Have we lost our way? The need for dynamic formulations of smoking relapse proneness

Thomas M. Piasecki, Michael C. Fiore, Danielle E. McCarthy & Timothy B. Baker

ABSTRACT

Current smoking cessation treatments seem to differ from one another in the proportion of ex-smokers who survive the first few days of the quit attempt. After this initial effect, parallel relapse processes appear to unfold in all treatment groups; no available treatments seem to alter the nature of this late relapse process. True relapse prevention will require that we obtain a better understanding of the forces contributing to relapse across the span of the cessation attempt. A working model of dynamic relapse processes may be necessary before treatments can be improved. In this paper, we suggest that the conceptual model of relapse proneness (RP) described originally by the National Working Conference on Smoking Relapse can serve as an ecumenical organizational framework that may be used to integrate and conceptualize relapse data in ways that could generate new strategies for relapse research and inform treatment design. As an illustration, we sketch a preliminary model of RP which postulates that physical withdrawal, stressors/temptations, and cessation fatigue each make independent, time-shifted contributions to relapse risk. A new round of descriptive research focused on relapse proneness processes may be a prerequisite for improving existing smoking cessation interventions.

KEYWORDS Relapse, relapse prevention, smoking cessation.

INTRODUCTION

Relapse to smoking claims the vast majority of quitters within a few weeks of the quit attempt. Although researchers have devised dozens of intensive behavioral treatments and developed an extensive pharmacopoeia to promote smoking cessation (Schwartz 1987; Fiore et al. 2000; USDHHS 2000), no currently available therapy consistently produces long-term abstinence rates greater than 50% (Shiffman 1993; Fiore et al. 2000). Despite rigorous efforts on the part of investigators and clinicians, we have not been able to understand or solve the problem of smoking relapse, which begs the question ‘Have we lost our way?’

We do not wish to dismiss the many strides that have been made in our understanding and treatment of tobacco addiction. Available cessation treatments have been found to double abstinence rates relative to no treatment conditions, and there is ample evidence that individuals can, and often do, quit smoking and remain abstinent indefinitely (e.g. Bjornson et al. 1995; Carlson et al. 2000). Additionally, many types of effective medical treatments produce only small but significant impacts on other chronic diseases (Hughes 1999). Nonetheless, the fact that, using current treatment approaches, post-treatment relapse is always more common than abstinence is a bedrock reality that must be acknowledged if further progress is to be made. We believe that our poor treatment success rates reflect a lack of understanding of addiction processes and will best be alleviated by a re-dedication to basic research efforts.

Some time ago, we accepted a request from the editor of this journal to provide a brief review of contemporary relapse prevention treatments for smoking cessation. Upon reflection, however, we found ourselves unable to identify a single pharmacological or behavioral intervention capable of truly preventing smoking relapse. Despite acceleration in research into smoking cessation, a pro-
found gap exists between what we can currently achieve clinically and what we want to achieve. How shall we close the gap? What is needed, we believe, is an organizational framework that can integrate diverse lines of relapse data and generate novel treatment hypotheses. In this paper, we suggest that the concept of relapse proneness (RP) curves (Shiftman et al. 1986) can provide such a framework. We sketch the outlines of an heuristic relapse model that postulates three RP forces (two of which are familiar variables), and try to show how construing commonly assessed variables in a joint dynamic formulation raises crucial questions for future basic and clinical research.

**RELAPSE PREVENTION AS FAVORABLY PATTERNED SURVIVAL**

Relapse to smoking is a dynamic process that may unfold idiosyncratically, and no single metric can perfectly summarize the relapse process (e.g. Shiffman, Hickcox et al. 1996). Nonetheless, smoking cessation researchers have developed and promulgated a number of useful data-analytic conventions for describing smoking behavior and cessation success (e.g. Ossip-Klein et al. 1986; Ockene et al. 2000). Although different conventions (i.e. point-prevalence versus continuous abstinence measures) may yield slightly different glimpses of the relapse process, they are largely fungible. Different measures tend to produce differing event rates, but all seem to reflect a similar pattern of behaviors in treated smokers—a reliable erosion of treatment gains over time. For the purpose of exposition, we will assume that curves derived from survival analysis yield reasonable portraits of the relapse process (see Curry et al. 1988; Willett & Singer 1993 for detailed reviews). In typical survival analyses, groups are compared across time on the proportion of ex-smokers who ‘survive’, where survival is defined as abstinence from smoking. When an individual reports smoking at a rate that exceeds some criterion value, that event represents a ‘death’, and the proportion of survivors is reduced accordingly. Each subject contributes data to a survival analysis until s/he smokes, and then does not factor in subsequent estimates of groupwise survival. Thus, survival analyses do not reflect renewed attempts at abstinence; the proportion of survivors may only decrease or remain stable. [This property of survival analysis, like the name itself, reflects its biostatistical origins as a measure of group differences in total mortality—a scenario in which ‘re-entry’ into the survivor class is nonsensical. At least one recent variant of survival analysis permits reversion to survivor status (e.g. Willet & Singer 1995), but we are not aware of any examples of its use in the smoking cessation field.]

While survival curves may be useful assays or markers of relapse, it is also clear that their form depends to some extent on investigators’ decisions regarding the types of data to be submitted to analysis. For instance, relapse rates in the same sample may look quite different if an investigator: (a) constrains the analyzed sample to individuals who achieved some sustained period of abstinence, or (b) includes all subjects who agreed to quit on a given date, even though many did not end up doing so. This is because in much of the ‘action’ in smoking treatment failure occurs at the outset of the quit period (Kenford et al. 1994; Stapleton et al. 1995).

Smoking that occurs very early in a quit attempt is both exceedingly common as well as conceptually troublesome. Early treatment failures can be construed either as failures of cessation or relapse (Ossip-Klein et al. 1986; Ockene et al. 2000). Decisions about the proper label for early smoking or delineation of its temporal parameters must be somewhat arbitrary. However, we hypothesize that it is meaningful to make process distinctions between early smoking events on the one hand and relapses occurring after some significant period of abstinence on the other. Individuals who smoke early in treatment may differ from later relapsers on personological variables and, in addition, the proximal instigators of smoking may also differ across time (i.e. as a function of abstinence duration). Therefore, in the discussion that follows, we will make distinctions between early and late treatment effects on survival. We refer to the first 7 days after the target quit date as the ‘cessation attempt’ period. We recognize that others could argue for an earlier cut-off (e.g. Ossip-Klein et al. 1986 defined 24–48 h as the period required for abstinence). We have chosen to consider the first week as part of the initial cessation attempt, however, to highlight the distinction between achieving early but substantial cessation and maintaining it over a longer period of time. The person who is abstinent over much or all of this 7-day cessation attempt period then enters a period of ‘relapse susceptibility’.

Armed with this heuristic distinction, we may ask what do survival curves tell us about relapse and relapse prevention? Smoking cessation clinical trial data tend to yield steep initial slopes in plotted survival curves (see for example Curry et al. 1988; Stevens & Hollis 1989; Transdermal Nicotine Study Group 1991; Fortmann & Killen 1995; Schneider et al. 1996; Zhu et al. 1996; Blondal et al. 1999; Jorenby et al. 1999). This pattern is relatively invariant across diverse smoking cessation interventions (although survival analyses tend to be reported most often in pharmacotherapy trials). Scrutiny of survival curves reveals, by and large, that treatment effects are attributable to differences in the cessation attempt period, not slope differences during the period of relapse susceptibility. In general, groups treated differ-
ently will tend to have similar slopes and rates of relapse deceleration, but may differ in survival rates very early in the quit attempt.

Figures 1 and 2 present survival curves from some recent clinical trials conducted (in whole or part) at the University of Wisconsin that illustrate this point. Figure 1 shows survival curves from two sister trials comparing the efficacy of a 22-mg nicotine patch with a placebo (Fiore et al. 1994). In one trial, the patch regimen was paired with intensive group counseling (60 minutes per week for 8 weeks). In the second trial, patch regimens were paired with brief individual counseling (10–20 minutes per week for 8 weeks). Both trials demonstrated a statistical advantage for the active patch regimen, but this advantage was due to early differences in cessation, and not due to late survival slopes. Figure 2 presents survival data from a clinical trial crossing nicotine patch dose (22 mg or 44 mg) with three levels of adjuvant counseling (Jorenby et al. 1995). Although six distinct combinations of treatment were implemented, relapse slopes were similar across groups. Of course, exceptions to the rule may sometimes be obtained. For instance, in this trial there was a hint of divergence in late relapse slopes (i.e. days 25–45) between the high-contact and low-contact groups in the 22-mg patch condition. The clinical importance and replicability of this finding is uncertain.

More potent treatments seem to shepherd a higher proportion of smokers through the first few days of a quit attempt without lapsing, and this small increment in the
The proportion of abstainers carries forward. In other words, when significant treatment effects are found, parallel relapse processes appear to unfold in groups with different base rates of survival after 7 days of the quit attempt. Treatments touted as effective do not appear to modify fundamentally the rate or nature of relapse after this initial effect. It is for this reason that we believe no currently available treatments can make a strong claim to preventing relapse.

A ‘true-blue’ relapse prevention treatment should yield a survival curve characterized by a distinctive slope and deceleration in the period of relapse susceptibility. A relapse prevention treatment might produce significant effects during the cessation attempt period, but its superiority to comparison treatments would not depend upon this difference. Even if the survival analysis were conducted on subjects with a week or more of abstinence, the relapse prevention group should still show evidence of superiority; i.e. it would alter the probability of late relapse events, and so new ‘deaths’ would occur much more slowly. One such hypothetical relapse prevention curve is schematized in Fig. 3. As the figure indicates, complete survival is not required to demonstrate relapse prevention—some smoking is likely to occur even in the most potent relapse prevention treatment. The major requirement is that the rate or probability of relapse over time is starkly modified. The hypothetical relapse prevention curve depicted in Fig. 3 may seem fanciful. Certainly,
Dynamic formulations of smoking relapse proneness

it implies a level of success that we cannot attain with current treatment methods. Of course, there exists a family of curves, many with lower absolute abstinence rates, that would be equally consistent with our requirement of nonequivalent late relapse slopes. We deliberately made our depiction dramatic to emphasize the improvement in treatment yields that might be possible if late relapse could be reliably prevented.

THE FUNDAMENTALLY ALTERED ORGANISM

Once we specify the kind of outcomes we expect from a true relapse prevention treatment, we can speculate a little more meaningfully about the kinds of patient change that would be required to produce such outcomes. One implication of the ‘favorably patterned survival’ conception of relapse prevention is that smokers receiving such a treatment must be fundamentally altered with respect to relapse proneness. In other words, a true relapse prevention treatment must somehow render the ex-smoker relatively impervious to the most typical instigators of smoking relapse. Relapse researchers have long understood this to be an important goal (e.g. Shiffman 1982; Marlatt & Gordon 1985), but it remains unattained.

The alterations produced by a real relapse prevention treatment would clearly need to be enduring if the are to affect patient survival across the cessation attempt. Enduring alterations of relapse risk seem most likely to arise from the more or less permanent alteration of an individual’s susceptibility to relapse instigators even after the offset of treatment delivery. Interventions that ‘prop up’ the ex-smoker temporarily (i.e. booster counseling sessions; Brandon et al. 1987) are certainly useful, but to the extent they require continued application of treatment to stave off relapse, they fall short of the desideratum.

The treatment effect produced by a true relapse prevention program should also be far-reaching or generalizable, such that it prepares the ex-smoker to anticipate and deflect the diverse challenges to abstinence s/he is likely to face. This requirement highlights the fundamental need for descriptive information about relapse risk processes. What are the most significant hurdles that ex-smokers face? How do they relate to one another or interact to prod a return to smoking? How do they unfold over time? Relapse prevention will remain a moving target until these questions are answered convincingly.

CLOSING THE GAP: AN ORGANIZATIONAL FRAMEWORK

To a large extent, relapse researchers have focused on developing descriptive taxonomies of relapse situations or testing bivariate relations between relapse and another interesting variable. We believe that such research has been fruitful, but an important next step for the field is to craft an organizational framework for compiling and synthesizing data pertinent to the entire smoking relapse process. It may be profitable here to draw a distinction between formal theories of relapse and an organizational framework for relapse research. Until now most accounts of smoking relapse have taken the form of theoretical models that posit specific mechanisms of drug motivation (e.g. Solomon & Corbit 1973; Prochaska & DiClemente...
RELAPSE PRONENESS AS ORGANIZATIONAL METRIC

In 1985, a task force of the National Working Conference on Smoking Relapse presented a family of dynamic relapse process models that had the potential to provide a skeleton for organizing theory and research on the determinants of relapse (Shiffman et al. 1986). Briefly, the task force introduced the concept of relapse proneness (RP), a continuously scaled dynamic latent variable indexing the probability of smoking recurrence in an abstinent individual over time. Smoking was hypothesized to occur if RP crossed critical thresholds. The task force used the RP concept as a thought-provoking device that could help to illustrate the kinds of dynamic forces that could give rise to relapse. They discussed three RP time signatures that seemed plausible in light of existing research and theory: a slow-building, tonic or cumulative process, a phasic or episodic process, and an interactive model in which episodic effects are superimposed on underlying cumulative risk processes (Shiffman et al. 1986).

The task force’s organizing framework highlighted two crucial problems in smoking relapse research: (1) the need to identify the manner in which particular relapse risks unfold over time, and (2) the need for greater understanding of how component relapse risks are combined to determine overall relapse susceptibility. These questions remain unresolved. While much contemporary relapse research is relevant to these questions, the field has not taken advantage of the task force’s framework to organize these findings in a way that implicates particular risk factors and that specifies their function forms.

In what follows, we offer a provisional heuristic model of relapse that specifies three relapse risks and their hypothesized RP function forms. The provisional model we present below does not exhaust the domain of relapse risks, and thus cannot be considered a comprehensive or final account of relapse proneness. We present it in the hope that our preliminary exposition will invite other investigators to elaborate on the skeletal structure we propose and to consider how their ongoing work fits within the Task Force’s RP framework.

THREE FORCES OF RELAPSE

We attempt to illustrate the value of the RP framework by providing a provisional heuristic model of relapse that emphasizes three RP forces. Two of these—physical withdrawal symptoms and stressor/temptation events—are cornerstones of relapse research and will undoubtedly require representation in any consensual framework of relapse forces. A third hypothesized force—cessation fatigue—is more speculative. Fatigue may or may not prove to be an important process in smoking relapse. We include it here, in part, because little is known about it; fatigue illustrates how an RP framework might suggest new putative relapse mechanisms and shape theorizing and research. The framework we are proposing, like that developed by the task force (Shiffman et al. 1986), begins with the assumption that an individual ex-smoker can be thought of as possessing a latent relapse proneness (RP)—a continuously scaled variable that describes the likelihood of smoking from moment to moment. RP is assumed to fluctuate over time, and smoking behavior is assumed to occur when the ex-smoker’s RP crosses some threshold value. [Shiffman et al. (1986) made a distinction between a low ‘slip threshold’ above which isolated smoking events occur that do not doom the cessation attempt, and a higher ‘relapse threshold’ above which an enduring resumption of regular smoking occurs. This distinction is an important one. For simplicity of exposition and for the sake of parsimony, however, we will depict only a single relapse threshold in our model; the reader can easily generalize the model to a two-threshold case.] We are proposing three RP forces, each of which is hypothesized to follow a unique time-course, and different forces are assumed to predominate at particular postquit latencies. We postulate that, at any given point,
an individual’s total RP is determined by a composite of his/her standing on each of the component processes. The composite RP status may represent a simple summation of RP scores, or may reflect some interactive (i.e. multiplicative) effects. [Regrettably, too little is known at present to determine how forces combine. Even though two of the forces in our heuristic model (physical withdrawal and stressors/temptations) are well-studied, they have typically not been studied in a way that would address whether they are summative or interactive dynamic risks. The problem is made trickier by the fact that common self-report measures of withdrawal symptoms probably reflect the impact of stressors/temptations and other affective events. A virtue of the RP framework approach is that it may goad us to develop cleaner measures of particular forces and to study them jointly over time.] In either case, the RP curve is hypothesized to assume the shape of a complex waveform. We assume that each ex-smoker encounters these three forces, but also recognize that the manifestation of each may be moderated by individual differences. Even this small family of postulated forces and basic assumptions generates a working model with the potential for tremendous complexity and a wide variety of individual RP curves.

Withdrawal

The withdrawal symptom trajectories of individual ex-smokers are idiosyncratic and volatile. However, when symptom profiles are averaged across many smokers who quit in unison, much of this idiosyncrasy ‘washes out’ to reveal a reliable groupwise withdrawal pattern characterized by a peak in symptoms within 1 week of the quit attempt, and a monotonic decrease in symptoms across the next 3–4 weeks (e.g. Piasecki et al. 1998).

One interpretation of both the heterogeneity in withdrawal symptoms and the signal among the noise is that withdrawal reports reflect both a physiological/pharmacological withdrawal process common to all smokers, and associative and affective events unique to each smoker’s environment and life circumstances (Piasecki et al. 1998). Averaging across smokers amplifies the common variance and reveals the signature of simple pharmacological withdrawal symptomatology (schematized in Fig. 4a). [It is somewhat risky to try to label some aspects of smoking withdrawal reports as ‘pharmacological’ or ‘physiological’ and others as ‘associative’ or ‘affective’. Of course, associative and affective processes are rooted in physiology, and their linkage to drug motivation requires experience with pharamacological processes. Our semantic distinction is convenient but, at some level, misleading. Nonetheless, similarly coarse distinctions are proving generative in the animal literature (Epping-Jordan et al. 1998; Watkins et al. 2000).] This early and reliable manifestation of smoking withdrawal probably reflects a physiological counter-reaction that is triggered by a physical event (i.e. declining drug levels in the body). It is temporally tied to the cessation event, is dampened by nicotine administration (e.g. Hughes et al. 1991) and its transient temporal pattern is similar to that seen in animal models (e.g. Isola et al. 1999).

We conceptualize this early, common withdrawal pattern as an automatic symptom burst that will reliably attend cessation of smoking. This symptom burst puts smokers at high risk of relapse very early in the quit attempt, although evidence on this score is somewhat controversial (cf. Hughes et al. 1990; Patten & Martin 1996). Statistical prediction of relapse from early withdrawal may be ‘swamped out’ in many data analyses because smoking rates are very high early in the cessation attempt period. During this critical window, absolute levels of withdrawal severity may vary across smokers
(undercutting traditional prediction models), but the relative, within-subject level of symptomatology is probably at its peak for most smokers. In addition, analytical strategies that jettison lapsing smokers will truncate the distribution of withdrawal scores, as lapsing smokers tend to have the highest levels of withdrawal (Piasecki et al. 2002).

This analysis of physical withdrawal suggests that it exerts its most powerful effects on RP very early in the quit attempt—during the ‘cessation attempt period’ where we believe current efficacious treatments exert their effects. The field has made impressive progress in devising interventions to combat this aspect of RP. One reason for this success is that our understanding of withdrawal (sudden de-occupation of drug receptor sites) led in a very direct way to the development of agonist replacement therapies that are efficacious in ameliorating early withdrawal symptoms. Another reason is that the process is common to nearly all smokers and occurs very early in treatment—it represents a clinically obvious target for intervention and may be the most probable candidate of any relapse instigator for a ‘one size fits all’ treatment maneuver.

One reason for the steady erosion of abstinence rates over time may be that treatments designed to combat physical withdrawal become less relevant as the quit attempt progresses; the presumed temporal signature of physical withdrawal suggests that its contributions to RP should diminish very early in treatment. The fact that new relapses occur even after physical withdrawal has been forded suggests the need to postulate the existence of additional forces that modify RP (assuming that associatively elicited withdrawal is incapable of accounting for these late relapses). Additionally, it seems reasonable to infer that, whatever these additional forces are, they are not responsive (or are minimally responsive) to the treatments currently in use to mitigate physical withdrawal; current treatments are often applied (e.g. continual dispensing of nicotine patches) long after the cessation attempt period has ended, and yet the toll of relapse continues to mount.

**Stressors/temptations**

As noted earlier, serial withdrawal reports collected from individuals do not always follow the clean, reliable pattern of steadily waning symptoms that we have characterized as ‘physical withdrawal’. For instance, in one sample, 68% showed a pattern of increasing or unremitting symptoms over 8 weeks of the cessation attempt (Piasecki et al. 1998). Slow oscillations and sudden spikes in symptomatology are both fairly common features of individual symptom profiles (Piasecki et al. 2002). If we assume that self-reports of withdrawal symptomatology index proximal motivation to smoke (and, thus, reflect RP), we may infer that RP is influenced by one or more processes distinct from those that produce the prototypical physical withdrawal pattern.

Just as it seems reasonable to look to pharmacological and physiological variables to explain commonalities in withdrawal across samples, it may be useful to turn to environmental and person variables to explain why symptom patterns diverge over time (e.g. Gilbert 1997; Piasecki et al. 1998). We hypothesize that the idiosyncratic features of withdrawal profiles may reflect the impact of stressors and temptations (e.g. Shiffman et al. 1996a, 1996b). Such provocative events may activate a latent vulnerability to RP, causing aperiodic surges in motivation to smoke. Because smokers’ environments will differ in the density and provocativeness/severity of these instigators, individuals will tend to show different patterns of RP over time. Moreover, because the form of stressors and temptations may change over time and across individuals, stressors and temptations can produce a multitude of RP waveforms. One possible pattern of stressor/temptation deflections of RP is schematized in Fig. 4b. This feature of our model resembles the ‘episodic model’ of RP described by Shiffman et al. (1986). Stressors and temptations have been the subject of much research in recent years, and many findings implicate such events as important instigators of relapse (e.g. O’Connell & Martin 1987; Shiffman et al. 1996b; Ockene et al. 2000). However, there is still more to learn about the determinants, temporal dynamics, variability and motivational impact of stressors/temptations.

Successful quitters often report that maintaining abstinence becomes easier over time. On the other hand, long-term quitters often report occasional strong cravings months or even years after quitting (e.g. Daughton et al. 1999). This leads us to hypothesize that, in general, the interval between significant stressor/temptation events may increase over time, and that the average exacerbation of RP provoked by these events may decrease over time. Although admittedly speculative, such a pattern would account for both the observation that some people eventually attain stable abstinence and the fact that relapses do occur weeks or months after cessation. A pattern of less-frequent, less-intense perturbations of RP would also be consistent with a number of plausible behavioral mechanisms. For instance, classically conditioned responses to omnipresent cues would eventually extinguish, but rare cues might occasionally evoke a relatively strong motivational response. Common stressful situations might result in extensive coping response practice that leads to routinization of coping after a period of abstinence, but unusual situations might still require effortful processing and press for less-developed responses.
Fatigue

We propose an additional RP-related process that reflects the cumulative toll of quitting and follows a distinctive course over time. Although we label this factor ‘fatigue’, we recognize that this is simply one possible late-emerging influence on relapse proneness that could be incorporated into the proposed framework. Multiple factors probably contribute to late relapses (e.g. overconfidence that fosters experimentation with smoking or recrudescence withdrawal episodes). We propose the fatigue factor as one candidate process to explain the occurrence of lapses or relapses, particularly those occurring weeks or months after a quit attempt. Late relapses have long puzzled addiction researchers and current models of addiction fail to account adequately for this phenomenon. We offer fatigue for its heuristic value, not because it is necessarily the most important determinant of late-occurring relapse.

Our clinical experience suggests that a smoker who attempts to quit experiences a cumulative toll or cost of quitting that saps their motivation to quit, coping skill utilization and ability to stay quit. Early in the quit attempt, quitting may be viewed as a ‘new adventure’ and each day without smoking as an exciting new accomplishment. Over time, however, the novelty of the cessation attempt is likely to dissipate, and dealing with stressors and temptations or simply maintaining vigilance against them is likely to be perceived as taxing. [One smoker going through our smoking cessation program alluded to this when he said ‘It was easy in the beginning...fighting a war is exciting and energizing...it’s standing “guard duty” the rest of your life that’s hard.’] We propose that the toll of continued efforts to resist temptations and stay quit could be captured in repeated measures of aspects of this latent process, such as decisional balance, motivation, self-efficacy, coping and self-control. We further postulate that changes in these measures would probably be predictive of increased risk of relapse.

Fatigue, as we define it, is a latent construct that could be operationalized by the following: loss of motivation, loss of hope in cessation success, a reduction in coping attempts, decreased self-efficacy and exhaustion of limited self-control resources. None of these indicators, by itself, is likely to adequately sample the fatigue construct (e.g. individuals may decrease their coping behavior because of a lack of provocations, or because of overconfidence rather than fatigue), but collectively they may serve as useful indicators. Adequate measurement of cessation fatigue therefore requires a synthesis of diverse, repeated assessments.

To date, very few researchers have used an approach that permits detection of dynamic patterns of motivational or fatigue constructs and how these relate to relapse proneness. Although much research suggests that precessation motivation to quit is predictive of readiness to change (e.g. McBride et al. 2001) or abstinence (e.g. Curry et al. 2001), motivation to quit has rarely been studied as a dynamic factor that can wax and wane over the postcessation period (see Hedeker & Mermelstein 1996 for a notable exception). Studies of motivation in other domains have demonstrated that motivation increases and decreases with certain situational factors (Brehm & Self 1989). We contend that smoking cessation research could benefit from the integration of these dynamic motivational approaches with traditional cessation research methodology. Certainly not all motivational influences will wax and wane over the postcessation period in a manner that influences relapse vulnerability (e.g. self-efficacy, Shiffman et al. 2000), but something must account for late-occurring relapse, and erosion of motivation constitutes a viable possibility. Future research needs to identify probable late-acting determinants of relapse vulnerability such as fatigue based upon relevant theory and data.

Theoretical justification for examining fatigue comes from basic and applied research on self-control that suggests that self-control resources may be limited and function in a way analogous to muscles (Heatherton & Baumeister 1996; Muraven & Baumeister 2000). Muraven & Baumeister (2000) explicated this muscle analogy, arguing that, like muscles, self-control resources are limited and subject to temporary depletion. They suggest that self-control resources can be exhausted by ongoing or rapidly successive demands, and may require a recovery period to return to full capacity. A good deal of evidence supports the contention that high self-control demand in one task is the specific factor that predicts subsequent loss of control on another related or unrelated task (see Muraven & Baumeister 2000 for a review).

Interestingly, Shiffman (1984) published data that seem to support the exhaustion of self-control resources during smoking cessation. In this study, quitters who experienced a strong temptation to smoke and resisted it were more likely to relapse subsequently than quitters who did not experience an earlier temptation. Shiffman (1984) contended that decreases in self-efficacy mediated the impact of the earlier lapse on subsequent relapse risk, but exhaustion of self-control resources might also play an important role in the causal chain leading up to relapse. Both changes in self-efficacy and availability of self-control resources might be manifestations of a latent factor, such as fatigue, that influences relapse proneness.

Additional research suggests that self-control performance deteriorates over time in the face of continuous demands. Research clearly shows that performance on tasks that require vigilance deteriorates over time (Davies
Quitting smoking often requires both relentless vigilance and the ability to cope with discrete stressors and temptation events that probably deplete self-control resources. Ironically, although current treatment approaches emphasizing self-control skills (vigilance, planning and coping) equip participants with tools to overcome common challenges in cessation, they may also increase a fatigue-prone individual’s long-term relapse risk by taxing limited self-control resources.

To our knowledge, no empirical studies have attempted to measure cessation fatigue directly. Future research should consider measuring fatigue; something as simple as a brief self-report assessment administered serially might prove informative. However, as is the case for stressor/temptation events, sensitive assessment of fatigue may require the assessment of both events (i.e. frequency/extensity of coping behavior) as well as the subjective toll exacted by such struggles.

While we know of no studies that have intentionally measured fatigue, it is probable that fatigue has been measured indirectly. As mentioned above, it seems probable that measures of motivation levels, decisional balance, confidence and affect/depression may all reflect fatigue in some way. Of course, it is possible that not only may fatigue ‘drive’ or influence these measures, but also that the constructs tapped by these measures, in turn, affect perceived fatigue. Also, examination of subjective and event measures may allow the investigator to identify unique patterns of vulnerability. High levels of perceived fatigue paired with low coping might mean something very different from fatigue in the presence of high coping rates. As with all complex psychological constructs, multiple convergent measures should yield the most faithful rendering of these relapse-relevant processes.

Although we may speculate generally that fatigue will increase over time, we do not know enough about this variable to specify the details of its time-course. For instance, the rise time of fatigue is difficult to estimate without empirical data. Moreover, we do not know whether fatigue reaches an asymptote and remains stable for long periods of time, or whether it fluctuates with life circumstances from day-to-day. In Fig. 4c, we have drawn the fatigue curve to show a rolling-off very late in cessation; it does not seem reasonable to expect that cessation fatigue is a permanent reaction to quitting. At some point, non-smoking must become habitual and relatively effortless. However, we do not know how long an individual needs to have remained abstinent before this decay of fatigue begins. Moreover, as with other cessation-instigated processes, fatigue may prove to be extremely heterogeneous.

In a recent cessation trial, we assessed repeatedly participants’ motivation to quit over a 7-week period following their quit date. A single item about motivation cannot sample the fatigue construct adequately; e.g. motivation might be affected by processes other than fatigue. Examination of these data suggests, however, that an erosion of motivation is occurring that is superficially consistent with the operation of a fatigue process. The data were drawn from a study of 55 smokers who completed multiple daily ecological momentary assessments in which they were prompted by palm-top computers at random times during the day to complete brief questionnaires. A daily score was calculated for each individual by averaging the motivation rating within each day. Then, the average level of motivation across people was calculated. Figure 5 shows the smoothed, average daily values for motivation to quit for those who relapsed (panel a) and those who maintained abstinence (panel b) during the assessment period. Although motivation remains high through the first month of the quit attempt for those who relapsed during the assessment period, there appears to be a sharp drop-off in motivation early in the second month.
Our nascent efforts to implicate fatigue in relapse are offered primarily for their heuristic value. Researchers have not intensively studied fatigue (or other probable factors influencing relapse proneness such as ‘overconfidence’) as a significant relapse determinant. We believe that this inattention reflects, in part, the failure of researchers to consider how different dynamic factors might unfold or behave over time, and the implications that this might have for relapse. For instance, could each successful effort to cope with smoking urges, or a lack of smoking urges, foster an inappropriate confidence in the ability to exercise self-control while smoking? While fatigue may or may not prove to be a potent determinant of relapse vulnerability, we believe that a consideration of dynamic processes, and their interactions with episodic events, will lead to new and more powerful models of relapse (e.g. Shiffman et al. 1986).

Relapse proneness as composite of forces

Figure 4d depicts the sum of the particular component curves shown in panels a–c of Fig. 4 and described above (it resembles the ‘interactive’ model presented by the task force (Shiffman et al. 1986). Although this composite curve and its components are highly schematic, they possess a number of notable features that accord with the clinical picture of smoking relapse. One clear implication of the composite curve is that RP is not expected to be constant or decreasing monotonically. New phasic challenges to abstinence may occur at any time after cessation, and may catch the ex-smoker off guard. Another feature of the composite curve is that the running average of total RP tends to be high compared to that of each component process (i.e. relative to the relapse threshold). This occurs because the three forces are summed, and because each is hypothesized to peak at different postcessation latencies. For instance, late in cessation, fatigue and the occasional stressor/temptation keep RP elevated relative to the level predicted by a simple model of relapse that concentrates only on physical withdrawal. It is possible to elaborate the composite curve further so as to incorporate interactive effects between RP components (see below); although not pictured here, scrutiny of force–force interactions would represent an important goal for descriptive research.

Interplay of relapse forces

We postulate that the three major RP forces may interact and influence one another to some extent, and that this interplay can increase the likelihood of relapse. For instance, severe physical withdrawal symptoms may create a motivational state that makes temptations more compelling—the withdrawal symptoms may serve as discriminative stimuli that signal the availability of potent negative reinforcement and thus amplify the salience of environmental smoking cues (e.g. Brandon et al. 1995). Severe physical withdrawal and frequent, strong stressors/temptations should increase the effort required to maintain abstinence, and so these features should be associated with heightened fatigue as cessation progresses.

More complex pathways are also possible if lapse events are considered in the model. For example, intermittent smoking triggered by stressors/temptations might retard the dissipation of physical withdrawal, allowing it to contribute to RP longer than would be expected under complete abstinence. This is consistent with recent data showing that low-level smoking during a cessation attempt is associated with increases in withdrawal symptoms (Piasecki et al. 2002). Negative or reciprocal relations could also obtain. For instance, initially low physical withdrawal might reduce or delay fatigue, but also might engender complacency that renders a smoker more vulnerable to temptation events. Such hypotheses require empirical testing and must await the development of measurement strategies for tracking RP forces. The main theoretical point is that one may need to explain the level and pattern of a given RP force with reference to the progress of the other forces.

QUESTIONS FOR FUTURE RESEARCH

Let us presume for a moment that our general framework holds, regardless of whether or not the specific relapse forces we have articulated are correct. If we assume that relapse proneness is driven by latent, dynamic forces, how can we study them? How will we measure these forces, test hypotheses about their nature, time-course and relative importance and, finally, build models to capture their dynamic influence on relapse? First, we need to learn how these forces influence relapse in the general
case, then we need to identify variables that moderate their influence and finally, we must modify smoking cessation treatments to reflect our new understanding of relapse processes. Below we outline some of the key questions to address in future research based on the framework proposed here and suggest some strategies that might be used to test these hypotheses.

The first challenge we face is to identify the specific relapse forces that are critical determinants of relapse proneness. We have proposed three specific forces here, but these and additional dynamic candidate factors must be examined in future research. One important objective of future research, therefore, should be to identify a taxonomy of relapse forces. In addition to identifying, defining and describing each force, we must also identify a chronology for each of these specific factors. Tracking determinants of relapse proneness over time is critical to our goal of understanding the process of relapse. Different factors will probably predominate at different points in time. Careful descriptive work that tracks changes in each factor will therefore be essential to our understanding of relapse. An additional goal would be to construct a hierarchy of RP forces, organized according to potency or ‘relapse riskiness’ (akin to a beta-weight) conferred by each factor.

Articulation of these goals raises some important methodological concerns. Do we currently have at our disposal the assessment and analytical tools necessary to achieve these goals? We must begin to track vulnerability to relapse throughout the postcessation period using techniques such as ecological momentary assessment (Stone & Shiffman 1994) rather than concentrating exclusively on predicting or describing discrete smoking events. To this end, we must develop repeatable measures of RP; instead of simply noting when patients smoke, we need to find out how close they are to smoking at many points along the way. Simple, face-valid items administered repeatedly may allow us to do this, either alone or in conjunction with repeated measures of commonly assessed constructs, such as withdrawal, urge or negative affect.

Once we have collected these repeated measures, how can we analyse these data to capture dynamic processes sensitively? Initially we could construct dynamic plots depicting the trajectories of putatively important variables over time to meet our goal of understanding the chronology of these relapse forces. Simple, descriptive analytical approaches such as this will not be sufficient, however. Examination of average values across time, or average trajectories will mask heterogeneity in responses to cessation along several dimensions. For this reason, we want to use an analytical approach that captures, rather than ignores, this information for each relapse force (Piasecki et al. 2002).

Recent developments in the analysis of data profiles over time afford the investigator powerful tools to explore RP forces in an appropriately sensitive manner. In particular, multi-level modeling (e.g. Bryk & Raudenbush 1992; Hedeker & Mermelstein 1996; Hedeker & Mermelstein 2000) and univariate variable growth models (e.g. Muthen & Curran 1997; Jackson et al. 2000; Jackson et al. 2001) and non-linear dynamical models (e.g. Vallacher & Nowak 1997) constitute an impressive armamentarium for characterizing the motivational influences and setting events for relapse. In addition, we may want to take advantage of advances in biostatistical methodology by using process modeling approaches, such as stochastic modeling, that have already been applied to other areas of clinical research (Carter et al. 1998). Researchers will need to select carefully the most appropriate and powerful way to model each RP force. For instance, in the current paper we have modeled stress/temptation events as a process—the manifestations of which are linked over time (see Fig. 4b). This is in keeping with the fact that consistencies in lifestyle and relatively stable person-factors create autoregressivity in such measures. However, it might be more appropriate to view such events as true episodic challenges that exert temporally discrete effects on relapse vulnerability. Thus, using the terminology of non-linear mathematical modeling, one would investigate whether the RP force, and its motivational impact, are captured adequately by its state value (episodic magnitude) or instead whether its phasic value is also critical (its dependency upon proximal values; Haynes et al. 1995; Heiby 1995).

Additional analytic complexity is added by the fact that every relapse alters the population of relapse-susceptible individuals. Thus, the potential relevance/potency of relapse component processes (e.g. fatigue) across a population of ex-smokers will be determined not only by dynamic processes affecting the individual smoker (e.g. the time-course of physical withdrawal) but also by the prior impacts of relapse processes. That is, at any point in time, the ‘distillate’ of ex-smokers reflects the prior influence of the diverse relapse components. For example, those individuals most susceptible to physical withdrawal will be eliminated from the pool of ex-smokers early in the quit attempt. We need to bear this in mind and develop research methods that will allow us to take this non-independence of relapse processes within individuals into account.

Once we have identified, characterized and modeled individual relapse forces, we will then need to consider how the factors are related to one another. Specifically, we will need to determine whether RP forces are fairly independent of one another or, alternatively, if they are correlated or causally related to one another, either unidirectionally or reciprocally. For instance, we might
want to determine whether withdrawal symptoms and fatigue were unrelated, whether withdrawal exacerbations were related to later increased fatigue, and whether increases in fatigue were similarly associated with subsequent worsening of withdrawal symptoms. When looking at the relations between factors, it will be important to remember that the nature of such relations might change over time. For instance, withdrawal severity might be primarily ‘driven’ by pharmacological factors early in the quitting process, but more reactive to stressors later in the postcessation period. A fine-grained analysis of the relationship between forces will be important in the construction of models of relapse proneness.

In addition to the straightforward additive and interactive models of relapse processes introduced above, investigators must also consider more complex models of the relationships among relapse factors. For example, interrelations might best be modeled using non-linear methods. Such models are appropriate where changes in causal factors produce disproportionate changes in target behaviors, where transitions in the target behaviors reflect the impact of historical values of both the target behavior and causal factors, and where transitions in the target behavior involve an ongoing positive feedback loop (e.g. Heiby 1995). In the case of relapse proneness, it is easy to imagine how the same stressor might cause a disproportionate impact on relapse proneness as a function of historical values (e.g. the level of other causal factors, prior exposure to stressors, existing levels of relapse proneness). Additionally, it is easy to see how reinforcement or priming mechanisms might produce positive feedback effects once a lapse has occurred.

Once the relationships between relapse forces have been explored and defined, then we must determine how these forces influence relapse proneness. Models depicting each of the relapse forces and the relationship between forces will ultimately need to be tied to cessation outcomes. The ability of models of relapse processes to predict relapse proneness and cessation outcome will need to be tested in prospective cessation trials.

An additional level of complexity will be added when the extensive research documenting individual differences in responses to cessation is applied to a relapse proneness framework. Individual ex-smokers will probably have unique areas of strength and weakness. Additionally, ex-smokers will differ from one another in terms of life circumstances (e.g. density of stressors). Both personological and situational variables may press for individual differences in the shape, level and persistence of RP component processes and total RP curves. Individual differences and their relations to smoking relapse have been the subjects of much theorizing and research (e.g. Gilbert 1995; Shiffman et al. 1996a; Pomerleau et al. 2000), and a thorough treatment of these influences is properly the subject of a different paper. The point to be made here is that the component RP curves are not likely to unfold in a lockstep fashion, but rather they should vary parametrically across smokers as an orderly function of some finite set of personological and situational variables. Models of relapse processes must reflect the influence of these variables on relapse forces and relapse proneness.

Ultimately, extensive research on relapse processes should lead to improved treatment of tobacco addiction by suggesting targets for intervention and ways to optimize the tailoring of intervention components and the timing of their delivery. If our supposition that relapse proneness is multi-faceted is correct, then treatment interventions should anticipate and address diverse relapse instigators. A potential benefit of dynamic models of relapse proneness is that they may suggest rational hypotheses for treatment matching. Moreover, they may specify the timing at which the treatment needs to be delivered for the matching effect to be obtained. Although we are not yet in a position to make specific recommendations for interventions, we anticipate that research addressing the questions outlined above will generate specific, testable hypotheses regarding treatment planning, staging and matching.

We have proposed that, in future, investigators take a more process-orientated approach to the study of relapse. We suggest further that treatment evaluation research methods should change in a parallel fashion to identify the mechanisms of treatment effects (Piasecki & Baker 2001). Such research should track not only conventional outcomes such as absolute abstinence, but also scrutinize the treatment effect on the RP force per se. Modification of the theoretically targeted force would be encouraging even if total abstinence were not dramatically increased.

A powerful demonstration of the efficacy of a relapse-targeted treatment would be the demonstration that treatment affects relapse within a particular postcessation epoch, that it affects relapse among individuals at elevated risk for a specific component force (e.g. fatigue), and that measures of the component force mediate changes in relapse rates.

CONCLUSION

Smoking cessation treatments may already be as effective as our current knowledge of relapse permits; our relative ignorance of the crucial factors in relapse represents a barrier that we cannot surmount without a vigorous rededication of effort to basic research and theory on relapse and relapse proneness. We have proposed both a framework and an heuristic model in the hopes that our speculation will spark debate and a return to a way of thinking well-captured by the 1985 task force but largely
abandoned since then. Even if the particular variables we have emphasized are in error, this basic outline of a multi-process model of RP illustrates how treatments can fail to prevent relapse—RP may remain elevated for many weeks or months, but the processes conspiring to raise RP may shift over time. Concomitant shifts in treatment strategy may be required to keep pace with the changing nature of relapse proneness. We need to understand what the relevant forces are, how they combine with one another and which forces predominate in the composite at any given time if we are to mine treatment lessons from relapse data.

ACKNOWLEDGEMENTS

Preparation of this manuscript was supported in part by a grant from the University of Missouri Research Board and by grant CA84724 from the National Cancer Institute. The authors wish to thank Stevens S. Smith PhD for his assistance with Fig. 1.

REFERENCES


Hughes, J. R. (1999) Four beliefs that may impede progress in the treatment of smoking. Tobacco Control, 8, 323–326.


Dynamic formulations of smoking relapse proneness 1107


effects and predictors of outcome in a randomized trial of transdermal nicotine patches in general practice. *Addiction*, 90, 31–42.


