Identifying and Screening for Psychological and Comorbid Medical and Psychological Disorders in Medical Settings

Rodger Kessler
University of Vermont College of Medicine

There is increased attention to the medical and economic consequences of psychological problems comorbid with medical issues. There is also a clear awareness that most psychological problems are assessed and responded to in nonpsychiatric medical settings. This has furthered interest and attention in implementing screening procedures to better identify psychological, behavioral, and substance abuse problems in medical settings. Such interest is taking the form of recommendations from federal government task forces, and the funding of large projects to include screening in medical settings. At the same time there has been further attention to brief, valid, and reliable measures with which to capture psychological comorbidities. However, there have been multiple concerns raised about a variety of issues concerning the utility and effectiveness of such screening procedures and the identification of multiple issues to be considered in screening design. The author outlines and reviews the rationale and concerns about screening, identifies the issues that need to be considered in screening program development, and describes the efforts to develop a screening capacity in a rural family practice. J Clin Psychol 65:253–267, 2009.

Keywords: screening; comorbid disorders; rural family practice; collaborative medical psychological care

Introduction

There is a clinical and empirical need to identify patients in primary care settings who have a psychological disorder independent of or as part of their medical

My appreciation to Dale Stafford, MD for his critical assistance in editing this manuscript.

Correspondence concerning this article should be addressed to: Rodger Kessler, Department of Family Medicine, University of Vermont College of Medicine, Berlin Family Health, Fletcher Allen Health Care, Montpelier, VT 05602; e-mail: Rodger.Kessler@uvm.edu

Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jclp.20546

presentation.

- At any given time 21–26% of patients in a primary care practice have a diagnosable mental disorder (Academy of Psychosomatic Medicine, 1996).
- Half of presentations to primary care for physical problems are idiopathic or psychiatric problems. Twenty-six percent of such patients reported no improvement at follow-up (Cowley et al., 2003).
- Most frequent presentations had somatic components including back pain, limb pain, headache, dyspnea, cough, abdominal problems, chest pain, dermatologic complaints, dizziness, sleep complaints, and fatigue (Cowley et al., 2003).
- Of the most frequently presented medical problems, the majority often has significant psychological components (Kroenke & Mangelsdorff, 1989).
- More people receive psychological services in the primary care system than the specialty mental health system (Kessler & Stafford, 2008; Regier et al., 1993).
- Fifty percent of people with a psychological disorder are treated exclusively by primary care (Narrow, Regier, Rae, Manderscheid, & Locke, 1993).
- Medical patients with psychological comorbidties have higher utilization of health care resources on the order of 50–100% higher non-mental health medical costs (Simon, Von Korff, & Barlow, 1995; Unutzer et al., 1997).
- Depression is the third most frequent reason for consulting a primary care physician (Mitchell & Coyne, 2007).

In addition, currently in medicine, there is a great deal of attention to optimal primary care treatment of patients with chronic medical disorders, such as diabetes. It is acknowledged, but often organizationally and clinically overlooked that among patients with chronic medical conditions, 25–38% have a comorbid psychological disorder (Academy of Psychosomatic Medicine, 1996). The prevalence of depression is roughly twice as high among diabetics as among the general medical population (Piette, Richardson, & Valenstein, 2004). Patients with diabetes who are depressed have poorer glycemic control and more severe diabetic symptoms and have higher health care costs relative to patients' with diabetes, but no depression (Lin et al., 2006). Among people with diabetes, depression was associated with increased disability (Von Korff et al., 2005). Treatment of their depression improves clinical outcomes and a health care cost reduction of $379–$952 per patient has been demonstrated over the course of 2 years (Simon et al., 2007).

Major depressive disorder is also a risk factor for the development of coronary events in healthy patients and for adverse cardiovascular outcomes in patients with established heart disease (Whooley, 2006). It is also associated with cardiac mortality (Lowe et al., 2002). When depression is effectively treated among medical patients there is a general decrease in use of medical care (Simon, Chisholm, Treglia, & Bushnell, 2002; Simon et al., 2000).

A major conundrum in medical practice is that although comorbidity rates are high and there are significant medical and cost consequences, rates of detection have been limited (Kaner, Lock, McAvoy, Heather, & Givarry, 1999; Lowe et al., 2002; Thompson, 2000). Typically, there is a failure to recognize 30–50% of depression (Coyne, Thompson, Palmer, Kagee, & Maunsell, 2000; Kaner, Lock, McAvoy, Heather, & Givarry, 2002). Less than 10% of panic disorder is generally detected in primary care (Lowe et al., 2002). Although substance use and abuse contributes to multiple medical problems and worsens health status, it is poorly detected (Brown, Leonard, Saunders, & Papasouliotis, 1997).

In summary, many patients attending primary care with a medical presentation either are masking a psychological disorder, or have a psychological disorder that is affecting their acute or chronic medical problem. The psychological dimensions are frequently undetected. If detected and appropriately treated, good psychological, medical, and cost outcomes are often generated.

Screening as a Method of Identification

The U.S. Preventive Services Task Force (USPSTF) recommends screening adults for depression in clinical practices that have systems in place to assure accurate diagnosis, effective treatment, and follow-up (U.S. Preventative Services Task Force, 2002). It has been suggested that screening may be a method to reduce health care cost by reducing need (Fries et al., 1993) if patients identified receive appropriate treatment.

In the report advocating routine screening for depression, The USPSTF reviewed 10 trials that measured the effect of screening and feedback given to physicians on outcomes or care. Five reported significant improvement and three others reported improvements that did not reach statistical significance (U.S. Preventative Services Task Force, 2002). Support for those findings suggest that if screening identifies patients with behavioral health problems and treatment is provided, such treatment works (Lowe et al., 2002) and that good screening information generates physician action (Kaner et al., 2002). It has been observed that screening and information improves detection in primary care at a low cost (Oslin et al., 2006).

Support for screening effectiveness has also been questioned. Valenstein, Vijan, Zeber, Boehm, and Butler (2001) have suggested that while most studies suggest if screening increases treatment provision, it does not effect outcomes. Coyne and Thompson (2003) have found a lack of evidence that under conditions of routine care detection actually improves outcome or improves care. There have also been findings that routine screening did not increase identification and that even if there was increased identification it did not increase intervention (Gilbody, House, & Sheldon, 2001), and may or may not effect treatment initiation (Pignone et al., 2002). Many authors have found that when patients assessed as needing services were referred for psychological treatment less than half of those make contact (Callahan et al., 2002; Williams, Palmes, Kurt, Pulley, & Meschan, 2005). Recently, a meta-analysis was conducted of the effectiveness of screening and case finding for depression in primary care that included the studies originally considered in the USPSTF analysis and more recent research. The conclusion was that if used alone, case finding or screening questionnaires for depression have little or no impact on either detection or management of depression because organizational resources to respond to screening findings are not provided (Gilbody, Sheldon, & House, 2008).

In support of those findings, screening has only been found effective if costs are low and effective treatment options are readily available (Valenstein et al., 2001). In reviewing postscreening actions, Pignone et al. (2002) have found that just providing feedback about results of screening has had mixed results; however, screening with prescribed intervention has had better results. Also, questionnaire data has been found to be more useful when a specific disorder is targeted and symptoms were operationalized (Hoyer, Becker, Neumer, Soeder, & Margrat, 2002). Unless screening questionnaires are routinely used and demonstrated to effect treatment they are unlikely to continue to be used (Gilbody et al., 2001).

In summary, the decision to screen and the issues to be considered in screening are complicated; there are a set of specific issues that should be considered that will probably determine the success or failure of screening projects.

Who Should Be Screened and What They Should Be Screened For

The first obvious question to be reviewed is who should be screened. The USPSTF has suggested routine screening of all adults for depression (U.S. Preventative Services Task Force, 2002) and has previously suggested screening for problem alcohol use (Brown et al., 1997). Although intuitively appealing there are consequences of such a strategy that should be considered. Screening costs real money. Costs of administering and communicating the results of screening have ranged from $5.00–$25.00 per screen (Oslin et al., 2006; Valenstein et al., 2001). The outcome of a universal screening will generate a high percentage of negative screens and over half of positive screens are subsequently found to be false-positives (Coyne et al., 2000; Gilbody et al., 2008; Palmer & Coyne, 2003).

However, there is a potential benefit of a universal strategy in that absent a definitive method of predicting which populations are at higher chance of screening positive, a universal screen will detect patients in populations not identified at risk. However, this strategy is particularly expensive because it will increase the amount of negative screens, while maintaining the base cost of the procedure.

The alternative strategy that has been suggested is to select populations to be screened that evidence suggests are at high risk—such as patients with chronic medical disorders, new significant medical diagnoses, or patients who recently have lost a long-term partner. Such a strategy requires a review of pertinent literature to select subpopulations associated with high frequencies of psychological comorbidities. Alternatively, one could sample by age, selecting populations that are associated with high frequencies of psychological disorders, such as adolescents or the elderly. A limitation of a selected screening model is the other side of universal screening. Patients from populations that are not selected for screening, but nonetheless have a psychological problem will not be screened and become at higher risk of nonidentification of their problem. In addition, it is organizationally simpler to intervene with an entire population than to sample parts of the population. Administratively, in many primary care settings, there is no simple way to determine diagnostic status a prior, making sampling by diagnostic category a significant administrative task.

Once a decision is made about who is going to be screened, the next issue to be considered is what is to be screened. Experience and evidence suggest that measures to be used in primary care must be brief. Nevertheless, although brevity is desirable, psychometrically brevity decreases specificity and accuracy.

Some years ago a method was developed for the identification of psychiatric diagnoses in primary care settings called the PRIME MD (Spitzer et al., 1994). The method entailed a patient filling out a questionnaire, staff scoring it, and the physician then asking a set of structured questions based on the questionnaire scoring to ascertain diagnosis. The method entailed up to 8–12 minutes of physician time. I got very excited and went into a staff meeting to explain it and generate support from my physician colleagues. Lise Kowalski, MD, who is one of the very best evaluators of my ideas about psychology in primary care said something like the following, "You want me to take the majority of my visit time to ask patients questions about something that they are not necessarily there to discuss, that will

produce more negative findings than positive, and that the patient may or may not be interested in or willing to pursue?" It was at that moment that I realized that in the primary care setting the trade off between specificity and brevity mitigated towards brevity. The authors of the measure must have also had similar conversations because a later iteration, the Patient Health Questionnaire (Kroenke, Spitzer, & Williams, 2001) is a completely patient-reported measure that can be completed and scored quickly producing highly reliable specificity (Kroenke et al., 2001).

So once again, if it is not viable to use measures that can provide an accurate assessment of a broad range of diagnoses, how does one select what to screen for? The most prevalent method is to screen for one or more of the psychological disorders that occur most frequently in primary care populations. This leads to selecting from depression (with primary care prevalence rates of from 5–12%), substance abuse (with prevalence rates from 10–30%), and anxiety (with prevalence rates of 6–14%) (Katon & Sullivan, 1990). Recent reports indicate an increasing rate of anxiety disorder of almost 20% (Spitzer, Kroenke, Williams, & Lowe, 2006). Because patients who have both depression comorbid with substance abuse have generally even poorer outcomes, screening for both has been suggested (Watkins, Paddock, Zhang, & Wells, 2006). Because of high prevalence and particularly low detection, it has been suggested that panic disorder should also be included (Lowe et al., 2002).

Method

Once the population and content area have been determined, a method of data collection needs to be identified and specific measures selected. Probably the most robust and traditional method for diagnostic screening is a clinical interview conducted by a physician or psychologist. However thorough as this method might be, let us keep in mind that the task is to rapidly identify that a given individual has a high probability of interfering psychological issues. Screening is generally not an effort to establish diagnostic certainty; rather it is a method to identify the need for further assessment. Therefore, a clinical interview is probably not necessary for the task. In addition, such a method limits the number of patients who can be screened and is an expensive task that is probably not going to be supported by most medical systems.

Most contemporary screening methodologies make use of some variant of paper-and-pencil patient report, scored some time after the patient visit with subsequent feedback provided to the physician. Paper-and-pencil self-report has been found to be a better predictor than physician interview (Olson, Dietrich, Prazar, & Hurley, 2006; Spitzer et al., 2006). However, such a method also has limitations. It requires manual scoring or someone to do electronic scanning. Feedback in this method is usually post hoc, delivered to the physician sometime after the patient appointment eliminating the opportunity to intervene rapidly.

Recently, electronic technology, often supported by federal initiatives for screening, have allowed for automated information collection and reporting. One method is the use of a fixed automated kiosk in which a patient responds to questions with the answers scored and output produced in another area for later communication to the physician. Yet another makes use of portable electronic tablets or notebooks that are handed to patients to be filled out in the waiting room or exam room. In this model, after completion, the tablet is put into a docking

station where the data is recorded and scored. In both electronic scenarios, an advantage is that scoring is done automatically, and data can be programmed to be stored in a database for later use in monitoring outcomes and for quality improvement use.

In our office, we have tried to use the tablet technology and generate output in the form of a behavioral health lab slip attached to the patient chart for physician review before the patient is seen. This provides the advantage of physician discussion of the results with the patient during the visit that generated the screening data. With the advent of electronic medical records, there is discussion of the lab slip being sent directly to the record that is reviewed before the visit, eliminating further levels of manual effort. The potential disadvantages of these models include upfront expense, and patient comfort level with electronic technology.

Measures

After methods of collection scoring and information dissemination are determined, measure selection is the next task. It appears that screening measures within a clinical domain (i.e., depression or anxiety) perform similarly (Coyne & Thompson, 2003; Valenstein et al., 2001). There is little evidence to recommend one screening measure over another, so clinicians can choose the measure that best fits (U.S. Preventative Services Task Force, 2002). Although measures of depression are interrelated with anxiety and substance abuse (Olson et al., 2006), it has been suggested that depression and anxiety are distinct entities for screening purposes and should be measured separately (Spitzer et al., 2006).

As mentioned earlier, brief measures are necessary in a primary care setting. However, the briefer the measure, the more important is face-to-face follow-up to confirm the screening results. If the amount of available follow-up time is less, then there will need to be a lengthening of the measure to increase sensitivity (Mitchell & Coyne, 2007). It is a delicate balance because if measures are too long or not specific enough, in both cases they will not be used (Spitzer et al., 2006). Mitchell and Coyne have contrasted three types of measures—ultra short, short, and standard. They note that single-item tests can identify 3 of 10 patients who test positive, 2- or 3-item screens can identify 8 of 10 who test positive. They note however, that shorter measures have higher false-positive rates. Half of positive screens are determined to be false-positives with shorter measures (Mitchell & Coyne, 2007). In addition, when selecting measures there is a need to pay attention to ethnic and racial diversity (Huang, Chung, Kroenke, & Delucchi, 2006) and sensitivity to sex, income, and education (Brown & Rounds, 1995). Currently, the most frequently used primary care screening measure is the 9-item Patient Health Questionnaire (PHQ; Huang et al., 2006). It offers the advantage that specific scores are associated with evidence-based recommendations for intervention.

Screening Criteria

As Coyne and Thompson have pointed out, do not make a decision to screen casually (Coyne et al., 2000). Compared to screening for other medical disorders, screening for behavioral health issues among primary care patients is higher in burden of performance, interpretation, and treatment. Performance requires time-intensive administration and scoring compared to blood tests. There is often a high false-positive rate with limited standard reference points for interpretation, and

treatment requires considerable patient activation and frequent provider monitoring (Nease & Malouin, 2003).

Summarizing, we have identified the following considerations that should be reviewed as screening programs are developed.

- What is the population is to be screened?
- What clinical domains should be screened for?
- What measures are to be used?
- How long should measures be?
- Additional issues to consider include
- How much time can be allocated to collect the data?
- How is the procedure going to happen in your setting?
- Who does it and where?
- How is it scored?

Even after making these core determinations there are multiple implementation considerations that will have a significant impact on the outcome of screening efforts. How and when physicians get the information and the degree of specificity or generalization of the screening information will affect how the information is used (Olson et al., 2006). Historically, screening information has been provided well after it was collected from the patient, giving the physician no opportunity to make use of the data until the next encounter. This places an added burden on the process, requiring recall to make use of the information at the next point in time, which may not occur for some time into the future. In addition, a summary measurement about the potential presence of a disorder (i.e., depression or substance abuse) may not be as useful as specific statements about impact on functioning, such as sleep interruption, concentration difficulties, avoidance of usual activities or places, etc.

Once provided, how is the physician going to use the screening information? Does the potential disposition of screening information include the availability of time and resources to further confirm the presence of diagnosis or the availability of interventions to respond to the identified clinical issues? This will affect the utility of providing the information in the first place (Coyne & Thompson, 2003). Are there rapid follow-up appointments available for the physician to be able to see the patient again quickly for further evaluation or intervention? If referral for collaborative treatment is indicated, are there referral sources available that the physician feels comfortable enough to recommend to their patient? Are there referral and scheduling procedures available within the office to assure the same continuity that is available when making referral to other subspecialties? Can a patient leave the office with a referral appointment in hand? What is the planned response to the 20% of patients who reject treatment after a positive screen (Holzapfel et al., 2007)? What training and planning is needed to assist physicians with patients who present somatic complaints and screen positive, but prefer a medical explanation rather than a psychological explanation (Mitchell & Coyne, 2007)?

It is prudent to plan for administrative questions and issues that are bound to be part of the suggestion to screen, or the decision to develop a screening program. In addition to screening data providing a method to assist in clinical care, it has the potential to promote quality-improvement activities. If one can gather aggregate screening data, it can be used for decision making. Depending upon the populations being screened, the method can identify segments of the local population that appear to be at higher psychosocial risk. This, in turn, can affect service planning. Ultimately of course, psychological screening and progress monitoring can be

contrasted with medical diagnostic and services data, allowing the capture of the holy grail—the identification of medical populations where psychological comorbidity has most impact and where treatment of the comorbidities has the most medical and cost effectiveness.

Of course, in a medical setting there will be the need to identify the costs of the screening effort, both upfront and over time. It is always important to recall that in a total practice population screening as recommended by the USPSTF (2002) there is going to be an extremely high percentage of negative screens. The negative screens still have a cost associated with collecting the data. It is reasonable to be asked to compute the real costs so that an informed administrative and financial decision can be made. Certainly, the broad range of staff involved in the planning and execution of the project has cost consequences. Importantly, particular attention needs to be paid to managing any resistances that emerge (Coyne & Thompson, 2003). Just because you and your medical colleagues are clear about the need and utility of the effort, it does not mean that the rest of the office support staff and nursing staff (who will have tasks in the project) will be equally enthusiastic. If they are not enthusiastic, they may be resistant. If resistant, they can scuttle the enterprise. In summary, the development of a screening program requires a complex set of decisions and is a daunting task.

Case Illustration

Even as there are national recommendations for screening and evidence-based treatments based on screening scores, the challenges of translating them into the reality of day-to-day medical practice is substantial. Evidence-based recommendations and treatments are broad interpretations of evidence. Translation into practice is local and must be responsive to organizational needs and preferences. The following local illustration of the complexity of the translation has consistently attempted to respect the balance between the recommendations from the evidence, and the translation into local needs values and organization. It has been in its initial phases for about 4 years.

Berlin Family Health is a family medicine practice located on the campus of Central Vermont Medical Center in northern Vermont. The practice is part of the University of Vermont's Department of Family Medicine, and its academic health center Fletcher Allen Health Center. The practice has five family medicine physicians, two midlevel practitioners, and two psychologists. There are 40,000–50,000 patient visits per year.

The first discussion within the practice concerning screening for psychological disorders was many years ago with the advent of Spitzer's PRIME MD (Spitzer et al., 1994). At the time, it was deemed unwieldy and too time consuming and had little clinical and administrative support. About 5 years ago I was introduced to electronic tablet data collection at a demonstration by POV Systems (Patient Tools, Littleton, CO; www.patienttools.com). The technology offers a multitude of advantages including ease of programming a variety of measures, and ease of use and completion. It has the ability to both produce summary hard copy of screening results within seconds and have the data became part of a storable database for later analyses. The method could also be used later to retrieve the measures to track clinical changes. It also offers certain disadvantages, such as the need for patients to respond to electronic media with minimal supervision, and it requires small amounts of both front desk and nursing assistance to generate the data, although presently we

are attempting to eliminate the need for front desk staff. Discovery of the technology roughly corresponded with the introduction of a revision of the PRIME MD and the introduction of the Patient Health Questionnaire, specifically the affective disorders module, the PHQ-9 (Kroenke et al., 2001). The PHQ quickly became widely used because of its psychometric strength, its brevity, and its association with empirically derived treatment intervention strategies, which correspond to different scores on the measure.

With a small amount of unrestricted pharmaceutical company funding came support for the hardware and database maintenance costs. The idea was reviewed once more with my family medicine colleagues. There was increased but modest interest because this was one of many good ideas in a busy, clinically based primary care practice. It had merit but no planning time or financial support for the efforts required for implementation. An obstacle to innovation of any sort in primary care is the lack of available professional time for thinking and planning, and understaffing of support staff, leading to a general sense that *anything* else is difficult to accommodate.

A primary care physician and a physician assistant volunteered to pilot the idea. Details were primarily left to me, as project champion. With consultation from my two colleagues who were participating in the pilot, it was decided to start with all patients 18 years of age and older whose appointment was for a physical examination, either preoperative or annual.

It seemed clear that using the PHQ made sense. We modified it, to exclude the ninth suicide risk section, since the score on the eight items predicts the ninth. That is if there is a high enough cumulative score on the eight items, there is a high risk of suicidal ideation, which needs follow-up. It was also felt that not asking about suicide explicitly made the measure more palatable to a primary care population. The practice had gone through some training in identification, assessment, and brief intervention with problem of alcohol abuse so we decided to also screen for substance abuse. We decided to use the Alcohol Use Disorders Information Test because of its brevity and the fact that it has been endorsed by the World Health Organization and the Substance Abuse, Mental Health Services Administration (Babor, de la Fuente, Saunders, & Grant, 1989). During practice administrations of the measures, the time needed to collect the patient data was from 30 seconds to 3 minutes. There is an initial screen on the tablet explaining the rationale and procedure for completing the data collection and a statement about what happens to the results.

At check-in for their appointment, the tablet was handed to the patient by the front office staff with a scripted statement requesting that patients fill out the information in the waiting room and noting that if there is any concern to discuss it with their physician or their nurse. Upon completion the tablet is handed to the nurse who puts it into a docking station where the measures are automatically scored, summary data is sent to a searchable database and a Behavioral Health Lab slip (BHLS) is generated. The BHLS is in the same format as a slip from the laboratory summarizing the results of a test, listing the name of the test (in this case depression and substance abuse) the normal range on the measure, the patient score, and a notation if the score is significantly elevated. If it is elevated, the questions and elevations that were endorsed by the patient are included.

Also included were two questions for later physician completion for research purposes, one concerning disposition of a positive screen and one asking if the data is new or familiar to them about the patient. The lab slip is placed on the physician

chart, to be reviewed in a small number of seconds before the physician goes into the examination room with the patient, to be further explored at the physician's discretion.

In our practice psychologists are part of the practice and we have occasional access to onsite psychiatric consultation. We also have systems in place to facilitate within-office referral and appointment scheduling.

How Screening Has Worked in Our Setting

The effort has been a very difficult frustrating process that has not yet reached a mature status. With all the thought attention and effort, it remains a work in progress. There have been two periods of regular usage. During the first usage phase, there was a flurry of activities creating and operationalizing processes. Scripts were created. The tablet was programmed. Nursing and front desk staff was trained. The project began and shortly thereafter, the practice manager left. With the organizational problems this generated internally, things just came to a standstill. After a new office manager was recruited, we broached the project again. The office manager was supportive as long as the task did not cause too much strain and as long as I maintained primary responsibility. Yet another round of training occurred. We then started data collection again. For about 4 months things went well. Data was collected and the process appeared to be working as planned. Then it stopped. No one officially stopped it—the data just no longer appeared. Because the rest of primary care life continued frantically, we just shrugged our shoulders and said we would come back to it.

As this is being written, it is roughly a year after the process stopped. The practice has been awarded a 3-year grant by the Vermont Department of Health to focus on quality improvement project to assist the recognition and treatment of primary care patients with mental health and/or substance abuse diagnoses. A primary work task of the project is to reinitiate the screening procedure and report on data collected. We have been funded to think, talk, and figure out obstacles to implementation and generate solutions. In early discussions since the funding, an interesting point has surfaced about how and why what was apparently working just stopped (Figure 1).

Organizationally it was identified that patient registration for visits was not working effectively and overwhelming the front desk, which was the starting point to access the screening procedure. During the same time we were implementing the screening procedure, the practice was transitioning to a different registration procedure. The staff on the front desk, who also initiated the screening procedure, were so overwhelmed that there just was not enough time to both register the patient and initiate the screening procedure so the less critical task, the screening procedure,

NAME: _____ Date: _____

	SCORE	Above Range (+)	(Normal Range)
Depression			(0 – 8)
Anxiety			(0 – 7)
ETOH			(0)

Figure 1. The Behavioral Health Lab slip.

got left behind. This was not an overt sabotage; rather the procedure did not fit into the flow of the organization. Sufficient organizational reengineering (Dietrich et al., 2004) did not occur and the predictable outcome of a lack of an engineered process resulted. We laid on a new task that overwhelmed the existing structure and it failed. We had committed two of the deadly sins of collaborative care.

The deadly sins are a set of actions often taken in attempts to implement collaborative psychological and medical care that usually doom the effort to failure (Kessler, 2008). Two of the sins are "Don't worry about organizational and financial issues—they will take care of themselves," and "Every system is perfectly designed to get the outcomes it gets." In retrospect, Kessler's observations are germane to what happened in our practice.

There has been agreement among professional, nursing, and administrative staff that better identification of medical psychological comorbidities is a good idea that will assist patient care. Screening as a method to accomplish that goal is not novel in a primary care setting: Screening occurs for many disorders. However, the method of collecting the information and the need for staff involvement in the data collection are not only novel, but they add a layer of complexity and additional effort to daily practice. Most of this effort falls on staff that are not primary decision makers—front office staff and nursing.

To have thought that the implementation was going to proceed, impacting on busy work lives because we had a staff meeting and explained that we were going to do this and it really wasn't going to take much time from them, assumed that a significant organizational and cultural change was going to occur. Little attention was paid to buy-in from the affected parties or the organizational infrastructure changes necessary to accommodate the change. The issues most certainly did not just take care of themselves. The outcomes that were produced were exactly what one would predict with such inattention. We achieved the outcomes we had inadvertently designed into the intervention—it did not work.

The grant is allowing us to explore the issue in some depth. Within the grant structure there is a physician champion. There is a champion in senior administration in the Department of Family Medicine, and the Department chair has given his blessing. We are after all this time, positioned to better plan, to pay for the real costs, to redesign, and pay even more attention to the design elements and organizational issues that have limited implementation, so we can shift the resistances and be successful in the next iteration.

Summary

Gilbody and colleagues recent meta-analysis of the effectiveness of screening as a sole intervention concluded that if used alone, case-finding or screening questionnaires for depression appear to have little or no impact on the detection and management of depression by clinicians. Recommendations to adopt screening strategies using standardized questionnaires without organizational enhancements are not justified (Gilbody et al., 2008).

It is apparent that here are at least three critical dimensions necessary for screening efforts to have impact. There must be adequate resources to interpret positive screens (Palmer & Coyne, 2003). There must be appropriate available interventions (Coyne & Thompson, 2003). Screening results must be routinely used to effect treatment (Gilbody et al., 2001). In addition, there is the need to pay attention to potential negative effects (Coyne et al., 2000) such as patient reaction or refusal, staff

resistance to the process, lack of clearly thought out implementation strategies (Kaner et al., 2002), and costs that are not readily apparent (Palmer & Coyne, 2003). In short, behavioral health screening is more complicated than it seems.

That is not so surprising, really. Implementation of screening in a medical practice is a complex organizational change, no matter how simple or complicated the procedure. Organizational change is hard to accomplish in any setting and perhaps particularly difficult in medicine. Medical practices are organizations that give much more attention to what they do, that is deliver patient care, than how they do it, that is the organization of how care is delivered. In addition, organizational change requires planning time and implementation resources, things that are often lacking in primary care practice.

In addition, organizational changes that focus on psychological dimensions of medical care are even more fraught with difficulty. Medical practice is currently struggling with the organization of how care is delivered. Such changes are mandated, usually not voluntary. Changes are mandated by the systems that own or manage practices, which must focus on efficiencies and cost savings to live long and prosper. Changes are mandated by a variety of different payers that each have different sets of pharmacy formularies, prior approval procedures, and quality improvement requirements. Changes are mandated by regulatory and national health policy dictations such as chronic care initiatives and preventable medical errors mandates, which intend to improve patient care and achieve financial efficiencies, but nonetheless create intrusions on every day clinical life.

Currently, psychological interventions in medical organizations have no mandate. If implemented, they are either the result of collegial support of good ideas, commitment to research findings, or in anticipation of clinical and cost efficiencies. They are often unfunded or inadequately funded. Frankly, structurally, among the mandated competing demands, they just are not that important. In addition, interventions that improve identification in medical patients may not be totally welcomed by physicians receiving the information.

In Jerome Groopman's recent volume *How Doctors Think*, he observes, "A wealth of research shows that patients thought to have a psychological disorder get short shrift from internists and surgeons and gynecologists. As a result, their physical maladies often never get diagnosed or the diagnosis is delayed" (Groopman, 2007, pp. 264–265). "The lingering stigma that exists in medicine… about psychological distress and its ramifications through the body, stands as a roadblock to relieving the pain and misery of many patients" (Groopman, 2007).

Therefore, even though the provision of behavioral screening data may well assist medical diagnosis, there may equally be a bias in a medical practice not to know the information. An obstetric colleague who frequently refers patients to me has remarked that if a woman starts talking about emotional material he redirects the conversation to something else because historically his involvement has taken too much time and effort with limited benefit. If the woman persists in discussing such issues, he calls me. Therefore, a physician or practice agreeing to participate in the development of screening requires a belief that the information will not open the "black hole" of mental health, that receiving the information can assist in the medical treatment of the patient and that she or he has confidence in the resources available to provide the psychological services.

So, should there be screening for psychological disorder or comorbidity in primary care settings? It should be considered, but only if it is part of an organizational change process that can improve identification and treatment of psychological issues

that are part of the medical presentation. Screening for the presence of psychological disorder can and should assist that process. However, to do so requires grappling with and resolving most of the issues discussed in this article. If not, then because every intervention gets the outcomes it is designed to get, the results will be predictably limited and patient care will probably not be positively affected—once more marginalizing the efforts to include psychology as an integral part of medicine.

References

Academy of Psychosomatic Medicine. (1996). Mental disorders in general medical practice: An opportunity to add value to healthcare. Behavioral Healthcare Tomorrow, 72, 55–62.

Babor, T.F., de la Fuente, J.R., Saunders, J., & Grant, M. (1989). AUDIT: The Alcohol Use Disorders Identification Test: Guidelines for use in primary health care. Geneva: World Health Organization.

Brown, R., & Rounds, L. (1995). Conjoint screening questionnaires for alcohol and other drug abuse:criterion validity in a primary care practice. Wisconsin Medical Journal, 94, 135–140.

Brown, R.L., Leonard, T., Saunders, L.A., & Papasouliotis, O. (1997). A two item screening test for alcohol and other drug problems. Journal of Family Practice, 44, 151–160.

Callahan, E.J., Bertakis, K.D., Azari, R., Robbins, J.A., Helms, L.J., & Leigh, J.P. (2002). Association of higher costs with symptoms and diagnosis of depression. Journal of Family Practice, 51, 540–544.

Cowley, D., Khan, A., Khan, A., Harezlak, J., Tu, W., & Kroenke, K. (2003). Somatic symptoms in primary care: Etiology and outcome. Psychosomatics, 44, 471–478.

Coyne, J.C., & Thompson, R. (2003). Psychologists entering primary care cannot be bought for $24 worth of beads. Clinical Psychology Science and Practice, 19, 102–110.

Coyne, J.C., Thompson, R., Palmer, S.C., Kagee, A., & Maunsell, E.M. (2000). Should we screen for depression? Caveats and potential pitfalls. Applied and Preventative Psychology, 9, 101–121.

Dietrich, A.J., Oxman, T.E., Williams, J.W., Schulberg, Herbert, C., Bruce, M.L., et al. (2004). Re-engineering systems for the treatment of depression in primary care: Cluster randomized controlled trial. British Medical Journal, 329, 602.

Fries, J.F.K., Koop, C.E., Beadle, C.E., Cooper, P.P., England, M.J., Greaves, R.F., et al. (1993). Reducing health care costs by reducing the need and demand for medical services. New England Journal of Medicine, 329, 321–325.

Gilbody, S., Sheldon, T.A., & House, A.O. (2008). Screening and case finding instruments for depression: A meta-analysis. Canadian Medical Association Journal, 178, 997–1003.

Gilbody, S.M., House, A.O., & Sheldon, T.A. (2001). Routinely administered questionnaires for depression and anxiety. British Medical Journal, 322, 406–409.

Groopman, J. (2007). How doctors think. Boston: Houghton Mifflin.

Holzapfel, N., Zugck, C., Muller-Tasch, T., Lowe, B., Wild, B., Schellberg, D., et al. (2007). Routine screening for depression and quality of life in outpatients with congestive heart failure. Psychosomatics, 48, 112–116.

Hoyer, J., Becker, E.S., Neumer, S., Soeder, U., & Margrat, J. (2002). Screening for anxiety in an epidemiological sample: Predictive accuracy of questionnaires. Anxiety Disorders, 16, 113–121.

Huang, F.Y., Chung, H., Kroenke, K., & Delucchi, K.L. (2006). Using the Patient Health Questionnaire-9 to measure depression among racially and ethnically diverse primary care patients. Journal of General Internal Medicine, 21, 547–552.

Kaner, E.F., Lock, C.A., McAvoy, B.R., Heather, N., & Givarry, E.G. (1999). A RCT of three training and support strategies to encourage implementation of screening and brief

alcohol intervention by general practitioners. British Journal of General Practice, 49, 699–703.

Kaner, E.F., McAvoy, B.R., Heather, N., & Givarry, E.G. (2002). An RCT of three training and support strategies to encourage implementation of screening and brief alcohol intervention by general practitioners. British Journal of General Practice, 49, 699–703.

Katon, W., & Sullivan, M. (1990). Depression and chronic medical illness. Journal of Clinical Psychiatry, 5, 3–11.

Kessler, R. (2008). How I learned about integration by failing miserably: The deadly sins of integration. In R. Kessler & D. Stafford (Eds.), Collaborative medicine case studies: Evidence in practice (pp. 39–50). New York: Springer.

Kessler, R., & Stafford, D. (2008). Primary care is the defacto mental health system. In R. Kessler & D. Stafford (Eds.), Collaborative medicine case studies: Evidence in practice. New York: Springer.

Kroenke, K., & Mangelsdorff, A.D. (1989). Common symptoms in ambulatory care: Incidence, evaluation, therapy and outcome. American Journal of Medicine, 86, 262–266.

Kroenke, K., Spitzer, R.L., & Williams, J.B. (2001). The PHQ-9: Validity of a brief depression severity measure. Journal of General Internal Medicine, 16, 606–613.

Lin, E.H., Katon, W., Rutter, C., Simon, G., Ludman, E.J., Von Korff, M., et al. (2006). Effects of enhanced depression treatment on diabetes self care. Annals of Family Medicine, 4, 46–53.

Lowe, B., Grafe, K., Zipfel, S., Spitzer, R.L., Herrmann-Lingen, C., Witte, S., et al. (2002). Detecting panic disorder in medical and psychosomatic outpatients. Comparative validation of the Hospital Anxiety and Depression Scale, the Patient Health Questionnaire, a screening question and physician's diagnosis. Journal of Psychosomatic Research, 55, 515–519.

Mitchell, A.J., & Coyne, J.C. (2007). Do ultra short screening instruments accurately detect depression in primary care? British Journal of General Practice, 57, 144–151.

Narrow, W.E., Regier, D.A., Rae, D.S., Manderscheid, R.W., & Locke, B.Z. (1993). Use of services by persons with mental and addictive disorders: Findings from the National Institute of Mental Health Epidemiological Catchment Area program. Archives of General Psychiatry, 50, 95–107.

Nease, D.E., & Malouin, J.M. (2003). Depression screening: A practical strategy. Journal of Family Practice, 52, 118–124.

Olson, A., Dietrich, A.J., Prazar, G., & Hurley, J. (2006). Brief maternal depression screening at well-child visits. Pediatrics, 118, 207–216.

Oslin, D.W., Ross, J., Sayers, S., Murphy, J., Kane, V., & Katz, L.R. (2006). Screening assessment and management of depression in VA primary care clinics. Journal of General Internal Medicine, 21, 46–50.

Palmer, S.C., & Coyne, J.C. (2003). Screening for depression in medical care: Pitfalls, alternatives and revised priorities. Journal of Psychosomatic Research, 54, 279–287.

Piette, J.D., Richardson, C., & Valenstein, M. (2004). Addressing the needs of patients with multiple chronic illnesses: The case of diabetes and depression. American Journal of Managed Care, 10, 152–162.

Pignone, M.P., Gaynes, B.N., Rushton, J., Burchell, C.M., Orleans, C.T., Mulrow, C.D., et al. (2002). Screening for depression in adults: A summary of the evidence for the U.S. Preventive Services Task Force. Annals of Internal Medicine, 136, 765–776.

Regier, D., Narrow, W., Rae, D., Manderscheid, R., Locke, B., & Goodwin, F. (1993). The de facto U.S. mental and addictive services system: Epidemiologic catchment area prospective 1-year prevalence rates of disorders and services. Archives of General Psychiatry, 50, 85–94.

Simon, G., Chisholm, D., Treglia, M., & Bushnell, D. (2002). Course of depression, health services costs, and work productivity in an international primary care study. General Hospital Psychiatry, 24, 328–335.

Simon, G., Revicki, D., Heiligenstein, J., Grothaus, L., VonKorff, M., Katon, W., et al. (2000). Recovery from depression, work productivity, and health care costs among primary care patients. General Hospital Psychiatry, 153, 153–162.

Simon, G.E., Katon, W.J., Lin, E.H., Rutter, C., Manning, W.G., Von Korff, M., et al. (2007). Cost-effectiveness of systematic depression treatment among people with diabetes mellitus. Archives of General Psychiatry, 64, 65–72.

Simon, G.E., Von Korff, M., & Barlow, W. (1995). Health care costs of primary care patients with recognized depression. Archives of General Psychiatry, 52, 850–856.

Spitzer, R., Williams, J., Kroenke, K., Linzer, M., deGruy, F., Hahn, S., et al. (1994). Utility of a new procedure for diagnosing mental disorders in primary care. The PRIME-MD 1000 study. Journal of the American Medical Association, 122, 1749–1756.

Spitzer, R.L., Kroenke, K., Williams, J.B., & Lowe, B. (2006). A brief measure for assessing generalized anxiety disorder. Archives of Internal Medicine, 166, 1092–1097.

Thompson, M. (2000). Five giant leaps to integrating health care delivery and ways to drive organizations to leap or get out of the way. Journal of Ambulatory Care Management, 23, 1–18.

Unutzer, J., Patrick, D., Simon, G., Grembowski, D., Walker, E.A., Rutter, C., et al. (1997). Depressive symptoms and the cost of health services in HMO patients age 65 nd over: A four-year prospective study. Journal of the American Medical Association, 277, 1618–1623.

U.S. Preventative Services Task Force. (2002). Screening for depression: Recommendations and rationale. Annals of Internal Medicine, 136, 765–776.

Valenstein, M., Vijan, S., Zeber, J.E., Boehm, K., & Butler, A. (2001). The cost utility of screening for depression in primary care. Annals of Internal Medicine, 134, 345–360.

Von Korff, M., Katon, W., Lin, E., Simon, G., Ludman, E., Oliver, M., et al. (2005). Potentially modifiable factors associated with disability among people with diabetes. Psychosomatic Medicine, 67, 233–240.

Watkins, K., Paddock, S.M., Zhang, L., & Wells, K.B. (2006). Improving care for depression in patients with comorbid substance misuse. American Journal of Psychiatry, 163, 125–132.

Whooley, M.A. (2006). Depression and cardiovascular disease: Healing the broken hearted. Journal of the American Medical Association, 295, 2874–2881.

Williams, J., Palmes, G., Kurt, K., Pulley, A., & Meschan, J.F. (2005). Referral by pediatricians of children with behavioral health disorders. Clinical Pediatrics, 44, 343–349.