most of which are summarized in the Tufts Medical Center Cost-Effectiveness Analysis Registry (CEAR).[27] The CEAR includes over 4,800 pharmaceutical cost-utility analyses in the peer-reviewed medical literature. It is intended to be a comprehensive data set of all cost-utility articles analyzed by trained professionals, who rate the quality of the study and provide information about the quality level and quality relative to the standard of care found in the study. Of the twenty-four MS studies that use MS DMTs, I rely on the fifteen that are for relapsing-remitting MS (RRMS) patients (which composes 85 percent of all MS patients).[28] The CEAR rates the studies on a scale from 1 to 7 depending on the quality of the analysis. All of the MS clinical studies used from the CEAR data set have a rating above average. The efficacy measures are relative to a patient taking no DMT, so a QALY of zero means that the treatment provides no benefit relative to not taking any DMT.

One consideration when using clinical trials data is that the distribution of QALYs differs from how patients use the treatment in the real world. Since the uniformity of a clinical trial is likely to reduce the variance of treatment effects, the use of clinical trials data would likely underestimate each treatment variance in the real world and, therefore, underestimate the potential health impact of personalized medicine.

7.4.2 Covariance Estimation

The estimates of the covariance between treatment outcomes are more difficult to measure because clinical studies generally provide information about how a patient responds to one treatment, not how each patient responds to multiple treatments.[29] However, there are observational studies

27. I assume the distributions of QALYs from clinical studies is equal to distributions of all patients in the disease category, that the QALY measure incorporates all side effects as well as treatment efficacy, and that QALY measures incorporate adherence and are not conditional on adherence.

28. RRMS is relapsing-remitting MS.

29. See Basu, Jena, and Philipson (2011) for a discussion of the effect of the joint distribution of treatment effects.

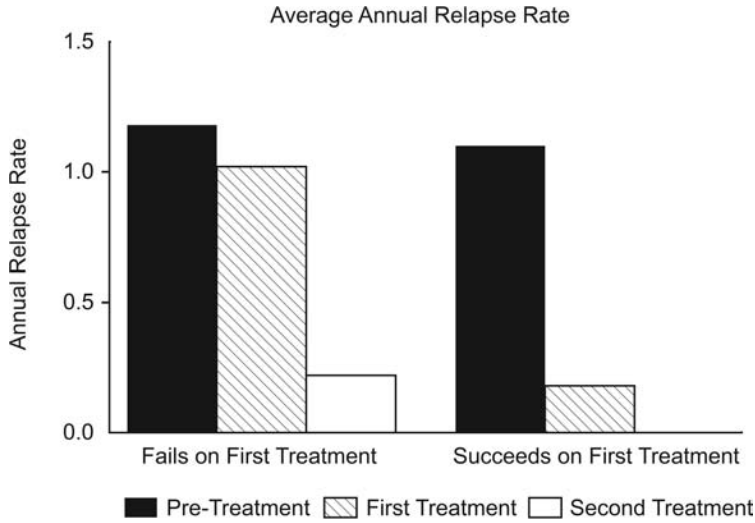

Fig. 7.5 Correlation example for MS treatments
Source: Río et al. (2012).

that measure the treatment effect of patients before and after a treatment failure, which is when a patient experiences a suboptimal response to a treatment and switches to a substitute treatment.[30] These studies show how a patient responded to two different treatments conditional on the patient failing at least one of the treatments, but do not show how many patients would have been successful on both treatments.[31]

I estimate the covariance between treatment impacts using two observational studies of patients who switch therapies (Gajofatto et al. 2009; Río et al. 2012).[32] Figure 7.5, from Río et al. (2012), illustrates the type of data that can be used to construct a covariance matrix and why the covariance is important for understanding the potential health impact of personalized medicine. This figure shows how average annual relapse rates, which are unpredictable acute attacks that are a main symptom of MS, vary for patients who fail or succeed on their first treatment. For patients that fail on their first treatment, their relapse rate on their first treatment is not much different from not being on any MS treatment at all. However, the relapse rate on the second treatment is similar to the relapse rate for patients who succeed on their first treatment. If treatment effects were highly correlated

30. See, for example, Río et al. (2012) and Gajofatto et al. (2009).

31. A patient's response to the second treatment may be affected by the first treatment. For example, I previously discussed the effect of neutralizing antibodies that could be produced during the first treatment and make the second treatment less effective.

32. The covariance could also be estimated using a micro-level data set that tracks a patient's treatment and response to treatment such as the Sylvia Lawry Centre MS Patient Database.

Table 7.1 **Failure correlation calculation**

First treatment	Result on first treatment	Second treatment	Result on second treatment	Share of patients (%)
Copaxone	Failure	Interferon	Success	6
Copaxone	Failure	Interferon	Failure	1
Copaxone	Success	Interferon	Success	12
Copaxone	Success	Interferon	Failure	1
Interferon	Failure	Copaxone	Success	2
Interferon	Failure	Copaxone	Failure	2
Interferon	Success	Copaxone	Success	38
Interferon	Success	Copaxone	Failure	38
				100
Interferon A	Failure	Interferon B	Success	16
Interferon A	Failure	Interferon B	Failure	9
Interferon A	Success	Interferon B	Success	48
Interferon A	Success	Interferon B	Failure	26
				100

across treatments, the patient who failed on their first treatment would be very likely to fail on their second treatment. The less correlated treatment effects are across treatments, the more differentiated a patient's treatment effect would be between the first and second treatment. This figure suggests that treatment effects for MS treatments are not very correlated and, therefore, there is potential for personalized medicine to help patients by matching them to the treatment on which they are more likely to succeed.

I estimate two covariances with this data. First, I estimate the covariance in treatment outcomes between a patient on two different interferons. Second, I estimate the covariance in treatment outcomes between a patient on an interferon and Copaxone. I do not have data on oral treatments so I assume that oral treatments have the same covariance as an interferon with Copaxone. Since I do not have the patient's treatment impact in QALYs, I use the covariance between treatment failure as a proxy. Although treatment failure is likely to be highly correlated with treatment impact, treatment failure outcomes are likely to be more correlated than health impacts within a patient due to the dichotomous nature of treatment failure. A higher correlation would understate the potential impact of personalized medicine.

Table 7.1 shows the share of the 597 patients from Gajofatto et al. (2009) that fall into each combination of treatment pair and treatment result (failure/success). I observe aggregate counts of patients who fail or not on their first treatments. However, I do not observe what a patient who has a success with their first treatment would do on a second treatment. As a result, I assume that the probability of success on Treatment A given the success of Treatment B is proportional to the probability of failure on Treatment A given failure of Treatment B.

The resulting covariances are listed in table 7.2.

Table 7.2 **Failure correlation calculation**

Treatment 1	Treatment 2	Correlation	Covariance
Interferon A	Interferon B	0.13	0.03
Interferon	Copaxone	0.10	0.02

There are several important considerations when estimating the covariance. The first consideration is disease progression. With MS, as with many diseases, the patient's disease may have progressed between the first and second treatment; since the correlation measures the same patient at different points of time, it could impact the measurement of the covariance matrix. Ideally, the covariance matrix would measure the treatment impact of two treatments at the same time, which is not possible to observe with most treatments. While there will be some disease progression between the two treatments in my covariance matrix calculation, the estimation is restricted to patients who are taking two first-line treatments with similar efficacy and side effects and, therefore, would not be likely to have had significant disease progression. The limitation of this restriction is that I assume that patients' alternative treatment choices are restricted to similar treatments. This assumption would underestimate the potential health impact of sorting because it could be valuable for patients on first-line treatments to switch to treatments that differ significantly in efficacy and side effects.

Another consideration is the representativeness of the patient behavior in the observational study. The studies provide breakdowns of the patient characteristics included in the sample (include gender, race, age, age of onset, and annual relapse rate). These data are consistent with the characteristics of the MS population as a whole, but it would require a more detailed data set of MS patients to understand any potential bias in the patient pool.[33]

I present a sensitivity analysis around the covariance matrix in the next section.

7.4.3 Health Impact

I measure the actual or potential health impact of seven events in the history of MS treatments: (a) the innovation of Betaseron, the first MS DMT; (b) the innovation of the other ABCRE DMTs; (c) the potential impact of improved matching between ABCRE treatments; (d) the innovation of oral DMTs; (e) the potential impact of improved matching between oral DMTs; (f) the innovation of infusion DMTs; and (g) the potential benefit of risk assessment for Tysabri.

33. For example, the average age of onset in the observational data is thirty-one and the female-to-male ratio is 1.4. Both are consistent with the MS population as a whole (see, e.g., https://www.healthline.com/health/multiple-sclerosis/facts-statistics-infographic).

1. Innovation of Betaseron. Betaseron, approved in 1993, was the first DMT for MS. As shown in table 7.5, Betaseron provides patients with 0.34 QALY relative to no DMT, has an adherence rate of 52 percent, and has a market share of 10 percent (or roughly 23,400 patients per year).

To measure the health impact of Betaseron, I compare the market for MS with no DMT and a but-for world where Betaseron is the only DMT. I assume that in this but-for world all interferon patients and 63 percent of Copaxone patients (the share of actual Copaxone patients that are tolerant of interferon treatments) would be on Betaseron (Bergvall et al. 2014).[34] As a result, the introduction of Betaseron provided 0.34 QALYs of treatment for 76 percent of the market (or 178,000 patients) for a total health impact of roughly 61,000 QALYs.

2. Innovation of other ABCREs. After Betaseron's entry into the market, the other ABCRE treatments (Avonex, Copaxone, Rebif, and Extavia) hit the market between 1996 and 2009. The introduction of these treatments had several effects. First, they expanded the market by the 21 percent of the market (or 63 percent of actual Copaxone patients) who could take Copaxone, but not an interferon. Second, the introduction of the other ABCRE treatments made higher efficacy and adherence treatments available. For instance, Copaxone, with an efficacy of 0.41 QALY and an adherence rate of 55 percent, has a higher efficacy and adherence rate than Betaseron.

To determine the increase in health impact from the other ABCRE treatments, I measure the share of patients who failed or did not fail on treatment.[35] Failure is defined by either switching to a different first-line treatment or switching to a second-line treatment after being on treatment for less than two years. I assume that patients who failed their treatment received a health impact similar to a patient who was randomly assigned a treatment (h_1), and a patient who did not fail the treatment received a health impact similar to a patient who was assigned the optimal treatment (h_2).

The introduction of the ABCREs increased the health impact of MS treatments by 22,000 QALYs or a 36 percent increase in the total health impact.

3. Potential of ABCRE Heterogeneity. Using the distribution of health impact for each of the treatments (shown in table 7.3) and the covariance table, I estimate the potential impact of personalized medicine to match patients to their highest individual treatment effect across the different ABCRE treatments.

I assume that the individual health impact from Avonex and Rebif (which are both interferon beta-1a) and from Betaseron and Extavia (which are both interferon beta-1b) are perfectly correlated because they are the same

34. There are four interferon treatments, Avonex, Betaseron, Extavia, and Rebif, which together comprise 55 percent of the market.

35. These shares are taken from Gajofatto et al. (2009).

Table 7.3 **Distribution of health impact of ABCRE treatments**

	Mean	Standard deviation
Avonex	0.20	0.08
Betaseron	0.34	0.14
Copaxone	0.41	0.20
Extavia	0.34	0.14
Rebif	0.20	0.08

molecule. Since Avonex and Rebif have different modes of administration, this assumption provides a conservative estimate of the potential impact of heterogeneity.

To calculate the two counterfactual health impacts (the health impact of patients being randomly sorted on treatments and the health impact of patients sorted to their highest health impact treatment), a simulated group of patients the size of the actual market receives a health impact for each treatment according to a joint normal distribution with the parameters listed in table 7.3 and the covariance table. The simulated patients are then either assigned a random treatment or the treatment that corresponds to their highest health impact.

The health impact of the actual distribution of patients across treatments increases the health impact by 63 percent relative to patients being randomly distributed across treatments. The maximum potential health impact given this distribution provides an 18 percent or 14,000 QALY increase compared to how patients and physicians choose treatments in the real world.

4. Innovation of Oral Treatments. The three oral treatments, Aubagio, Gilenya, and Tecfidera, entered the market between 2010 and 2013. These treatments altered the MS landscape by offering an alternative form of treatment administration. They also offered improved efficacy over the existing ABCRE treatments for early line patients. The oral treatments expanded the market by the 6 percent of oral treatment patients that were not on any MS treatment before taking an oral treatment.[36] In addition, the oral treatments increased the maximum health impact by a first-line treatment by 0.19 QALY.

As a result, the introduction of the oral treatments increased the health impact by 18 percent over ABCRE treatments, which resulted in an increase of 15,000 QALYs.

5. Potential of Oral Treatment Heterogeneity. As with the ABCRE treatments, properly matching patients with treatments considering patient heterogeneity has a potential to increase the health impact of treatments. The distribution of health impact across the oral treatments is listed in table 7.4.

The health impact of the current distribution of patients across treat-

36. See MS in America (2014).

Table 7.4 Distribution of health impact of oral treatments

	Mean	Standard deviation
Gilenya	0.60	0.27
Aubagio	0.32	0.34
Tecfidera	0.59	0.23

ments increases the health impact by 100 percent relative to patients being randomly distributed across treatments.[37] The maximum potential health impact given this distribution provides an 18 percent increase over the current distribution. As a result, personalized medicine that improves the ability of patients to identify the oral treatments with the highest impact can improve the average health impact of these patients by 17,700 QALYs.

6. Innovation of Infusion Treatments. The innovation of infusion treatments, especially Tysabri, not only brought a new form of treatment administration to the market but also an increase in efficacy. The infusion treatments increased the market by 3 ppts, since 35 percent of infusion patients were new to the market (in other words, they were not patients that would have been on another treatment in the absence of the infusion treatments) and infusion treatments comprise 9 percent of the market (Biogen 2008). In addition, the infusion treatments provide 1.70 QALY over the next highest treatment in terms of efficacy.

As a result, the infusion treatments increased the total health impact by 60 percent or almost 58,000 QALYs.

7. Potential of Tysabri Risk Assessment. As discussed previously, Tysabri not only brought an increase in efficacy but also the potential for very serious side effects.[38] Tysabri's PML side effect was not known at the time of the FDA approval. Instead, the treatment was on the market for almost three months when Biogen, the maker of Tysabri, learned about one confirmed and two suspected cases of PML. As a result, Tysabri was temporarily pulled from the market until it was allowed to be reintroduced to the market roughly one and a half years after learning about the PML side effect.

Since the PML side effect was learned after Tysabri was on the market, I back out the effect that PML has on Tysabri consumption to measure the potential effect of a PML diagnostic test. First, prior to learning about PML, industry analysts expected Biogen sales to exceed 87,000 patients per year, which amounts to over 40 percent of market share.[39] Second, when

37. Health impact under a random treatment distribution and maximum potential health impact are calculated using the same methodology that was discussed in the subsection on ABCRE heterogeneity, but using the parameters in table 7.3.

38. It was recently discovered that Tecfidera also poses a PML risk (Van Schependom et al. 2016).

39. *Wall Street Journal* (2005). Tysabri was expected to sell over $2 billion per year at $23,000 per year. For a market with 210,000 patients, which is roughly the market size in 2009, this would amount to over 40 percent of the market.

information about Tysabri's link to PML came out, Biogen's stock dropped 44 percent or $10 billion, which is consistent with an expected market share of Tysabri above 33 percent.[40] Third, before the information about Tysabri's link to PML was known, industry projection models predicted that Tysabri would have a market share that rose from 15 percent in 2005 (the treatment's first full year on the market) and would stay around 35 percent through 2015.[41] Finally, these estimates are consistent with an estimate based on physician perceptions. If the only patients in the current market that are prescribed Tysabri are patients with physicians who feel the benefits of Tysabri outweigh the costs (roughly 65 percent of physicians) and are JCV negative (55 percent of patients), then Tysabri's 9 percent market share would be over 25 percent in a world with perfect information about a patient's PML risk.[42]

All of these examples suggest that Tysabri would have a market share between 25 and 40 percent of the market with a diagnostic test that provides perfect information about a patient's PML risk. To be conservative and to allow for the introduction of other treatments that were not on the market in 2005, I assume Tysabri would have a 20 percent market share if it did not have any PML side effect. By comparison, the number of PML cases in the United States from 2005 to 2015 was 165.[43]

As a result, a perfect PML diagnostic test that could correctly identify the PML side effect would have allowed over 12 percent of the MS market to take Tysabri while restricting it to the hundreds of patients that were subjected to PML.[44] The health impact of putting 12 percent of the market that is not at risk for PML onto Tysabri relative to the treatment with the next highest health impact (0.60 for Gilenya compared with 2.30 for Tysabri) would increase the total health impact by almost 47,000 QALYs or 30 percent from the current market.[45]

Although conservative in the 20 percent market share, this estimate serves as an upper bound for a PML diagnostic test since the diagnostic test would

40. See http://www.fool.com/investing/high-growth/2005/03/08/after-the-crash-is-biogen-idec-a-buy.aspx. If Avonex was responsible for the entire $12 billion remaining market share, had 40 percent market share, and had a nearly identical price to Tysabri, this suggests that Tysabri's market share would be in excess of 33 percent market share.

41. http://www.fool.com/investing/small-cap/2004/12/07/spin-the-medicine-bottle.aspx.

42. http://i.bnet.com/blogs/tysabri-confidence-survey_figure-2.jpg.

43. http://wasmain.nationalmssociety.org/site/DocServer/PML._MS_Summit_2015.pdf?docID=75816.

44. There is already a diagnostic test on the market for JCV. In a step toward incorporating personalized medicine into MS treatments, in 2012, the FDA approved the Stratify JCV Antibody ELISA test, which helps identify patients who are more prone to PML. This diagnostic test tells if a patient is anti-JCV antibody positive or negative. If the patient is anti-JCV antibody negative, they have a lower than 1 in 1,000 risk of developing PML. If the patient if anti-JCV antibody positive, that risk is between 6 and 13 in 1,000 depending on prior treatments. However, 70 to 90 percent of the population has the JCV virus, so the test is not very informative about a patient's actual risk factors (Holland and Nall 2016). However, this test was not on the market for most of the period of interest so the vast majority of MS patients did not have access to JCV diagnostic testing before taking Tysabri.

45. Lemtrada is currently the second-highest treatment on the market, but it has not been on the market long so it would not have a significant impact during the 2005 to 2015 time period.

Table 7.5 Health impact by type of innovation as a share of total health impact

Novel and incremental innovation	(%)
Innovation of Betaseron	26
Innovation of ACRE	6
Innovation of oral treatments	6
Innovation of infusion treatments	25
Total	64
Potential impact of personalized medicine	
Within ABCRE treatments	9
Within oral treatments	6
Potential of Tysabri risk assessment	21
Total	36

not perfectly sort patients. The health impact of an actual diagnostic test would depend on its accuracy.

Breakdown. Table 7.5 breaks down what share of the total health impact discussed in the previous sections come from each of the seven events.

This breakdown shows that personalized medicine events in MS have the potential to increase the health impact of treatments by over 50 percent (= 0.34/0.66). The potential health impact of personalized medicine is split between improving the matching process of patients to treatment through the individual treatment effect and risk assessment for serious side effects.[46]

Betaseron had the largest impact, even in large part because it was the first treatment on the market, and the infusion treatments had the second largest impact because they had the highest efficacy. The potential impact of a Tysabri risk assessment shows that Tysabri would have by far the largest health impact if the PML risk were better identified.

The impact of perfectly sorting patients on both ABCRE and oral treatments is roughly equivalent to the impact of the seven treatments (Avonex, Copaxone, Rebif, Extavia, Gilenya, Aubagio, and Tecfidera).

7.5 Cost Considerations in Personalized Medicine Innovation

Up to this point, I have focused on the health impact of personalized medicine. In this section, I discuss potential cost considerations for personalized medicine, including the cost of development and the cost to patients.

46. To address the sensitivity of the results with respect to the covariance matrix, I include the results with no correlation and perfect correlation. With perfect correlation, there would be no potential health impact of sorting within ABCRE treatments and within oral treatments. As a result, the potential health impact of personalized medicine would all come from the potential of Tysabri risk assessment and would increase health impact by 30 percent. With no correlation between treatment effects, the potential impact of personalized medicine would increase by a percentage point for both within ABCRE treatments and within oral treatments. As a result, the potential of personalized medicine would increase the health impact of treatments by 55 percent.

In the multiple sclerosis example, there are two types of personalized medicine being considered: a diagnostic blood test and improved sorting of patients on existing treatments. Development costs for a diagnostic test are generally between $250 million and $300 million (or 10 to 12 percent of the cost of a new treatment; McKinsey and Company [2013]). Diagnostic blood tests are also significantly cheaper to consumers, generally costing less than $1,000 in a one-time fee compared to the roughly $60,000 per year cost of MS treatments (Hartung et al. 2015). Since patients spend around six years on a given treatment (Río et al. 2012), these results suggest that the diagnostic test for Tysabri costs around 0.3 percent of the total cost of Tysabri for the average patient.

Innovations that improve patient sorting by predicting a patient's response to a treatment are harder to characterize. Innovations that use data on patient characteristics are likely to be significantly cheaper to develop. For example, the Rio Score is a scoring system that combines patient characteristics, including clinical and MRI parameters, to predict whether a patient will fail on a treatment (Río et al. 2012).[47] After one year of therapy, 92 percent of patients with a Rio Score of 2 or 3 failed on their treatment, while 8 percent of patients with a Rio Score of 0 or 1 failed their treatment (Hyun et al. 2015). Tests that rely on patient data are relatively inexpensive to develop and inexpensive to implement. The main cost to these types of innovations is likely informing patients and consumers.

The multiple sclerosis example may not be representative of other types of personalized medicine. Many forms of personalized medicine use genetic, epigenetic, and protein biomarkers and require development costs that are similar or slightly below the development costs of standard treatments (see discussions in chapter 5, this volume, and Gupta et al. 2004). Understanding the cost impact of these personalized treatments for the patient is still speculative, but due to the targeted nature of these treatments they are likely to have higher costs to patients than more traditional treatments.

7.6 Conclusion

The potential of personalized medicine comes from its ability to either create treatments that address the heterogeneity across patients or its ability to provide information to patients that can improve the health impact of existing treatments. This chapter explores the potential magnitude of the latter effect for MS treatments.

I find that several factors influence the health impact of personalized medicine. Personalized medicine has a greater potential health impact when treatment effects are less correlated across treatments, the variance of the

47. The Rio Score is the count of how many of the following conditions are met: (a) more than two active T2 lesions on an MRI, (b) at least one relapse, and (c) an increase of EDSS score by at least 1 point sustained over at least six months. Failure was defined as having any of the following: switched therapy due to failure, clinical relapse, or EDSS progression.

distribution of health impacts is larger, there is less noise in an individual's signal of their treatment effect, and there are more treatment options.

These results suggest that there is significant potential for personalized medicine in MS due to the heterogeneity in the MS population, disease course, and treatment response and twelve DMTs that vary in efficacy and administration. I find that personalized medicine has the potential to increase the health impact of MS patients by over 50 percent.

One extension of this work is understanding the value of me-too innovations or evergreening, which are innovations that are considered to be slight modifications of existing treatments. The conventional wisdom is that these innovations provide little to no value and waste resources (see, e.g., Collier 2013). With personalized medicine, me-too innovations can provide a health impact even if they have a lower average treatment effect than similar existing products if the treatment effects are not well correlated across treatments. This result suggests that me-too innovations are more valuable in a world with personalized medicine.

There are several areas for future research in personalized medicine. First, it would be valuable to gain a better understanding of the research and development (R&D) costs of personalized medicine in order to measure the productivity of innovations in personalized medicine. This research would inform whether personalized medicine has a higher rate of return on R&D compared to other types of medical innovations and which types of personalized medicine have the highest rates of return.

Second, it is important to understand why there is not more innovation in personalized medicine if, as this chapter suggests, it has so much potential health impact. Two potential explanations are that innovations in personalized medicine, like diagnostic tests, are difficult to create and that firms do not have an incentive to engage in personalized medicine innovation because they are not able to capture the gains from the innovation. This chapter touches on the second point. Some forms of innovation redistribute patients between treatments, which may have lower returns on innovation for a firm since they may only capture a share of the redistributed patients, while other forms of innovation, like a diagnostic test, are more likely to be tied to a specific treatment and, therefore, allow the innovating firm to capture the gains from the innovation.

Third, future research could use more patient-level data to understand how and why patients switch between treatments. This research would allow for a more nuanced model of how patients choose and move between treatments, and how personalized medicine can improve the matching process between patients and treatments. This research could also provide more empirical research on the correlation between treatment effects.

Finally, there are currently eight established biomarkers and at least six potential biomarkers in MS, and it would be valuable to understand how much health impact could be gained if these biomarkers could be more effectively integrated into determining individual treatment effects (Derfuss 2012).

Appendix A

Summary of MS Treatments

Table 7A.1 **Summary of MS treatments**

Treatment	Company	Molecule	Year of FDA approval	Mode of administration	Health impact (QALY)	Standard deviation of health impact	Adherence (%)	Market share (2005–2015) (%)
Betaseron	Bayer	Interferon beta-1b	1993	Subcutaneous injection	0.34	0.14	52	10
Copaxone	Teva	Glatiramer acetate	1996	Subcutaneous injection	0.41	0.20	55	33
Avonex	Biogen	Interferon beta-1a	1996	Intramuscular injection	0.20	0.08	62	28
Rebif	Serono/Pfizer	Interferon beta-1a	2002	Subcutaneous injection	0.20	0.08	59	17
Tysabri	Biogen	Natalizumab	2004	Infusion	2.30	—	75	8
Extavia	Novartis	Interferon beta-1b	2009	Subcutaneous injection	0.34	0.14	52	0
Gilenya	Novartis	Fingolimod	2010	Oral	0.60	0.27	55	0
Aubagio	Sanofi	Teriflunomide	2012	Oral	0.32	0.34	55	0
Tecfidera	Biogen	Dimethyl fumarate	2013	Oral	0.59	0.23	55	4
Lemtrada	Sanofi	Alemtuzumab	2014	Infusion	1.54	—	93	1
Plegridy	Biogen	Pegylated interferon beta-1a	2014	Subcutaneous injection	0.20	0.08	62	0
Glatopa	Sandoz	Glatiramer acetate	2015	Subcutaneous injection	0.41	0.20	55	0

Sources: See appendix B.

Appendix B

Data Summary

Health Impact/Efficacy Data. For the efficacy measurement, I mainly use the Tufts Medical Center Cost-Effectiveness Analysis Registry (CEAR) from the Center for the Evaluation of Value and Risk in Health. The CEAR includes over 4,800 pharmaceutical cost-utility analyses in the peer-reviewed medical literature. It is intended to be a comprehensive data set of all cost-utility articles analyzed by trained professionals, who rate the quality of the study and provide information about the quality level and quality relative to the standard of care found in the study. The data set lists the drug's name or active ingredient; the drug's disease class, which can be uniquely mapped into my nineteen disease classes; and the year of the study. The data set includes fifteen studies that list the QALY of treatments for all ABCRE treatments and Tysabri. I take the average across studies for treatments that have multiple studies. For the oral treatments, I use estimates from Pistoresi (2015) and for Lemtrada, I use an estimate from the Scottish Medicines Consortium (2014).

The estimates of standard deviations are taken from estimates in Prosser et al. (2004) and Pistoresi (2015).

Adherence Data. Adherence is a measure of whether patients are taking their treatment as prescribed and with the proper frequency. A patient is generally defined to be adherent if he possesses medication for at least 80 percent of the time they are active on treatment. I get adherence estimates from published studies in medical journals. Specifically, for all ABCRE treatment, I use adherence estimates from Halpern et al. (2011), which estimates adherence rates from 6,680 MS patients from 2000 to 2008 on ABCRE treatments. For oral and infusion treatments, I use estimates from Dionne et al. (2015), which compares adherence rates for 209 MS patients. These rates are generally consistent with those found in other published studies including Treadaway et al. (2009), Devonshire et al. (2011), and Reynolds et al. (2010).

Patient Count Data. I take patient count published estimates from Symphony Health Solutions.[48] Since these data are in revenues, I convert them to patients using cost estimates from Hartung et al. (2015). I supplement this data with estimates from Biogen documents, the producer of Avonex, Plegridy, Tecfidera, and Tysabri, published by the SEC, which primarily use IMS data.[49]

48. See http://symphonyhealth.com/wp-content/uploads/2013/06/Tecfidera.inThought.4Mar.pdf.

49. See Biogen (2008).

References

Basu, Anirban. 2013. "Personalized Medicine in the Context of Comparative Effectiveness Research." *Forum for Health Economics and Policy* 16 (2): S73–86.

Basu, Anirban, Anupam B. Jena, Dana P. Goldman, Tomas J. Philipson, and Robert Dubois. 2014. "Heterogeneity in Action: The Role of Passive Personalization in Comparative Effectiveness Research." *Health Economics* 23 (3): 359–73.

Basu, Anirban, Anupam B. Jena, and Tomas J. Philipson. 2011. "The Impact of Comparative Effectiveness Research on Health and Health Care Spending." *Journal of Health Economics* 30 (4): 695–706.

Basu, Anirban, and David Meltzer. 2007. "Value of Information on Preference Heterogeneity and Individualized Care." *Medical Decision Making* 27 (2): 112–27.

Bergvall, Niklas, Charles Makin, Raquel Lahoz, Neetu Agashivala, Ashish Pradhan, Gorana Capkun, Allison A. Petrilla, Swapna U. Karkare, Catherine Balderston McGuiness, and Jonathan R. Korn. 2014. "Relapse Rates in Patients with Multiple Sclerosis Switching from Interferon to Fingolimod or Glatiramer Acetate: A US Claims Database Study." *PLoS ONE* 9 (2): e88472.

Biogen. "Biogen Idec Q1 2008 Earnings Conference Call and Webcast, April 23rd, 2008." http://library.corporate-ir.net/library/14/148/148682/items/289373/Q108BIIB.pdf.

Center for the Evaluation of Value and Risk in Health. The Cost-Effectiveness Analysis Registry, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center. Accessed on Sept. 3, 2016. www.cearegistry.org.

Chin, Lynda, Jannik N. Andersen, and P. Andrew Futreal. 2011. "Heterogeneity in Action: The Role of Passive Personalization in Comparative Effectiveness Research." *Nature Medicine* 17 (3): 297–303.

Coles, Alasdair. 2015. "Newer Therapies for Multiple Sclerosis." *Annals of Indian Academy of Neurology* 18 (suppl. 1): S30.

Collier, Roger. 2013. "Drug Patents: The Evergreening Problem." *Canadian Medical Association Journal* 185 (9): E385–86.

Derfuss, Tobias. 2012. "Personalized Medicine in Multiple Sclerosis: Hope or Reality?" *BMC Medicine* 10 (1): 1.

Devonshire, V., Y. Lapierre, R. Macdonell, C. Ramo-Tello, F. Patti, P. Fontoura, L. Suchet, R. Hyde, I. Balla, E. M. Frohman, and B. C. Kieseier. 2011. "The Global Adherence Project (GAP): A Multicenter Observational Study on Adherence to Disease-Modifying Therapies in Patients with Relapsing-Remitting Multiple Sclerosis." *European Journal of Neurology* 18 (1): 69–77.

Dionne, Caitlin A., Rik Ganguly, Ann Camac, and Claudia Chaves. 2015. "Do Oral Disease Modifying Agents (DMTs) Improve Adherence to MS Treatment? A Comparison of Oral and Injectable Drugs." https://cmsc.confex.com/cmsc/2015/webprogram/Paper3275.html.

Egan, Mark, and Tomas J. Philipson. 2014. "Health Care Adherence and Personalized Medicine." NBER Working Paper no. 20330, Cambridge, MA.

Espinoza, Manuel A., Andrea Manca, Karl Claxton, and Mark J. Sculpher. 2014. "The Value of Heterogeneity for Cost-Effectiveness Subgroup Analysis: Conceptual Framework and Application." *Medical Decision Making* 34 (8): 951–64.

Gajofatto, A., P. Bacchetti, B. Grimes, A. High, and E. Waubant. 2009. "Switching First-Line Disease-Modifying Therapy after Failure: Impact on the Course of Relapsing-Remitting Multiple Sclerosis." *Multiple Sclerosis Journal* 15 (1): 50–58.

Goldman, Dana P., Charu Gupta, Eshan Vasudeva, Kostas Trakas, Ralph Riley, Darius Lakdawalla, David Agus, Neeraj Sood, Anupam B. Jena, and Tomas J.

Philipson. 2013. "The Value of Diagnostic Testing in Personalized Medicine." *Forum for Health Economics and Policy* 16 (2): S87–99.
Grutters, Janneke P. C., Mark Sculpher, Andrew H. Briggs, Johan L. Severens, Math J. Candel, James E. Stahl, Dirk De Ruysscher, Albert Boer, Bram L. T. Ramaekers, and Manuela A. Joore. 2013. "Acknowledging Patient Heterogeneity in Economic Evaluation." *PharmacoEconomics* 31 (2): 111–23.
Gupta, Ranjan, J. P. Kim, Jack Spiegel, and Steven M. Ferguson. 2004. "Developing Products for Personalized Medicine: NIH Research Tools Policy Applications." *Personalized Medicine* 1 (1): 115–24.
Halpern, Rachel, Sonalee Agarwal, Carole Dembek, Leigh Borton, and Maria Lopex-Bresnahan. 2011. "Comparison of Adherence and Persistence among Multiple Sclerosis Patients Treated with Disease-Modifying Therapies: A Retrospective Administrative Claims Analysis." *Patient Preference and Adherence* 5:73–84.
Hartung, Daniel M., Dennis N. Bourdette, Sharia M. Ahmed, and Ruth H. Whitham. 2015. "The Cost of Multiple Sclerosis Drugs in the US and the Pharmaceutical Industry: Too Big to Fail?" *Neurology* 84 (21): 2185–92.
Holland, Kimberly, and Rachel Nall. 2016. "The JC Virus and Risks for MS Patients." Accessed Sept. 3, 2016. http://www.healthline.com/health/multiple-sclerosis/jc-virus-risks-for-ms-patients.
Hult, Kristopher. 2014. "Incremental Innovation and Pharmaceutical Productivity." Working paper.
Hult, Kristopher J., Sonia Jaffe, and Tomas J. Philipson. 2016. "How Does Technological Change Affect Quality-Adjusted Prices in Health Care? Systematic Evidence from Thousands of Innovations." NBER Working Paper no. 22986, Cambridge, MA.
Hutchinson, Ryan A., Richard A. Adams, Darragh G. McArt, Manuel Salto-Tellez, Bharat Jasani, and Peter W. Hamilton. 2015. "Epidermal Growth Factor Receptor Immunohistochemistry: New Opportunities in Metastatic Colorectal Cancer." *Journal of Translational Medicine* 13 (1): 1.
Hyun, Jae-Won, Su-Hyun Kim, In Hye Jeong, Suk-Won Ahn, So-Young Huh, Min Su Park, Young In Eom, et al. 2015. "Utility of the Rio Score and Modified Rio Score in Korean Patients with Multiple Sclerosis." *PLoS ONE* 10 (5): e0129243. https://www.doi.org/10.1371/journal.pone.0129243.
IMS. 2016. "Medicines Use and Spending Shifts." Accessed Sept. 3, 2016. https://www.imshealth.com/ files/web/IMSH%20Institute/Reports/Medicines_Use_and_Spending_Shifts/Medicine-Spending-and-Growth_1995–2014.pdf.
Kravitz, Richard L., Naihua Duan, and Joel Braslow. 2004. "Evidence-Based Medicine, Heterogeneity of Treatment Effects, and the Trouble with Averages." *Milbank Quarterly* 82 (4): 661–87.
Lucchinetti, Claudia, Wolfgang Bruck, Joseph Parisi, Bernd Scheithauer, Moses Rodriguez, and Hans Lassman. 2000. "Heterogeneity of Multiple Sclerosis Lesions: Implications for the Pathogenesis of Demyelination." *Annals of Neurology* 47 (6): 707–17.
Malucchi, S., A. Sala, F. Gilli, R. Bottero, A. Di Sapio, M. Capobianco, and A. Bertolotto. 2004. "Neutralizing Antibodies Reduce the Efficacy of βIfn during Treatment of Multiple Sclerosis." *Neurology* 62 (11): 2031–37.
McKinsey and Company. 2013. "Personalized Medicine: The Path Forward. 2013." http://www.mckinsey.com//media/McKinsey/dotcom/client_service/Pharma%20and%20Medical%20Products/PMP%20NEW/PDFs/McKinsey%20on%20Personalized%20Medicine%20March%202013.ashx.
MS in America. 2014. https://multiplesclerosis.net/ms-in-america-2014/.
Nadarajah, Saralees, and Samuel Kotz. 2008. "Exact Distribution of the Max/Min

of Two Gaussian Random Variables." *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 16 (2): 210–12.
Nelson, Phillip. 1970. "Information and Consumer Behavior." *Journal of Political Economy* 78 (2): 311–29.
———. 1974. "Advertising as Information." *Journal of Political Economy* 82 (4): 729–54.
Pavlovic, Dejan, Andriani C. Patera, Fredrik Nyberg, Marianne Gerber, Maggie Liu, and Progressive Multifocal Leukeoncephalopathy Consortium. 2015. "Progressive Multifocal Leukoencephalopathy: Current Treatment Options and Future Perspectives." *Therapeutic Advances in Neurological Disorders* 8 (6): 255–73.
Pietrangelo, Ann, and Valencia Higuera. 2016. "Multiple Sclerosis by the Numbers: Facts, Statistics, and You." Accessed Sept. 3, 2016. http://www.healthline.com/health/multiple-sclerosis/facts-statistics-infographic.
Pistoresi, Ryan Hay. 2015. "Cost-Utility Analysis of Medications for Relapsing-Remitting Multiple Sclerosis: A United States Third-Party Payer Perspective." PhD diss., Washington State University.
Prosser, Lisa A., Karen M. Kuntz, Amit Bar-Or, and Milton C. Weinstein. 2004. "Cost-Effectiveness of Interferon Beta-1a, Interferon Beta-1b, and Glatiramer Acetate in Newly Diagnosed Non-Primary Progressive Multiple Sclerosis." *Value in Health* 7 (5): 554–68.
Reynolds, Matthew W., Reejis Stephen, Chris Seaman, and Kitty Rajagopalan. 2010. "Persistence and Adherence to Disease Modifying Drugs among Patients with Multiple Sclerosis." *Current Medical Research and Opinion* 26 (3): 663–74.
Río, J., M. Tintoré, J. Sastre-Garriga, C. Nos, J. Castilló, C. Tur, M. Comabella, and X. Montalban. 2012. "Change in the Clinical Activity of Multiple Sclerosis after Treatment Switch for Suboptimal Response." *European Journal of Neurology* 19 (6): 899–904.
Rush, Carolina A., Heather J. MacLean, and Mark S. Freedman. 2015. "Aggressive Multiple Sclerosis: Proposed Definition and Treatment Algorithm." *Nature Reviews Neurology* 11 (7): 379–89.
Scottish Medicines Consortium. 2014. Database. https://www.scottishmedicines.org.uk/files/advice/alemtuzumab_Lemtrada_FINAL_April_2014_amended_24.04.14_02.07.14_for_website.pdf.
Segal, Jodi B., C. Weiss, and Ravi Varadhan. 2012. "Understanding Heterogeneity of Treatment Effects in Pragmatic Trials with an Example of a Large, Simple Trial of a Drug Treatment for Osteoporosis (White Paper)." Center for Medical Technology and Policy. http://www.cmtpnet.org/docs/resources/Segal-Heterogeneity-in-Pragmatic-Trials.pdf.
Sørensen, Per Soelberg, Antonio Bertolotto, Gilles Edan, Gavin Giovannoni, Ralf Gold, Eva Havrdova, Ludwig Kappos, Bernd C. Kieseier, Xavier Montalban, and Tomas Olsson. 2012. "Risk Stratification for Progressive Multifocal Leukoencephalopathy in Patients Treated with Natalizumab." *Multiple Sclerosis Journal* 18 (2): 143–52.
Sylvia Lawry Centre for Multiple Sclerosis Research. Sylvia Lawry Centre MS Patient Database. Accessed Apr. 1, 2018. http://www.slcmsr.net/en/about/start.html.
Trapp, Bruce D., and Klaus-Armin Nave. 2008. "Multiple Sclerosis: An Immune or Neurodegenerative Disorder?" *Annual Review Neuroscience* 31:247–69.
Treadaway, Katherine, Gary Cutter, Amber Salter, Sharon Lynch, James Simsarian, John Corboy, Douglas Jeffery, et al. 2009. "Factors That Influence Adherence with Disease-Modifying Therapy in MS." *Journal of Neurology* 256(4): 568–76.
UNAIDS. 2015. "New Formulation of HIV Treatment to Save More Children's

Lives—UNICEF and UNAIDS." Press Statement. http://www.unaids.org/en/resources/presscentre/pressreleaseandstatementarchive/2015/june/20150605_PS_children.

Van Schependom, Jeroen, Jeroen Gielen, Jorne Laton, and Guy Nagels. 2016. "Assessing PML Risk under Immunotherapy: If All You Have Is a Hammer, Everything Looks Like a Nail." *Multiple Sclerosis Journal* 22 (3): 389–92.

Wall Street Journal. 2005. "Best 10-Year Performer: Biogen Idec." Feb. 28. http://www.wsj.com/articles/SB110929411399963912.

Zaheer, Fariha, and Joseph R. Berger. 2012. "Treatment-Related Progressive Multifocal Leukoencephalopathy: Current Understanding and Future Steps." *Therapeutic Advances in Drug Safety* 3:227–39.