Decisional Dilemmas in Discontinuing Prolonged Disease-Modifying Treatment for Multiple Sclerosis

Executive Summary

Background

Multiple sclerosis (MS) is a variably debilitating disease characterized by demyelination (deterioration of the protective myelin sheaths covering nerve cell processes in the brain and spinal cord) and axon loss within the central nervous system. The lesions created by the myelin destruction and resulting scar tissue interfere with normal transmission along nerve fibers within the brain and to and from the brain. This results in classic symptoms associated with MS. The condition affects 2.5 million individuals worldwide and approximately 400,000 in the United States. About 40 percent of people with MS receive some form of disability income. Twice as many women as men are affected, and diagnosis usually occurs between the ages of 20 and 50.1

Symptoms and disease course are highly individual, depending on where the lesions occur within the central nervous system and the type of MS. Clinically definite MS types include the following:

- Relapsing-remitting MS (RRMS) is the most common form, affecting approximately 85 percent of patients. Patients typically are diagnosed in their 20s or 30s. Neurologic symptoms of a relapse typically develop over a course of days, stabilize, and spontaneously improve. However, over time permanent disability often accrues, with further relapses. Many patients with RRMS eventually transition to secondary progressive MS (below). Estimates of
the median time from RRMS onset to this transition range from 15 to 29 years.\(^3\)

- Secondary progressive MS (SPMS) is characterized by worsening disability with or without relapses. Patients may have exacerbations, but the trend over time is a relatively steady progression of disease and disability.\(^1\)

- Primary progressive MS (PPMS) represents about 15 percent of patients and affects women and men about equally. This form has the worst prognosis and is characterized by gradual and progressive worsening of function without distinct relapses.\(^1\)

- Progressive relapsing MS (PRMS) affects about 5 percent of patients. This form is usually diagnosed first as PPMS due to a steady worsening of functioning and changed to PRMS when the patient experiences a relapse. Recently, a recommendation has been made to eliminate PRMS as a type, classifying these patients as having PPMS.\(^4\)

People with clinically isolated syndrome (CIS), a first neurologic episode consistent with an MS relapse, may or may not go on to develop MS. CIS involves neurologic symptoms such as vision loss, numbness, or weakness that last at least 24 hours and are caused by inflammation or demyelination in one (monofocal) or more (multifocal) sites in the central nervous system. In a cohort of 107 CIS patients followed for 20 years, 60 patients with three or more lesions (seen via magnetic resonance imaging [MRI]) converted to definite MS, while only 7 with normal baseline MRI converted.\(^5\)

MS cannot be cured with current therapies. Disease-modifying treatments (DMTs) comprise immunomodulating and immunosuppressant medications aimed at slowing the progression of MS and improving quality of life. The working hypothesis is that reducing or preventing new lesions and their sequelae slows the worsening of the disease. DMTs currently approved by the Food and Drug Administration (FDA) for RRMS include interferon (IFN) beta-1a and -1b (some formulations also approved for CIS), glatiramer acetate, mitoxantrone (also approved for SPMS and PRMS), natalizumab, fingolimod, and dimethyl fumarate.

A 2013 Cochrane overview review and network analysis of 44 2- to 3-year trials of DMTs for MS found moderate-to high-quality evidence that DMTs are effective against recurrence of relapses in RRMS during the first 24 months of treatment compared with placebo.\(^6\) The network analysis ranked natalizumab as the most effective drug, followed in order by IFNbeta-1a (Rebif\(^8\)), mitoxantrone, glatiramer acetate (Copaxone\(^8\)), and IFNbeta-1b (Betaseron\(^8\)). Confidence in the evidence dropped to moderate for direct comparisons of mitoxantrone or IFNbeta-1b versus placebo and very low for glatiramer acetate versus placebo. Further, natalizumab and IFNbeta-1b were more effective than IFNbeta-1a in reducing the number of RRMS participants with disease progression, as measured with surrogate markers. In patients with progressive MS, both pairwise and network analysis found that no DMT analyzed prevented disability progression over 2 or 3 years. The overview and network analysis were too recent to include the newest approved drugs, such as fingolimod or dimethyl fumarate.

Unfortunately, the efficacy of MS treatments appears to correlate with the frequency and severity of side effects.\(^7\) The injectable treatments, the IFN drugs and glatiramer acetate, were modestly efficacious and side effects were tolerable by many patients.\(^6\) Mitoxantrone, an escalation medication, has a lifetime maximum dosage due to cardiotoxicity and risks of leukemia.\(^7\) Natalizumab, the first monoclonal antibody approved for treating MS, can induce the potentially fatal brain infection progressive multifocal leukoencephalopathy (PML). Risk for PML increases with natalizumab use longer than 2 years, anti-JC virus antibody status, and prior use of immuosuppressive agents.\(^8\) People taking natalizumab may take a drug holiday or discontinue use completely if their risk for PML increases, assessed by a positive test for the anti-JC virus antibody status.

Women considering pregnancy face special considerations for drug holidays. There are no class A drugs (drugs safe for use during pregnancy according to the FDA) for MS. Women and their physicians must weigh the possible risks of DMT exposure to the unborn fetus against the maternal risk of disease progression if they discontinue DMT.

The optimal duration of DMT use remains an open and controversial question. Many patients do not use these medications throughout their entire life after diagnosis. However, with few exceptions (such as natalizumab use or intended pregnancy), patients who opt for DMT for MS may end up using it for several years to decades, as long as they tolerate the treatment and the DMT seems effective. Patients may switch between DMTs in order to find one that is more effective or more tolerable, and studies have found high rates of switching between drugs.\(^9\) Some patients cannot tolerate any of the DMTs, but if a tolerable drug regime is determined, treatment generally continues until the individual reaches a disease stage where DMTs are no longer considered to be helping.
Such a point may be reached when it is determined that a person is nonresponsive to the medication due to disease progression. Determining when DMT is no longer helpful is challenging. Thus, major questions of interest are whether or not DMTs for MS alter the natural history of the disease in the long run and when (if ever) to discontinue DMT. The related question addresses the influence of patient values, beliefs, and preferences regarding discontinuing DMTs. Such information should support clinicians, patients, consumer advocates, and other decisionmakers on decisions to discontinue treatment.

**Scope and Key Questions**

This review examines the long-term (more than 3 years) consequences of continuing or discontinuing DMT. We looked for evidence that directly assessed discontinuing versus continuing DMT and also evidence for long-term (more than 3 years) benefits and harms for either continuing or discontinuing, since that decision can be informed by the benefits or harms directly linked to either course of action. This information would extend understanding beyond the short-term trials examined in the 2013 Cochrane review. We were also interested in the reasons for discontinuing treatment reported in the long-term studies.

We concentrated on outcomes that patients notice or factor directly into their decisionmaking, such as relapse rates and changes in disability level, rather than intermediate outcomes such as lab tests for neutralizing antibodies. MRI to identify MS-related lesions has been shown to correlate with short-term relapse rates (6 months to 2 years). However, long-term MRI followup results as a surrogate marker for relapse rates or, more importantly, disease progression, currently lack evidence. Further, short-term MRI followup as a predictor of long-term disability progression answers a different research question—i.e., does short-term treatment affect long-term outcomes—than the research questions asked for this review. Thus, we did not use MRI results as a long-term outcome in this review. However, we included MRI results as a short-term outcome in the subset of patients discontinuing natalizumab due to risk of PML.

People with MS commonly switch between the available DMTs, depending on tolerance, presence of adverse effects, and perceived helpfulness of the treatment. The pertinent clinical question for switching medications is how to define the threshold of disease activity for changing medications. This important question is qualitatively different from that of when, if ever, to stop DMT completely. To adequately address the question of when to switch medications will likely require a review of both short- and long-term research. Therefore, questions related to switching between DMTs are outside the scope of this review.

We synthesized the evidence in the published literature to address the following two Key Questions (KQs):

**KQ1: What are the consequences of discontinuing disease-modifying treatments in adult patients?**

a. What is the evidence for benefits for continuing versus discontinuing treatment?

b. What is the evidence for long-term harms?

c. What reasons for discontinuation of disease-modifying treatments have been reported in long-term observational cohort studies?

**KQ2: What are individual values, beliefs, and preferences regarding discontinuing disease-modifying treatments?**

a. What are patient and provider preferences for discontinuation of disease-modifying treatments?

b. What are patient and provider preferences for participation in shared decisionmaking to discontinue disease-modifying treatments?

Figure A provides a conceptual framework that links the KQs. At the top it depicts the logic path both physicians and patients must travel when considering DMT:

- Does it work?
- What drug should I start with?
- When should I switch to a new drug and what should that drug be?
- When should I discontinue DMT?

This logic path describes the context within which patients and clinicians consider clinical factors—tolerability of the medication, disease characteristics at the time of discontinuation (relapses, progression, MRI activity), risk of ongoing disease treatment, other impediments to continued medication use (difficulty in obtaining, injecting, or ingesting, cost, etc.)—and make decisions about DMT or, in the case of this review, discontinuation (KQ1). The lower part of the figure, the conceptual basis for KQ2, depicts the progression from an individual’s internal decision context and process (such as preferences, values, knowledge, beliefs, and cognitive behaviors and habits) to an interpersonal decision context and processes
between the physician and patient. The overlapping ovals representing the clinician and the patient indicate information shared between the two parties versus information and other cognitive processes specific to one individual. Any overlap depends in part on the level of sophistication a patient brings to the decisionmaking process and in part on how well a physician understands a patient’s beliefs, values, goals, and preferences. For example, a patient newly diagnosed with MS in the novice phase of learning about MS would likely have a smaller overlap. The interaction between the physician and patient results in decisions that can vary in their level of concordance.

Figure B provides an analytic framework describing the treatment path and long-term benefits and harms of continuing versus discontinuing DMT for KQ1.

**Figure A. Conceptual framework for Key Questions**

- **KQ1**
  - Does immunotherapy work?
  - What drug do I start with?
  - If/when should I switch to what drug?
  - When should I discontinue?

- **KQ2**
  - Physician Decision (Not Concordant)
  - Shared Decision (Concordant)
  - Patient Decision (Not Concordant)

- Preferences
- Values
- Knowledge
- Beliefs
- Cognitive Behaviors/Habits

**KQ = Key Question**
Figure B. Analytic framework for discontinuing disease-modifying treatments for MS

DMT = disease-modifying treatment; KQ = Key Question; MS = multiple sclerosis

Methods

The methods for this review follow the methods suggested in the Agency for Healthcare Research and Quality “Methods Guide for Effectiveness and Comparative Effectiveness Reviews” (available at www.effectivehealthcare.ahrq.gov). We used Technical Brief methods for KQ2. A complete description of the methods can be found in the full report. All methods and analyses were determined a priori.

Literature Search Strategy

We used bibliographic databases to identify publications on randomized controlled trials (RCTs), systematic reviews, and observational studies with control groups published from 1990 to August 2014 that enrolled adults with CIS or MS. Relevant bibliographic databases for this topic include MEDLINE®, the Cochrane Central Register of Controlled Trials (CENTRAL), PsycInfo®, and Scopus. We supplemented bibliographic database searches with backward citation searches of highly relevant systematic reviews.

Eligibility

For KQ1, we included studies from 1990 through August 2014 that examined patient-centered outcomes for patients with CIS or MS in outpatient settings using FDA-approved (through August 2014) DMT compared with placebo, other active DMT, or no DMT. We excluded studies of pediatric MS patients, studies of mitoxantrone (since it has a maximum lifetime dosage), and studies with 3-year or less followup. However timing was relaxed for women who were considering pregnancy or were pregnant, or patients discontinuing natalizumab due to changes in risk of PML.

For KQ2, we included studies of any design that examined individuals’ attitudes, values, preferences for discontinuing treatments and health states, perceptions of risk and seriousness of health states, or factors and processes patients with MS and clinicians use in shared decisionmaking.

Two independent investigators independently determined study eligibility and resolved disagreements through discussions (possibly with a third adjudicator) until consensus was achieved. Study selection involved an
extensive full-text review process to identify adult subgroups, since subgroup reporting was commonly not evident in titles and abstracts.

**Data Extraction**

We extracted data from included studies into standardized evidence tables. Extracted data included relevant population, intervention, baseline, and outcomes data on the adult subgroups of interest. Initial data abstraction was quality checked by a second investigator.

**Risk-of-Bias Assessment of Individual Studies**

The risk of bias of eligible studies was assessed by two independent investigators using instruments specific to each study design for KQ1. The two investigators consulted to reconcile any discrepancies in overall risk-of-bias assessments and, when needed, a third investigator was consulted to reconcile the summary judgment. For KQ1, we developed an instrument to assess risk of bias for observational studies using the RTI Observational Studies Risk of Bias and Precision Item Bank. We selected items most relevant in assessing risk of bias for this topic, including participant selection, ascertainment, attrition, performance, and appropriateness of analytic methods. Following Technical Brief methods, risk of bias was not assessed for KQ2.

**Data Synthesis**

For KQ1, we summarized the results into evidence tables and qualitatively synthesized evidence for comparisons for specific disease-modifying medications, unique populations, duration of DMT, length of study followup, and outcomes. We used the best of the evidence provided by the identified observational literature. So, while all identified articles underwent abstraction, only the best evidence, based on those studies closest to an “ideal” study design (those studies with the lowest risk of bias), are included in the evidence synthesis.

For KQ2, we summarized the results into evidence tables and conducted a qualitative synthesis. We grouped the literature by mapping the included studies to the conceptual framework (Figure A) and analyzed the study findings for emergent patterns in patient perspectives, clinician perspectives, and clinician/patient interpersonal interactions.

**Strength of the Body of Evidence**

The overall strength of evidence for selected outcomes for KQ1 (relapse rate, change in disability, progression of disease, time to sustained disease progression) within each comparison was evaluated based on four required domains: (1) study limitations (internal validity); (2) directness (single direct link between intervention and outcome); (3) consistency (similarity of effect direction and size); and (4) precision (degree of certainty around an estimate). A fifth domain, reporting bias, was assessed when strength of evidence based on the first four domains was moderate or high. Based on study design and conduct, risk of bias was rated as low, medium, or high. Consistency was rated as consistent, inconsistent, or unknown/not applicable (e.g., single study). Directness was rated as either direct or indirect. Precision was rated as precise or imprecise. Other factors that may be considered in assessing strength of evidence include dose-response relationship, the presence of confounders, and strength of association. Based on these factors, the overall evidence for each outcome was rated as follows:

- High—Very confident that estimate of effect lies close to true effect. Few or no deficiencies in body of evidence; findings believed to be stable.
- Moderate—Moderately confident that estimate of effect lies close to true effect. Some deficiencies in body of evidence; findings likely to be stable, but some doubt.
- Low—Limited confidence that estimate of effect lies close to true effect; major or numerous deficiencies in body of evidence. Additional evidence necessary before concluding that findings are stable or that estimate of effect is close to true effect.
- Insufficient—No evidence, unable to estimate an effect, or no confidence in estimate of effect. No evidence is available or the body of evidence precludes judgment.

Following Technical Brief methods, strength of evidence was not assessed for KQ2. This KQ was approached in a hypothesis-generating manner.

**Applicability**

Applicability of studies was determined according to the PICOTS (population, intervention, comparison, outcomes, timing, setting) framework. Study characteristics that may affect applicability include, but are not limited to, type of MS or CIS, unobserved differences in patient preferences, or country in which treatment is provided, given differences in international regulations and treatment preferences.
Results
We identified 4,359 unique citations in searching from 1990 to August 2014. After excluding articles at the title and abstract phase, full texts of 198 articles were reviewed to determine final inclusion. Seven articles were added through hand-search. Of the 61 articles retained for KQ1, 11 were specific to discontinuing natalizumab due to increased risk and 12 were specific to discontinuing due to pregnancy. Of the remaining 38 articles comprising 27 unique studies, only 16 studies contained complete information to allow for full analysis. All 38 articles were reviewed for information on reported reasons for discontinuation. For KQ2, 30 articles comprising 27 unique studies were included. Detailed tables and synthesis can be found in the full report.

KQ1a. Benefits of Continuing Versus Discontinuing DMT
The key points for KQ1a are as follows:
• No studies directly assessed the consequences of continuing versus discontinuing DMT in comparable populations.
• Low-strength evidence from one study with moderate risk of bias suggests that, for RRMS patients, long-term all-cause survival is higher for treatment-naïve patients who did not delay starting IFNbeta-1b by 2 years and used DMT for a longer duration than those who started later.
• Low-strength evidence from one study with moderate risk of bias suggests that IFN use did not change disability progression for RRMS patients.
• Insufficient evidence was available to assess long-term benefits of DMT for SPMS patients and most outcomes for RRMS patients. Except for those noted above, studies were high risk of bias and had small sample sizes, and reported effects were small in magnitude.

Results are summarized in Table A.

<table>
<thead>
<tr>
<th>DMT and Author</th>
<th>Type of MS at Baseline</th>
<th>Median or Mean Years to Final Assessment</th>
<th>All-Cause Mortality</th>
<th>Convert to SPMS</th>
<th>Strength of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>IFNbeta-1b: Goodin, 2012(^{18,19})</td>
<td>RRMS</td>
<td>21</td>
<td>All-cause mortality: 250 mg arm vs. placebo—HR, 0.532 (95% CI, 0.31 to 0.90) 50 mg arm vs. placebo—HR, 0.54 (95% CI, 0.32 to 0.92) Favors treatment</td>
<td>NR</td>
<td>Low</td>
</tr>
<tr>
<td>IFNbeta, mixed: Shirani, 2012(^{20})</td>
<td>RRMS</td>
<td>4.5–10.5</td>
<td>NR</td>
<td>Time to sustained EDSS 6: no difference from contemporary control (HR, 1.30; 95% CI, 0.92 to 1.83) or historical control (HR, 0.77; 95% CI, 0.58 to 1.02)</td>
<td>Low</td>
</tr>
</tbody>
</table>

CI = confidence interval; DMT = disease-modifying treatment; EDSS = Expanded Disability Status Scale; HR = hazard ratio; IFN = interferon; NR=not reported; MS = multiple sclerosis; RRMS = relapsing-remitting multiple sclerosis; SPMS = secondary progressive multiple sclerosis
No studies directly assessed continuing versus discontinuing DMT in comparable populations. We therefore turned to literature examining benefits for continuing DMT long term. Variation among the included studies on long-term benefits of DMT for patient populations, interventions, outcome measurements, and timeframes precluded meaningful pooling. Only two studies provide a low strength of evidence for two benefit outcomes measured long term for IFNs.

One study with moderate risk of bias examined all-cause mortality over a 21-year period for 366 patients who had enrolled in an RCT (98.4% of the original RCT participants) testing IFNbeta-1b for treatment-naïve RRMS patients in 11 clinics in North America.\(^\text{18}\) The study’s strength lies in the nearly complete followup of patients and the objective outcome measure. Participants in the two treatment arms (50 mg and 250 mg) showed lower all-cause mortality compared with the placebo arm. The survival rate for the placebo arm was consistent with survival rates reported in MS natural history studies. Median treatment duration for the three groups ranged from 7 years for the placebo group to 14 years for the 50 mg arm and 12 years for the 250 mg arm. Patients assigned to placebo had both later starts and shorter exposure to DMTs. Thus, the study cannot distinguish between the effects of early use and the effects of long-term use.

One study with moderate risk of bias examined the association between IFNbeta use and progression to a sustained Expanded Disability Status Scale (EDSS) score of 6 for 2,656 RRMS patients in Canada.\(^\text{20}\) Three arms were used: a treatment cohort followed for 5.1 years, a contemporary cohort followed 4 years, and a historical cohort (drawn from the pre-IFN period) followed 10.8 years. The strength of this study lies in the almost complete capture of MS patients, since patients were unable to obtain DMTs other than from the participating clinic, and the multiple statistical approaches used to test for association, including use of comorbidities (Charlson score) and socioeconomic status along with age, sex, disease duration, and EDSS. Propensity score adjustments did not substantially change the results. The study did not find statistically significant differences in hazard ratios for reaching a sustained EDSS score of 6 for either contemporary or historical cohort comparisons.

Insufficient evidence exists to address long-term benefits for glatiramer acetate, teriflunomide, and natalizumab for either RRMS or SPMS, as well as important MS outcomes for IFNbeta for RRMS beyond all-cause mortality or 5-year disability progression.

**KQ1b. Evidence for Harms**

The key points for KQ1b are as follows:

- Limited low-strength evidence suggests that harms for injectable DMTs do not differ between short term (2-3 years) and long term (up to 16 years for IFN, 22 years for glatiramer acetate, and 8.5 years for teriflunomide).
- The majority of discontinuation tends to occur in the short term (2-3 years from start).
- Broad variation in harms reporting precludes informative aggregation and summary.
- Because of high risk of bias and small sample sizes, evidence is insufficient for whether rebound after discontinuing natalizumab exists.
- Because of high risk of bias and small sample sizes, evidence is insufficient to address the risks of fetal exposure to DMT during pregnancy in women with MS or the risks to the mother from the drug holiday.

Results are summarized in Table B.
<table>
<thead>
<tr>
<th>DMT</th>
<th>Number of Studies; Total N; Followup</th>
<th>Any Adverse Event</th>
<th>At Least 1 Serious Adverse Event</th>
<th>Treatment Discontinuation for Adverse Event</th>
<th>Comparator Groups</th>
<th>Reported Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>IFNbeta-1a(^{21})</td>
<td>1</td>
<td>Most common AEs: injection site reactions, headache, flulike symptoms</td>
<td>NR</td>
<td>NR</td>
<td>No</td>
<td>Long-term events do not differ from short-term events.</td>
</tr>
<tr>
<td>IFNbeta-1b(^{19,22})</td>
<td>2</td>
<td>Most common AEs: injection site reactions, depression, flulike symptoms, headache</td>
<td>21% to 24%</td>
<td>Discontinuation rates “high” but numbers not reported</td>
<td>No</td>
<td>Long-term events do not differ from short-term events. Frequency declined over 16 years in continuers.</td>
</tr>
<tr>
<td>IFNbeta, mixed(^{9,23,24})</td>
<td>3</td>
<td>Most common AEs: injection site reactions, depression, flulike symptoms, headache</td>
<td>NR</td>
<td>3% during long-term followup; discontinuation for serious AE more likely to happen early in treatment course (1 year)</td>
<td>No</td>
<td>Headache more likely for IFNbeta-1a; injection site reactions more likely for IFNbeta-1b. No other differences between type of IFNbeta. Majority of discontinuations occur early/short term.</td>
</tr>
<tr>
<td>IFNbeta, mixed(^{25}) SPMS</td>
<td>1</td>
<td>Only 1 reported overall rate: 87.3%</td>
<td>NR</td>
<td>3.4%, although timing is not clear</td>
<td>No</td>
<td>Majority of discontinuations occur early/short term.</td>
</tr>
<tr>
<td>Glatiramer acetate(^{26-28})</td>
<td>3</td>
<td>Most common AE: injection site reactions</td>
<td>NR</td>
<td>Only 1 reported overall rate: 4.9% in long-term extension</td>
<td>No</td>
<td>Majority of discontinuations occur early/short term.</td>
</tr>
<tr>
<td>Teriflunomide(^{29})</td>
<td>1</td>
<td>98% of 7 mg dose and 100% of 14 mg dose experienced treatment-emergent AE</td>
<td>36% of 7 mg dose and 29% of 14 mg dose</td>
<td>13.6% of 7 mg dose and 13.6% of 14 mg dose</td>
<td>1 comparison to general population rates for cancer</td>
<td>Long-term events do not differ from short-term events.</td>
</tr>
</tbody>
</table>

AE = adverse event; IFN = interferon; NR = not reported; SPMS = secondary progressive multiple sclerosis
Eleven of the 16 unique studies reported harms in enough detail for abstraction. Only one of the studies was moderate risk of bias; all others were rated as high risk of bias.

The included studies used a wide range of reporting methods and adverse event categories that precluded simple aggregation over the studies. The most commonly reported adverse events were injection site reactions, flulike symptoms, depression, and headache. Serious adverse events were generally not reported, although two studies gave rates of about 25 percent to 30 percent of participants. Discontinuations tended to occur during the first or second year of the study. When reported, discontinuation rates during long-term followup were low, about 3 to 4 percent, but rates due to adverse events were not separate from total discontinuation rates, which would also include perceived lack of efficacy and other reasons not necessarily related to adverse events or side effects. Further, all studies lost participants to attrition. Dropouts from observational studies are more likely to bias reporting toward lack of adverse events. Patients on long-term treatment are self-selected for positive outcomes, even though this may be due to any combination of treatment effects and benign disease course. The studies also did not use large enough patient populations to adequately detect rare events.

Eight observational studies (all high risk of bias) addressed the risks of rebound disease activity with natalizumab treatment interruption. Determining whether rebound exists requires comparing disease activity prior to receiving natalizumab and disease activity after interrupting treatment. None of the studies used appropriate comparison groups. All but one study used a subjective definition for rebound.

Twelve observational studies (all high risk of bias) addressed the benefits and risks to mothers and fetuses of IFN, glatiramer acetate, or natalizumab treatment discontinuation due to pregnancy or intended pregnancy. Women who discontinue DMT with the intention of becoming pregnant risk increased relapses between discontinuation and pregnancy, as well as postpartum. Given that the studied populations are those who became pregnant, none of the studies capture what happens to women who discontinue DMT but do not become pregnant. Therefore, no research has observed whether such women are at increased risk of relapse.

KQ1c. Reasons for Discontinuing DMT

The key points for KQ1c are as follows:

- The broad variation in discontinuation reporting prevented useful aggregation of studies.
- All studies reported one or more adverse events and inefficacy or progression of disability as reasons to discontinue.
- Patient reasons for discontinuing DMT were not explored.

Results are summarized in Table C.

<table>
<thead>
<tr>
<th>DMT</th>
<th>Total Number of Studies</th>
<th>Adverse Event</th>
<th>Inefficacy or Progression of Disability</th>
<th>Intended Pregnancy</th>
<th>Long-Term Stable MS</th>
<th>Death</th>
<th>Protocol Violation</th>
<th>Patient Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glatiramer acetate</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Teriflunomide</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Interferon beta-1a</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Interferon beta-1b</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Interferon beta, mixed</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>DMT, mixed</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

DMT = disease-modifying treatment; MS = multiple sclerosis

*Category includes counts of discontinuation based both on clinician evaluation of disease progression and patient evaluation of lack of efficacy
Twenty articles of the full reporting set reported reasons for discontinuing treatment. The wide range of reporting methods and discontinuation categories prohibited detailed quantitative aggregation over the studies. Most articles reported numerous reasons for discontinuations. Unfortunately, the reason for the patient’s decision to discontinue remained largely unexplored. Minimal text in this category generally used phrases such as “by own will,”25,37 “withdrew consent,”22,28 or “voluntary withdrawal.”34,38,39

KQ2. Preferences for Discontinuing DMT

The 25 included unique studies (28 total articles) represented a wide range of study aims. Designs ranged from factor analysis of questionnaires to experimental psychology lab tests to trials of shared decisionmaking interventions. Study locations were international, including the United States,40-47 the Netherlands,48-50 Germany,51-61 Norway,62 a consortium of European countries,63 Canada,64,65 Italy,66 and Ireland.67

Given the complexity of understanding preferences and behaviors, and the wide range of study designs used over a small literature set, all KQ2 key points should be viewed as preliminary.

KQ2a. Intrapersonal Aspects

The key points for intrapersonal aspects of KQ2 are as follows:

- Patients overestimated intermediate-term risk of wheelchair use but underestimated the lifetime risk. This underestimation may indicate the uncertainty MS patients felt when contemplating their personal trajectories rather than lack of knowledge (2 studies).
- Patients are likely to use heuristics in risk assessments (1 study).
- With training, patients can improve risk understanding and sense of informed choice (1 study).
- Quantified preference studies suggest that patients are willing to make risk tradeoffs for benefits only to the point where the discomfort from side effects and treatment are equal to or worse than the disease symptoms (2 studies).
- Increasing out-of-pocket cost reduces DMT purchases (2 studies).
- Common reasons for discontinuing include side effects, uncertainty about or perceived lack of efficacy against disease progression, administration method and frequency, and cost (5 studies).
- MS patients tended to take responsibility for the decision to discontinue (3 studies), while viewing their neurologist as the driver for decisions regarding choice of DMT (1 study).
- Psychological models of behavior support the presence of rational processes contributing to patient decisionmaking (2 studies).

We found 14 studies in the literature on values and preferences to populate the intrapersonal portion of Figure A addressing intrapersonal factors. The literature tended to examine attitudes and cognition rather than patient knowledge and how that knowledge affected decisions. Studies examined risk expectation, preferences for DMT and treatment tradeoffs, knowledge of cost factors, reasons for using or discontinuing DMT, and theoretical approaches to understanding decisionmaking and behavior processes.

KQ2b. Interpersonal Aspects

The key points for interpersonal aspects of KQ2 are as follows:

- MS patients and their physicians can differ significantly in their perceptions of the relative importance of health states and risks (2 studies).
- Physicians and patients must communicate in order to clarify differences in perceptions and preferences (1 study).

Much less literature populates the interpersonal than intrapersonal portion of Figure A. Three studies examined interpersonal concerns, including the knowledge, values, beliefs, and preferences that both the patient and physician bring to a decisionmaking encounter, and also the extent to which this information is shared between the two. Communication issues also are important at the interpersonal level.

KQ2b. Patient and Provider Preferences for Participation in Shared Decisionmaking

The key points for KQ2b are as follows:

- Different MS patients may bring different information-seeking orientations to shared decisionmaking (1 study).
- Mildly cognitively impaired MS patients show a significantly reduced capacity to understand treatment disclosures, but understanding may be brought back to the level of healthy controls through repetition and recognition cuing (1 study).
• The large majority of people with MS prefer a collaborative or active role in treatment decisions (3 non-U.S. studies).

• Physicians cannot reliably predict patient preferences for an active participation role and may inadvertently pull patients away from their preferred treatment (2 studies).

• Both patient and third-party observers rated physicians as showing limited skill at involving patients in shared decisionmaking (1 study).

• Providing balanced evidence-based information alone is not sufficient to alter decisionmaking processes to help patients achieve their preferred participation role (1 study).

Literature for this KQ subquestion relates to shared decisionmaking for patients and providers. All but one of the identified studies populated the center box in the shared decisionmaking portion of Figure A. Five studies addressed shared decisionmaking from the patient side, four addressed the physician side, and one tested a decision aid to improve shared decisionmaking.

**Discussion**

Effective health care relies on the three legs of physicians’ clinical experience, patients’ knowledge of their specific health situations and preferences, and an evidence base. Together, these three components provide the input for medical decisions. In the absence of a clear unambiguous path to follow, patients are best served by shared decisionmaking, which requires clinicians to provide the best available information against which patients can weigh their preferences and risk tolerance.

The decisions around discontinuation of DMT are extremely personal and individual. It is hard to envision ever having enough information to cover all contingencies. Providers and MS patients who have followed a prolonged DMT treatment plan have little information to guide decisions regarding discontinuing DMT. Thus, personal preferences about risks take on more weight.

No literature directly compared continuing versus discontinuing DMT in comparable populations. Only sparse information was available to address one part of the decisionmaking picture faced by providers and patients: long-term benefits and harms. As summarized in Table D, low-strength evidence showed increased all-cause mortality for patients who started IFNbeta-1b 2 years earlier than the comparators, but no differences between treated and comparator groups in time to progression to SPMS (as measured by a sustained EDSS score of 6). Similarly, overall long-term harms were found to be no different from short-term harms. Low-strength evidence implies low confidence in the findings and the expectation that future research could change the findings. Evidence is insufficient to assess long-term benefits and harms for any other patient population, type of DMT, or outcome.
Table D. Summary of KQ1 findings with sufficient evidence

<table>
<thead>
<tr>
<th>DMTs Used in Long-Term Studies Assessing Discontinuing or Continuing DMTs</th>
<th>Number of Studies; Number of Participants</th>
<th>Findings</th>
<th>Strength of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cause survival: interferon beta-1b</td>
<td>1 study;\textsuperscript{18} N = 366 RRMS</td>
<td>All-cause mortality: 250 mg arm vs. placebo—HR, 0.532 (98% CI, 0.31 to 0.90) 50 mg arm vs. placebo—HR, 0.54 (95% CI, 0.32 to 0.92) Favors treatment</td>
<td>Low (moderate risk of bias, unknown consistency)</td>
</tr>
<tr>
<td>Time to progression to SPMS: interferon, mixed</td>
<td>1 study;\textsuperscript{20} N = 2,656 RRMS</td>
<td>No difference from contemporary or historical control</td>
<td>Low (moderate risk of bias, unknown consistency)</td>
</tr>
<tr>
<td>Overall harms: interferon, glatiramer acetate, teriflunomide</td>
<td>3 studies;\textsuperscript{18,28,29} N = 746 RRMS, interferon beta-1b, 16 years; N=46 RRMS, glatiramer acetate, 22 years; N=131 RRMS, 16 SPMS, teriflunomide, 8.5 years</td>
<td>Long-term harms not different from short-term harms (qualitative finding)</td>
<td>Low (high risk of bias, consistent, indeterminate precision)</td>
</tr>
</tbody>
</table>

CI = confidence interval; DMT = disease-modifying treatment; HR = hazard ratio; RRMS = relapsing-remitting multiple sclerosis; SPMS = secondary progressive multiple sclerosis

The current literature did not examine whether long-term benefits for DMTs remain after a patient converts to SPMS. For the special cases of natalizumab and discontinuation for planned pregnancy, evidence was insufficient to answer whether discontinuation is problem free.

In the absence of evidence, providers and patients are left with little to inform their preferences and guide their decisions regarding when to discontinue treatments. The majority of included studies reported reasons for patients discontinuing treatments, but the information provided was without detail. Adverse events and inefficacy or progression of disability were two expected categories. Other possible reasons for discontinuation, such as a patient’s desire to try alternative medicine approaches, perceived risk of long-term use, or financial concerns such as out-of-pocket costs or loss of insurance, are not noted. The “patient decision” category for discontinuing was consistently unexplored.

Harms from long-term DMT use, as is true for many treatments and medical conditions, is poorly reported in the literature. The low-strength evidence showing long-term harms to be generally similar to short-term harms may very well be upset by improved tracking and reporting. For example, a November 2014 FDA Drug Safety Communication reported the first confirmed fatal case of PML for an MS patient using dimethyl fumarate (Tecfidera\textsuperscript{®}).\textsuperscript{68} The patient had used dimethyl fumarate for 4 years.

KQ2 aimed to delve into what is known about patient and provider preferences. While the literature was sparse, with only 28 studies available to populate the conceptual map provided in Figure A, each of the three major conceptual areas was at least partially represented. No study directly asked why people are reluctant to discontinue when treatment seems no longer effective, but taken as a whole, the literature set provides some insight.

Overall, one can weave together the general themes found in the KQ2 literature. Admittedly, physicians cannot reliably predict patient preferences for shared decisionmaking, and often physicians and patients perceive the relative importance of health status or acceptable risks differently. However, when it comes to the decision to discontinue DMT, the patient drives the decision, and this preference and role are generally unchallenged by the physician. In some DMT discontinuations, the balance of
shared decisionmaking may shift to discordance between the physician and patient, with the physician deferring to patient preferences for continuing treatment or not. The quantified preferences work by Prosser and colleagues illustrates a paradox—patients are less likely to prefer DMT during the early course of the disease, when disease symptoms are lower than the side effects of the DMT, and more likely to use it at later stages of the disease, when the side effects are less than disease symptoms. This behavior is counter to the hypothesis under which DMTs are assumed to work, which is by using DMTs to reduce relapses early in the disease course to prevent or delay disease progression. Without more solid evidence for the long-term net benefits or the thresholds at which treatment is no longer effective in preventing disease progression, the decision to discontinue treatment remains preference sensitive.

The preferences literature underscores the complexity of the topic and the processes underlying decisionmaking. Both rational and nonrational (such as heuristic) processes came into play, and neither had primacy over the other. Cost was a factor in both self-report and through observation of purchasing behavior. Cognitive deficits impairing decisional capacity may be overcome with adequate cuing. Information is a necessary component of decisionmaking, yet nonrational factors can influence what information is sought at what time.

Preferences, values, and beliefs are highly variable, may change over time, and are linked to the nature of the patients’ relationships with their doctors. There may well be differences based on age, sex, race, class, and other factors. A patient’s preference position between “treat my MS at any cost/comfort from knowledge of receiving treatment” and “need strong evidence that the medication will help and be worth the cost/side effects” may change over time and as the disease changes.

Changing perceptions regarding health states were common across different parts of the intrapersonal literature. Risk perceptions and quantified preferences (which are risk based as well) both suggested that people with longer MS experience assigned higher values to disabled states or viewed them as less serious than people with shorter MS experience did. This is a finding consistent with other research into how people value different health states. Many people overestimate their aversion to hypothetical states of disability and hence eliminate treatment options that might lead to such disability, especially if it could be long term. The hypothetical disutilities for these states are consistently higher than the actual disutilities for those experiencing the state.

**Issues**

Several challenges impede the gathering of evidence to inform decisions to discontinue DMT. First, the potential differential effectiveness of DMTs for different patient subpopulations is unclear because of the lack of studies examining the questions as well as the use of unsatisfactory study designs. Whether DMTs for CIS patients are effective remains an open question. DMTs may offer little benefit in exchange for side effects and potential harms for patients with a benign MS course. Conversely, it is not known which patients are at risk of worsened disease activity (such as a rebound effect) when DMTs are discontinued, possibly prematurely. We cannot currently predict early or benign disease courses.

Second, the transition from RRMS to SPMS is difficult to ascertain and therefore poses challenges in the decision to discontinue treatment. Clear biomarkers do not exist, and neither do distinct boundaries for the transition. Currently, clinical judgment and EDSS changes or an EDSS score of 6 or 7 are generally used. Furthermore, some patients with RRMS never transition to a clear secondary progressive phase. Since relapses tend to decrease in frequency with advancing age (being rare after the sixth decade of life and very rare after the seventh decade), the problem arises of determining whether a patient’s lack of relapses is due to ongoing DMT or to the natural history of the disease. For example, consider a 75-year-old patient who developed RRMS at age 30, has been taking DMT since 1994, has had no relapses or new MRI lesions since 1996, and has shown no evidence of secondary progression (stable EDSS). Is the lack of relapses due to ongoing DMT use, or has this patient’s MS reached the stage where the risk of relapse is passed and there is no ongoing neurologic deterioration beyond what would be expected in normal aging (sometimes referred to as “burned out” MS)? Is it safe to discontinue DMT in such patients? Adequate data to answer this question are not yet available.

This observation leads to the third major challenge: measuring disability. The EDSS is the most commonly used scale in research, in part because it is the longest standing. Because the EDSS is largely driven by mobility assessment, available research is generally silent on potential benefits of DMT other than ambulation, such as upper limb function and cognitive impairments. Other validated measures of health status in MS that incorporate
more function domains include the MS Quality of Life-54 (MSQOL-54), the Functional Assessment of Multiple Sclerosis (FAMS), and the Multiple Sclerosis Quality-of-Life Inventory (MSQLI).\textsuperscript{73-75} As seen in KQ2, given that people with MS can value health domains differently than physicians (or perhaps researchers),\textsuperscript{64} the broader range of disability assessment should be pursued regardless of any potential limitations in comparing results with studies that used the EDSS exclusively.

Without adequate measures of quality of life, balancing the benefits of treatments against harms becomes challenging, especially across different drug regimens. DMTs are not benign with regard to side effects and risk profiles. The degree to which quality-of-life benefits of treatment are offset by quality-of-life decreases due to side effects and risk profiles is important.

Much remains to be done to understand patient preferences. Emerging but useful information was available to explore KQ2, but no study directly asked the question about preference for discontinuing treatment or explored why patients may be unwilling to discontinue even when treatment no longer appears effective. Lonergan and colleagues approached the question tangentially, asking physicians about how they counsel patients when considering discontinuation.\textsuperscript{69} Providers who are involved with such counseling sessions would also benefit from research that separates understanding of preferences, which may be clear to the patient, and the mixed feelings such preferences may generate, ranging from fear or grief related to “giving up” on the disease to relief at no longer carrying the burden of DMTs.

Newly approved drugs, such as fingolimod, and drugs in the development pipeline are emphasizing oral administration to improve medication uptake and adherence to treatment programs. Self-injection can be a deterrent to patients with MS starting injectable DMTs, and “shot-fatigue” is a significant factor for adherence. Oral medications will certainly have implications for preferences for continuing and discontinuing DMTs.

**Future Research**

Since only three areas of evidence for KQ1 were sufficient to provide answers with only low strength of evidence, essentially all questions related to KQ1 would benefit from further study. The utility of studies for estimating long-term treatment effectiveness in MS can be improved by using prospective population-based designs with appropriate comparators and standardized data collection methods. Study cohorts must be better characterized with respect to demographic and clinical characteristics, as well as other factors that may influence outcome, such as socioeconomic status, access to care, health behaviors, and comorbidities. Near-complete patient retention with regularly scheduled patient visits is also necessary. The ability to account for treatment effects would improve with better models to predict disability outcomes in MS, including disentangling the young versus old from the new versus long-term disease presence, since the two overlap. Techniques to adjust for selection bias, such as regression analysis or propensity scores, are more easily accomplished with rich datasets. Since the pharmaceutical industry would not benefit from strong comparator studies focusing on treatment discontinuation, other funding sources will need to be identified. With regard to the question of discontinuing for pregnancy, appropriate comparison groups need to include women who discontinued DMT to attempt pregnancy but did not conceive.

Some efforts to improve longer term research are underway—for example, the prospective 5-year OPT-Up study.\textsuperscript{70} While the study is more geared toward initial treatment and switching choices, understanding discontinuation within that context is one of its goals. A prospective 10-year observational study based on the United Kingdom’s MS risk-sharing scheme is evaluating the effectiveness of the first DMTs, IFN and glatiramer acetate. After the National Institute for Health and Care Excellence (NICE) recommended against DMTs in 2002,\textsuperscript{77} a pricing scheme was negotiated with participating pharmaceutical companies whereby the drug prices would be reduced if patient outcomes were lower than expected;\textsuperscript{78} thus, the United Kingdom National Health Service and the pharmaceutical companies shared the financial risk for cost-effective treatment. The initial 2-year results, published in 2009, found that patient outcomes were worse than predicted.\textsuperscript{79} However, results were controversial; an independent review of the data identified intrinsic flaws in both the control dataset and analysis model selected when setting up the risk-sharing scheme.\textsuperscript{80} Four-, 6-, and 8-year data have been collected and are being analyzed using an updated modeling methodology. This research initiative should help inform the long-term benefits of these injectable treatments and may suggest improvements to current MS registries or methods, making analysis of such registries more fruitful.

KQ2 covered a broad array of relevant topics, and investigator-driven research remains a likely source for innovative and interesting approaches to continued exploration. The AutoMS project, an international
A consortium of six European locations and Australia, was formed in 2010 to explore MS patient preferences for shared decisionmaking. Confirming the generalizability of their findings to the United States would be beneficial. Also useful would be well-designed qualitative and survey research, perhaps as a mixed-methods study, exploring why and under what circumstances a patient might seek to terminate treatment, and why people are reluctant to discontinue when treatment appears no longer effective.

Attention to areas such as implementation science and quality improvement, which combined are often referred to as health care improvement science, should be included in the future of MS research and the improvement of MS treatment decisions and outcomes. Such areas may contribute systems-level factors to DMT selection and adherence, and to the successful implementation of shared decisionmaking.

Limitations

Literature on preferences is not indexed to permit easy identification of relevant articles. Search strategies to capture the diffuse literature used natural language as keywords. While we tested multiple terms before settling on the final algorithm, relevant articles were likely missed, and thus the included literature set must be viewed as comprehensive but not exhaustive. Likewise, setting the review scope to exclude adherence literature, as adherence by definition connotes a decision to continue DMT use, may have precluded some relevant literature examining lack of adherence as a de facto decision to discontinue use.

References


Full Report