THE CAMPBELL COLLABORATION

Protocol:
Deployment of Personnel to Military Operations: Impact on Mental Health and Social Functioning

Martin Bøg, Maia Lindstrøm, Trine Filges, Stephanie Vincent Lyk-Jensen, & Anne-Marie Klint Jørgensen

Submitted to the Coordinating Group of:

☐	Crime and Justice
☒	Education
	☐ Disability
☐	International Development
	☐ Nutrition
☒	Social Welfare
☐	Other:

Plans to co-register:

☒	No		
☐	Yes	☐ Cochrane	☐ Other
☐	Maybe		

Date Submitted: 10 April 2014
Date Revision Submitted: 15 May 2014
Approval Date: 22 September 2014
Publication Date: 1 November 2014

1 Background

1.1 DESCRIPTION OF THE CONDITION

The landscape of military deployment has changed dramatically since 1989. While military deployment by its very nature carries substantial risk of exposure to events and experiences that may physically and mentally impair those directly engaged in combat and those engaged in support functions, there are a number of reasons why the extent of this issue is arguably of even greater importance than previously.

First, the types of security threats and military operations have changed substantially since 1989. The fall of the Iron Curtain released the tight grip of the former USSR on the Eastern Bloc, and in many developing nations across the globe where the Cold War had been fought by proxy. As a consequence, a dramatic increase in peacekeeping and peace-enforcing operations ensued, many under the auspices of the United Nations. In the period from 1989 to 1994, the UN Security Council authorized a total of 20 new operations, raising the number of peacekeepers from 11,000 to 75,000 (United Nations [UN], 2014a). In addition, the post-Cold War period is characterised by a shift away from more conventional conflicts between nations towards, for example, the War on Terror, support for insurgencies or counterinsurgencies, and operations to protect civilians in countries under threat of civil war. During this time ad hoc international military coalitions have been formed, including during the Gulf War (1990-91) and in the recent operations in Iraq and Afghanistan.

Second, many nations have gone from having a large proportion of their military personnel on active duty to having a large proportion of military personnel serve as reservists (Department of Defense [DoD], 2008). This has affected the types of people who are deployed; reservists differ from active duty personnel in training received, physical and mental screening, frequency of civil careers, and probably in unobserved ways as well. For example, as a consequence of the large scale commitment of the US, and in turn that of her allies, to military operations in Afghanistan and Iraq, an increasing share of those deployed consists of reservists and national guards. Currently,[1] the UN has 15 peacekeeping missions worldwide[2] and a special political mission in Afghanistan[3] (UN, 2014b), and NATO has 5

---

[1] January, 2014.

[2] UN peacekeeping operations include: UNTSO, UNMOGIP (since 1949), UNFICYP (since 1964), UNDOF (since 1974), UNIFIL (since 1978), MINURSO (since 1991), UNMIK (since 1999), UNMIL (since 2003), UNOCI (since

operations internationally with multilateral participation[4] (NATO, 2014). Furthermore, there are international coalitions with active military operations in other countries, e.g. Iraq.[5] In the period 2001 to 2008, approximately 1.64 million US troops were deployed to Afghanistan or Iraq (Tanielian & Jaycox, 2008). As recently as 2011, Canada had over 3,500 military personnel deployed overseas (National Defence and Canadian Forces, 2011), Australia had approximately 2,900 soldiers deployed to international missions (Australian Department of Defence, 2011), and Denmark had approximately 1,500 military personnel deployed to international missions and has had approximately 50,000 military personnel deployed to international military missions since 1992 (Danish Armed Forces, 2011). As a consequence of the growing demand for international security, larger proportions of military personnel are being deployed, re-deployment is becoming more frequent, and breaks between deployments are being shortened in order to meet demand (Hosek, Kavanagh, & Miller, 2006).

Third, technological advancements in body armor and other military technology, medical technologies, and military tactics, have increased chances of survival among those who face military deployment (Regan, 2004; Warden, 2006).

Set against this change in the scope and nature of military deployments are some of the consequences of deployment. Returning personnel will, depending on mission type and/or assignment, have been exposed to a wide variety of stress factors, such as witnessing violent death, physical abuse, dead and/or decomposing bodies, maimed soldiers and civilians, been captured as prisoners of war, or witnessed movements of refugees as a result of civil war. Frequently, military personnel face deployment to theatres of insurgence or counter-insurgency, where insurgents operate using guerrilla tactics such as suicide missions and road-side bombings (e.g., improvised explosive devices). Deployments to such theatres carry their own risk scenarios, including traumatic brain injury (TBI) and loss of limb (Tanielian & Jaycox, 2008). The lack of a real frontline of war implies that even support troops face increased risk of combat exposure (Street, Vogt, & Dutra, 2009). In addition, such experiences frequently take place within a context that is very different from everyday life: much of the social support network of the deployed is missing, as family and friends are far away, use of civilian communication systems may be difficult, there is often a lack of personal space, and living quarters may be unhygienic (MHAT-IV, 2006).

2004), MINUSTAH (since 2004), UNAMID (since 2007), MONUSCO (since 2010), UNISFA (since 2011), UNMISS (since 2011), MINUSMA (since 2013).

[3] UNAMA, Afghanistan (since 2002).

[4] NATO operations include: Kosovo (since 1999), Monitoring the Mediterranean Sea (since 2001), Afghanistan (since 2003), Supporting the African Union (since 2007), and Counter-piracy off the Horn of Africa (since 2009).

[5] The Multi-National Force – Iraq (MNF–I) was a military command, led by the United States, which was responsible for Operation Iraqi Freedom. Multi-National Force – Iraq replaced the previous force, Combined Joint Task Force 7, on 15 May 2004, and was later itself reorganized into its successor, United States Forces – Iraq, on 1 January 2010. The Force was significantly reinforced during the Iraq War troop surge of 2007. As of August 2009, all non-U.S. coalition members had withdrawn from Iraq.

The majority of deployed military personnel who return home will return and readjust successfully to either civilian life or life on a military base (Tanielian & Jaycox, 2008). However, combat exposure and other stressors increase the risk of physical and psychological trauma and, as a result, a substantial proportion of those returning from deployment to military operations abroad face the real risk of developing adverse effects to mental health and social functioning (Hoge, Auchterlonie, & Milliken, 2004). These include increased risk of posttraumatic stress syndrome (PTSD), depression, anxiety, anger, and substance abuse (Andersen, 1998; Dohrenwend, Turner, Turse, Adams, Koenen, & Marshall, 2006; Helzer, Robins, and McEvoy, 1987; Hoge et al., 2004; Ishøy et al., 1999; Jacobson et al. 2008; Jakupcak et al., 2007; Larson, Highfill-McRoy, & Booth-Kewley2008; O'Brien & Hughes, 1991; O'Toole, Schureck, Marshall, Grayson, & Dobson, 1999; Pitman, Altman, and Macklin, 1989; Rona et al., 2009; Tanielian & Jaycox, 2008). Estimates presented in Tanielian and Jaycox (2008) on the prevalence PTSD among those serving in Afghanistan and Iraq presents numbers ranging from 4-45% of those deployed[6].

The severity and duration of PTSD symptoms can vary greatly between individuals. PTSD symptoms generally begin shortly after the trauma, however the onset may be delayed for months or years (Bisson, 2007; Friedman; 2006). Atkinson, Guetz, and Wein (2009) estimate that, due to the lag in diagnosing PTSD and the fact that, as a result, some servicemen with undetected PTSD are re-deployed to combat zones, the cross-sectional estimates may underestimate the true incidence of PTSD by as much as 100%. Their results suggest that as many as 300,000 US Army and Marine service members may be affected by PTSD.

Adverse mental health outcomes have a direct effect on individual wellbeing, but, in addition, have a number of indirect effects. For example, the families of deployed personnel who return with mental health problems may face strain in family relations, and mental health problems may be transferred between generations, particularly if the affliction is untreated (Davidson and Mellor, 2001; Dekel & Goldblatt, 2008; O'Toole, Outram, Catts, & Pierse, 2010). Furthermore, additional resources at the societal level must be allocated to screen, diagnose and treat the consequences of deployment. The experience of trauma can also lead to posttraumatic growth, understood as an individual sense of strength upon having successfully coped with trauma (Ramos & Leal, 2013). In the context of military deployment, personnel have reported positive post-deployment outcomes such as increased self-esteem, personal development, and strong peer bonding (Danish Ministry of Defense, 2010; Rona et al., 2005; Thoresen, 2006).

These complicated issues raise important questions about the effects of deployment on those servicemen and -women who return from military deployment. In particular, we are interested in the effects of deployment on mental health and social functioning. The answers

---

[6] We note that these estimates rely on different samples and methodologies. For example some studies are restricted to looking only at deployed personnel, and thus lack a comparison group. Therefore the estimates cannot be given a causal interpretation in the sense of measuring the effect of deployment on PTSD.

to these questions are important both for those on active duty, and for those who return to face reintegration into civilian life. As such, the main objective of this review is to assess what is known about the causal effect of deployment on mental health and social functioning. One of the objectives of this review is that it should inform current decision making. In other words, our effect measures should provide information regarding the type of deployments and operations and their consequences that are likely to occur today. We therefore believe it is timely and important to consider and assess effect measures given the new conditions under which military deployment has taken place since the fall of the Iron Curtain in early 1989.

We are aware that a tight causal link between deployment and mental health outcomes will be difficult to establish. Due to the nature of research in the field, we are unlikely to uncover any studies based on controlled trials. Instead we expect to uncover observational studies. Within the broad class of observational studies, a distinction can be drawn between (a) correlational studies which simply seek to measure incidence in the military population regarding outcomes of interest (e.g., PTSD), and its correlates ("risk" and "resilience" factors), and (b) studies that attempt to create a credible comparison group to the sample of deployed. A credible comparison group might consist of a sample from the military population that did not deploy. If this group is balanced on important confounders, we may have credible reasons as to why a comparison can be interpreted as giving *some* evidence regarding the causal effect of deployment. While conclusions about causal effects must be very tentative, it is important to extract and summarize the best evidence available. In our efforts to present the best available evidence we will consider evidence from any nation that has deployed troops to international military operations since 1989.

1.2 DESCRIPTION OF DEPLOYMENT

The primary condition under consideration is deployment to an international military operation. Deployment to a military operation is not a uniform condition; rather, it covers a range of scenarios. Military deployment is defined as performing military service in an operation at a location outside the home country for a limited time period, pursuant to orders.

Staff at military bases across the globe experience varying degrees of extreme and challenging conditions. Military deployment differs from most civilian tasks due to the risk to loss of life, and more generally to the palette of risk factors that military personnel are exposed to while deployed. The particular focus of this review is on the effects of deployment on the mental health and social functioning of personnel returning from deployment. The multi-faceted nature of deployment incorporates a wide variety of conditions and circumstances due to variation in the conflict and setting, type of operation, tasks, combat intensity, characteristics of those deployed, and duration of deployment itself. We expect many of these factors to moderate the effects of interest to this review.

1.2.1   Type of Mission

International military missions may be peacekeeping or peace-enforcing missions under the UN, NATO, multilateral military coalitions or other international agencies. Such operations may involve the deployment of air, navy and army personnel as well as support staff who are not directly involved in military confrontation.

International military operations as studied in this review fall in the categories of low intensity conflicts (LIC) or conventional warfare. Conventional warfare is defined as warfare, other than guerrilla/counterinsurgent warfare, conducted without the use of nuclear, biological, or chemical weapons (Oxford University Press, 2001). Low intensity conflict is defined by the US department of defense: *"Low intensity conflict is a political-military confrontation between contending states or groups below conventional war and above the routine, peaceful competition among states"* (Department of Defense, 2011). LIC can be classified in four major operational categories (Department of Defense, 2011; Global Security, 2011):

- support for insurgencies and counterinsurgencies

- combating terrorism

- peacekeeping operations

- peacetime contingency operations

LIC frequently involves protracted struggles between competing principles and ideologies. LIC ranges from subversion to the use of armed force. It is waged by a combination of means, employing political, economic, informational, and military instruments. Low intensity conflicts are often localized, generally in developing countries, but contain regional and global security implications (Global Security, 2011).

Within low intensity conflict and conventional war there can be great variation in levels of combat intensity. Peacekeeping missions may be non-combat or low combat intensity deployments and relatively free from conflict exposure and be relatively stress free. An example of a non-combat mission is the UN peacekeeping mission in Cyprus. At the other end of the range we find high combat intensity missions as seen in missions by NATO and coalition forces deployed to the Middle East. Military operations in Afghanistan and Iraq are examples of deployments that involve high and frequent combat exposure.

Military personnel may also be deployed to areas struck by natural disaster or military bases overseas not included in a military mission or operation, e.g., US military bases in Europe. This type of deployment is outside the scope of the review.

1.2.2 Type of Service

Deployed personnel may come from different military populations, depending on country of origin and scale of operations. Some countries operate volunteer military such as the UK and the US, while other countries rely fully or partially on mandatory conscription to the military. As a consequence, the underlying characteristics of military personnel, including mental health characteristics, are not uniform. All deployed servicemen and women, regardless of deploying nation, are included in the review.

Deployed personnel may derive from different military population pools including personnel on active duty, reserves forces, and national guards. Military reserve forces are members of the armed forces who combine a civil career with a military career. Reserve forces have a signed contract with the military and may be called in for active duty if needed. At any given time the demand for military personnel affects which subpopulation is at risk of deployment. For instance, at times of low demand only active duty members face risk of deployment, whereas under times of high demand, such as for example under Operation Enduring Freedom to Afghanistan, large contingents of military reserve forces and national guards of the US face risk of deployment (Tanielian & Jaycox, 2008). Military personnel, regardless of type of service, are relevant to this review provided they have faced deployment to international military operations.

1.2.3 Type of Branch

All branches of deployed military personnel are relevant for this review. That is to say, personnel from army, navy, Marine Corps, air force, Special Forces, etc. are all relevant. Both combat and non-combat personnel face deployment to international missions and are thus relevant to the review. Deployed non-combat personnel include personnel with transport/logistics functions, health and medical staff, information technology and communication staff, and general technical support personnel. As an example, the NATO mission to Afghanistan included components of deployed personnel overseeing training of local national police and military units (NATO, n.d.).

Deployed personnel regardless of military rank are relevant to the review. Military rank affects the types of operation-specific tasks assigned to deployed personnel, and therefore rank can affect the mix of stress factors to which personnel are exposed. We therefore consider military rank to be an important control.

1.3 POSSIBLE CONSEQUENCES OF DEPLOYMENT

Deployed personnel participate in military operations in order to fulfil a military objective. Military deployment itself involves increased risk of exposure to a number of stress and risk factors which are linked to mental health outcomes. Mental health problems may be long lasting, can influence the daily life of individuals profoundly, and can impact reintegration into civil society.

To highlight these issues, Figure 1 presents a schematic overview linking events during deployments to outcomes. The events or series of events of interest to this review are events that may occur during deployment to war zones: the experience of physical and/or psychological trauma. The American Psychiatric Association (1994) defines (psychological) *trauma* according to two criteria: (i) a person "experienced, witnessed or was confronted with an event or events that involved actual or threatened death or serious injury, or a threat to the physical integrity of self or others," and (ii) "the person's response involved intense fear, helplessness, or horror" (pp. 427-28). Notice that the experience of the event alone is not sufficient. In order to induce trauma the individual must have an emotional response of fear, horror, or helplessness. The experience of trauma does not necessarily lead to any mental disorders, but, if it does, it can lead to a multitude of disorders such as: major depression, specific phobias, panic disorder, personality disorders, and PTSD (Hyams, Wignall, & Roswell, 1996). The individual experience of physical trauma, such as loss of limbs, may also produce psychological trauma, but need not.
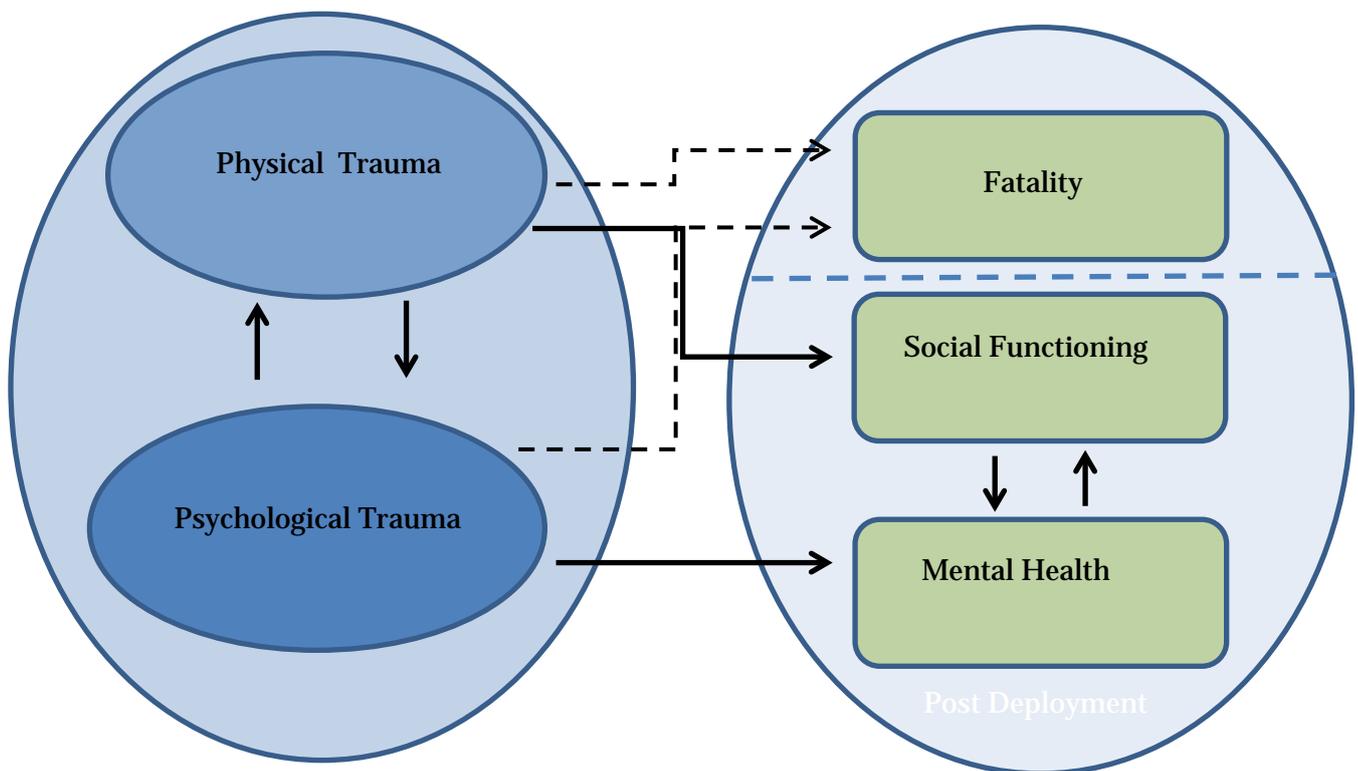

Figure 1. Military Deployment and Post-deployment Outcomes

Deployment entails an increased risk of exposure to such events compared to civilian life. Physical trauma may be experienced by the individual who is, for example, wounded in combat or injured by suicide bombers, improvised explosive devices (IEDs), etc. The individual may also witness physical harm to others, such as the killing of civilians, mass graves, explosions, watching fellow combatants sustain injury or death during combat, etc. The witnessing of such events may cause psychological trauma to the individual.

The experience of physical and/or psychological trauma may have a number of consequences. First, the experience of personal trauma may directly lead to fatality. Such a dire consequence of trauma is outside the scope of this review. Second, the experience of trauma may affect the mental health and social functioning of those returning from service. Adverse mental health outcomes may in turn indirectly affect measures of social functioning such as employment, homelessness, ability to sustain marital relations, etc. The experience of physical trauma can have a direct effect on social functioning. For example, the loss of a limb may make it impossible to return to specific civil careers, which may in turn indirectly affect the mental well-being of the veteran.

1.3.1   Impact on Mental Health

The effects of deployment on mental health received relatively little attention prior to World War I (WW1), although attempts at diagnosing war syndromes go at least back as far as the U.S. Civil War (Hyams et al., 1996). During WW1, deployed personnel suffered from what became known as "soldier's heart." Affected soldiers reported, amongst other symptoms, fatigue, headache, confusion, and lack of concentration. An acute combat stress reaction (CSR) was also diagnosed ("shell shock"). In response, a clinical research program was developed during the war. After the Vietnam War, returning veterans were diagnosed with post-traumatic stress disorder (PTSD), which was subsequently detected among veterans of the Korean Conflict and WW2 (Weisæth, 2002; Jones et al., 2002). Unlike acute combat stress reaction, which is an immediate result of psychological trauma, PTSD refers to the long term consequences of psychological stress (Hyams et al., 1996). PTSD is characterised by a range of symptoms: "[...] distressing thoughts, feelings, and images that recapitulate the traumatic event, avoidance of stimuli associated with the event, emotional numbing of responsiveness, [...]" (Fairbank, Ebert, & Cadell, 2001, p. 184). PTSD is diagnosed only after symptoms have persisted for at least 30 days (Tanelian & Jaycox, 2008). Another comorbidity of deployment is major depression disorder (MDD). MDD is a mood disorder that affects the everyday functioning of an individual. Amongst the common symptoms are disinterest in activities, significant weight gain/loss, insomnia/hypersomnia, feelings of worthlessness, thoughts of suicide, and/or suicide plans.

Deployment mainly affects the mental health of those deployed through increased risk of exposure to trauma (Dohrenwendt et al., 2006; Larson et al., 2008). Exposure to trauma, either by participating in armed combat or as a witness to combat or the consequences of combat, e.g., military medical teams and logistic officers, may adversely affect mental health.

Deployed military personnel, particularly combat soldiers, experience risk of fatality, and some will see their peers wounded, maimed or killed.

Specific circumstances may represent specific stress factors. An example is the UN peacekeeping mission in former Yugoslavia. The mission operated under a limited UN mandate, which allowed only neutral surveillance and the prohibition of the use of force. There, deployed soldiers witnessed extremely stressful situations when they observed direct attacks upon the civilian population. Another potentially contributing factor is the fact that deployment by its very nature implies the (temporary) removal of individuals from their everyday environment, including separation from family, friends, colleagues and support network. Relations to friends and family during, and particularly after deployment, appear to be correlated with psychological stress reactions and impact social functioning (Christensen, 2001; King, King, Fairbank, Keane, & Adams, 1998). Similarly, Fontana and Rosenheck (1994) found that the homecoming experience of veterans played in important role in the development of PTSD.

The epidemiological literature on PTSD has revealed a number of risk factors for the development of PTSD following trauma in the general population as well as in military populations. Some of the relevant risk factors for this review are: being female, having a pre-existing disorder, having a family history of anxiety and depression, past exposure to trauma, degree and nature of war zone exposure (Fairbank et al., 2001). All else equal, longer individual deployments result in higher risk of occurrence of at least one traumatic incident. Deployment length is therefore an important variable of interest. Deployment length varies by country and by type of service (Buckman et al., 2010). For example the UK Army typically deploys for 6 months, while the US Army typically deploys for a period of between 12 to 15 months. The UK Navy may face deployments of up to 22 months for every 36 month period (unless deployed on land), while the US Navy and Marines deploy for 7 months. The UK Air Force deploys for 9.3 months in every 24-month period, while in the US deployments are of 4 month duration. Within these expected deployment lengths duration may vary considerably from sometimes only a few days to well above a year (US Army, 2011). Both troop demands and the nature of operation affect the length of deployment. Generally, however, deployment lengths vary between 6 and 12 months. Street et al. (2009) argue that mental "resilience" may decrease over time, therefore time between re-deployments and number of deployments may also be important factors for predicting the experience of trauma. This is particularly relevant since, as a result of the large scale troop demands for recent operations in Iraq and Afghanistan, military personnel in general are deployed longer than before and have shorter rest-periods between deployments (Tanielian & Jaycox, 2008).

Deployment may also lead to substance abuse, which can be considered a mental disorder (American Psychiatric Association, 1994). Long term effects of deployment may be increased alcohol intake, especially linked to combat exposure, PTSD, and depression (Jacobson et al., 2008). Deployed military personnel are more likely to experience new onset heavy weekly drinking, binge drinking, and alcohol related problems, compared with non-deployed

military personnel (Hoge et al., 2004; Jacobson et al., 2008). This is particularly evident for deployed individuals with combat exposure (Fear et al., 2010; Jacobson et al., 2008). Deployed individuals diagnosed with PTSD and depression have increased odds of new onset and continued alcohol-related problems (Fear et al., 2010; Jacobson et al., 2008). A few studies suggest that drug use is an increasing concern, particularly for individuals deployed to Afghanistan and Iraq, where heroin is readily available (Karney, Ramchand, Osilla, Caldarone, & Burns, 2008; Thomsen et al., 2010).

Traumatic Brain Injury (TBI), more specifically the cognitive consequences of TBI, has received attention following the military operations in Iraq and Afghanistan. In those campaigns, insurgents have increasingly made use of IEDs (Improvised Explosive Devices) in their fight against international coalitional forces. Tanielian and Jaycox (2008) found that 19.5% of a representative sample of the US military population which had served in Iraq or Afghanistan had probable TBI.

The experience of trauma can potentially have positive consequences, termed posttraumatic growth (PTG). Tedeschi and Calhoun (1996, 2004) identify five domains where PTG may occur: "personal strength," "new possibilities," "relating to others," "appreciation of life," and "spiritual change." In their review of PTG, Ramos and Leal (2013) identify five validated self-report measures to asses PTG.

1.3.2 Impact on Social Functioning

There is little evidence for the direct link between deployment and social functioning. Nevertheless, we are interested in exploring the effects of deployment on social functioning, which may work through multiple channels. For example, social functioning may be affected both directly and indirectly by the traumas of war. Social functioning may be impacted indirectly by psychological trauma experienced during deployment. Such trauma may result in the development of mental afflictions which may, in turn, affect the social functioning of veterans. More direct channels in which deployment affects social functioning recognise that deployment to a combat zone leads to increased risk of loss of limbs, paralysis, TBI, etc. Such factors may affect the chance of gaining employment upon returning from military service abroad, maintaining marital bonds, friendships, etc. Deployment may also entail loss of firm specific human capital leading to adverse effects on salary or employment (Harvey et al., 2011).

We understand social functioning as the ability to undertake tasks and duties in civil society. Social functioning includes the ability to carry out work and home tasks, the degree to which financial concerns are experienced, relationships with family, satisfying sexual relationships, social contacts, and pleasure in spare time activities. Here we specifically focus on a subset of objective indicators of social functioning namely: employment and homelessness.

Employment is linked to deployed military personnel with PTSD and mental health problems, showing increased unemployment in deployed populations with PTSD and

depression (Karney et al., 2008; Zatzick et al., 1997). Homelessness is more prevalent in post-deployed military personnel compared to the general population (Perl, 2007). It is, however, unclear whether these indicators of reintegration into civil society are linked directly to deployment (Karney et al., 2008). Absence from the civil labour market during deployment (and military service in general) may affect employment after deployment. Studies indicate that deployed military personnel with PTSD are less likely to be employed than deployed military personnel without PTSD (Karney et al., 2008; Zatzick et al., 1997). There is little data on the rates of homelessness among individuals with PTSD and the link between deployment and homelessness (Karney et al., 2008).

1.4 PREVIOUS REVIEWS

Two previous reviews have studied the mental health effects of deployment. The review most closely related to the topic of interest in the current protocol is a systematic review of psychiatric disorders in veterans of the Persian Gulf War of 1991 (Stimpson, Thomas, Weightman, Dunstan, & Lewis, 2003). The aim of Stimpson et al. (2003) was to review all studies comparing the prevalence of psychiatric disorders in Gulf War veterans with a group of service personnel not deployed to the Gulf War. The review authors identified 20 studies in which the desired comparison was present. Although heterogeneity between studies was significant, all studies reported increased prevalence of PTSD and common mental disorder in Gulf War veterans compared with the prevalence in other active service personnel not deployed to the Gulf War. For a description of the methods used in the primary studies, see section 1.5. The present review goes beyond the Stimpson review in scope in that it goes beyond studying veterans of the Gulf War only. One advantage is that this review will be able to go beyond the specifics of the Gulf War to uncover more general patterns of the effects of deployment. For example, the Gulf War was characterised by a number of unique stressors such as usage of uranium depleted shells and the burning of Kuwaiti oil wells.

The review by Buckman et al. (2011) examines the relationship between mental health and deployment length. Their review is broader than this review in that they consider all deployed personnel including private security personnel and journalists. In addition, they examined the effect of a discrepancy between expected deployment length and actual length on health and well-being. A systematic search of studies measuring deployment length and the issue of 'mismatch' between expected and actual deployment length was conducted. Nine studies were reviewed. The review suggests that, as deployment length increases, the potential for personnel to suffer adverse health effects also increases. Our review will be broader in scope in that it is more broadly interested in measuring the effects on mental health of deployment. Deployment length may be an important factor in this respect in that it may moderate the effects of deployment.

1.5 METHODS USED IN PRIMARY STUDIES

Studies of the effects of deployment on military personnel to international military missions with respect to mental health and social functioning typically rely on observational data. The most common method is to rely on data obtained from administrative registers or from surveys. Stimpson et al. (2003) reviewed 20 studies that were all based on cross sectional surveys. The studies reviewed in Stimpson et al. selected participants from the military databases using a stratified sampling procedure and/or were frequency-matched on characteristics such as age, gender reserve/regular, etc. in order to adjust for confounding (Cherry et al., 2001; Iowa Persian Gulf Study Group, 1997; Kang, Mahan, Lee, Magee, & Murphy, 2000; Pierce, 1997; Proctor et al., 1998; Steele, 2000; Unwin et al., 1999; Wolfe et al., 1999). Some studies further adjusted for confounding factors; two studies explicitly stated they used logistic regression (Proctor et al., 1998; Unwin et al., 1999) whereas three studies presented adjusted odds ratios without specifying the adjustment method (Goss Gilroy Inc., 1998; Iowa Persian Gulf Study Group, 1997; Steele, 2000). One study used only logistic regression (Holmes, Tariot, & Cox, 1998), another study used multiple regression techniques to adjust for confounders (Stuart & Bliese, 1998) and finally one study used a hierarchical multiple regression model (Stretch, Bliese, Marlowe, & Wright, 1996a). Seven studies used neither a stratified sampling procedure/frequency matching nor adjusted for confounding factors in other ways (Bartone, 2000; Gray, Kaiser, Hawksworth, Hall, & Barett-Connor, 1999; Perconte, Wilson, Pontius, &Dietrick, 1993; Stretch, Marlowe, Wright, & Bliese, 1996b; Stuart & Halverson, 1997; Sutker, Uddo, Brailey, & Allain, 1993; Sutker, Uddo, Brailey, Vasterling, & Errera, 1994).

More recently, a number of studies have used large population-based military cohort data: the Millennium Cohort Study, 2001-2008, is a longitudinal database consisting of participants drawn from a randomly selected sample of all US military service members on rosters as of October 2000 (Panel 1), October 2003 (Panel 2), and October 2006 (Panel 3). Health status is assessed at baseline using self-reported questionnaire data linked to supplemental data from various Department of Defence administrative and health databases. Follow-up data will be collected every 3 years through the year 2022 with periodic queries of the same electronic databases.[7]

The effects of deployment, measured based on the Millennium Cohort Study, on a number of different outcomes are calculated using multivariate logistic regression adjusting for a number of covariates; see for example, Jacobson et al. (2008), Smith et al. (2007), Wells et al. (2010) and Seelig et al. (2010). The study by Sareen et al. (2007) uses multivariate logistic regression adjusting for a number of covariates to study the effects of deployment among Canadian military personnel. Their data are, however, not longitudinal. They apply cross-sectional population-based survey in a multistage sampling framework to ensure representativeness in relation to the Canadian military population.

---

[7] For more information on this database see Smith et al. (2011) and Ryan et al. (2007).

1.6 WHY IT IS IMPORTANT TO DO THIS REVIEW

Deployment to military missions affects many people across the globe with the increased international military activity. In 2011, the missions to Afghanistan involved approximately 132,000 deployed from 48 nations, and in Iraq 49,700 deployed (USF-IRAQ, 2011). To the best of our knowledge the evidence post-dating the Gulf War has not been systematically reviewed. We seek to synthesize the existing research and systematically organise knowledge about important consequences of deployment. In addition, we want to explore in a comprehensive way how important consequences of deployment co-vary with policy instruments such as deployment length. Furthermore, we seek to identify gaps in existing knowledge in order to enhance future options for prioritizing research in the field.

This review is also conducted in order to inform potential next steps in policy development in the area of deployment and post-deployment support. By identifying the major effects of deployment, the review can inform policy development on deployment and military activity as well as post deployment support for veterans.

2 Objective of the Review

The objective of this review is to synthesise relevant studies in the research literature on the effects of deployment of military personnel to international military operations after 1989 with respect to the impact of deployment on mental health and social functioning.

3 Methods

3.1 CRITERIA FOR CONSIDERING STUDIES FOR THIS REVIEW

3.1.1 Types of Studies

We expect the majority of studies to be based on observational data; it is difficult to imagine that a military decision-maker would be willing to randomise between deploying and not deploying individual military personnel to an operation. In other words, even though assignment to military service may have derived from a draft lottery (such as e.g. the US Vietnam Draft Lottery), assignment to the deployment condition is likely not random. The decision procedures of military commanders typically involve a trade-off between the benefits of deploying experienced and effective personnel, often based on group cohesion, to the cost of deployment on physical and mental stress (Wessely, 2006). In other words, individual assignment to deployment, even if that individual is eligible as per contract, is not random. Naturally, randomised studies would be eligible for this review should our search strategy uncover any such studies. It is more likely that there are studies where assignment locally produces high quality quasi-experiments. Such studies may be available if the authors of primary studies have direct access to the decision rules of the decision maker responsible for assignment. For example, individual deployment may be contingent on reaching a particular fitness score.[8] If this is the case, and authors have access to this score, then it is possible to construct a valid comparison group by comparing individuals just above a pre-specified cut-off with those just below the cut-off. In cases where the authors of primary studies do not have access to such information, or procedures do not allow authors to identify the relevant sub-population, identification must rely on other strategies for constructing a credible comparison group. For example, Angrist, Chen, and Frandsen (2010) use the Vietnam Draft Lottery as an instrumental variable to identify the causal effect of exposure to the draft on labor force participation in the 1990s. Each such study will be judged on how well it addresses the identification problem based on the risk of bias model outlined below in Section 3.3.3.

The study designs that are eligible for this review are:

- *Controlled trials:* Randomised controlled trials, quasi-randomised trials (where participants are allocated by non-random means such as alternate allocation, birth

---

[8] For example, Hotopf et al. (2006) use a variable "fitness to deploy" to adjust for confounding.

date, day of the week, case number or alphabetical order), non-randomised trials (where participants are allocated by other actions controlled by the researcher).

- *Non-randomised studies:* the decision to deploy or not deploy individuals is not under the control of the researcher. The study must use a credible comparison group to be eligible.

Military populations can be expected to differ from the general population. This is particularly relevant to this review, since it affects what constitutes an appropriate comparison group for deployed military personnel. Military personnel are on average in better mental and physical health than the civilian population. This is due to a selection mechanism: entrance into military service requires passing a medical screening. This screening device sorts out applicants who are not fit to serve in the military for physical or psychological reasons (Haley, 1998). Two factors contribute to this difference. First, individuals who select themselves into a military career are not a random sample of the general population. Second, military training of recruits acts as a selection device that tends to select those who are physically and mentally strong. Individuals who are unable to cope with the stress of military training leave active duty or are discharged and return to civilian life. In addition, deployed military personnel have undergone additional physical and (sometimes) mental screening prior to being selected for deployment (Warner, Appenzeller, Parker, Warner, & Hoge, 2011). As a result, it can be expected that even within the subpopulation of military personnel, deployed and non-deployed will differ in both observed and un-observed characteristics. This effect is well known in occupational health studies, and is known as the "healthy worker effect" (see, e.g., Li & Sung, 1999). We return to this point in Section 3.3.

We will only include studies that use a well-defined comparison group. Such studies, for example, may

- compare deployed military personnel with non-deployed military personnel serving the same nation in the same era;

- compare military personnel deployed to high combat intensity missions with military personnel deployed to low combat intensity or non-combat deployments serving the same nation;

- compare deployed personnel to non-deployed personnel; or

- compare two groups of deployed personnel deployed under different combat exposure intensities.

While on the one hand we can expect that military decision makers select whom to deploy and whom not to deploy, there is also evidence to suggest that the deployment of individual servicemen to either combat vs. non-combat is unrelated to individual characteristics (e.g., Cesur, Sabia, & Tekin, 2012). We argue that a distinction must be made between non-

randomised studies that simply seek to document correlation between observable characteristics of participants and outcome(s), and non-randomised studies that attempt to mimic an experimental situation by constructing and documenting a plausible comparison group. Only the latter type of study design is eligible for this review. We now elaborate on the characteristics of plausible comparison conditions. As stated earlier, the military population differs in important ways from the civil population, not least because servicemen are frequently screened mentally and physically. Therefore, a credible comparison group to a deployed group must also be sampled from the military population. Even within the military population, selection procedures, including physical and mental screening, determine who does and who does not get deployed. Therefore restricting the comparison group to be sampled from the military population is not sufficient to remove selection bias, but goes some way toward resolving this problem. In addition, a comparison sample must be balanced on important confounders viz-a-viz the deployed sample, such that at least in observable characteristics, the deployed and the comparison group differ only in the assignment to having been deployed and not having been deployed. Some important confounders are: gender, age, rank, type of service, health, etc. We elaborate on the set of confounders in Section 3.3.3. An example of a primary study that is eligible for this review is Unwin et al. (1999). They assess mental health outcomes of UK servicemen who deployed to the Gulf War using postal survey. As a comparison group, they survey UK servicemen who were serving during the Gulf War but did not deploy there. Another study that would be eligible for inclusion is Cesur et al. (2012). They compare personnel that were deployed to combat zones during recent military operations in Afghanistan and Iraq with non-combat deployed personnel. They argue that in their sample the assignment of personnel to either combat zone or non-combat zone is orthogonal to individual characteristics. Their estimates therefore have a (plausible) causal interpretation.

It is also possible that mandated government policies assist researchers in creating a reasonable comparison group. For example, servicemen who would otherwise be eligible for deployment may not have been deployed at the time of the study because they were not *yet* eligible for re-deployment. If the deployed sample and the comparison sample are balanced with regard to number of previous deployments then this may be a valid comparison group. An example of a study that appears to have this type of information available to the researchers is Hotopf et al. (2006).

Another type of study that would be eligible for inclusion are studies that compare two (or more) different deployed samples that have (on average) been exposed to different degrees of combat exposure. If deployment mainly affects mental health and other outcomes through an increased risk of trauma, then a larger degree of combat exposure increases the likelihood of trauma, all else equal. In other words, it may be reasonable to argue that the two different samples have received different *dosages* of deployment. Also here caution must be exercised since commanding officers may use selection procedures to assign individuals to more or less severe combat exposures based on characteristics that are unobserved to researchers.

Studies that contain only a sample of deployed and compare outcomes before and after deployment are not eligible.

3.1.2   Types of Participants

The populations that are eligible for this review are military personnel, from any nation, who have faced deployment to international military operations since 1989. As detailed in section 1.1 there are several reasons for this choice of period. First, 1989 marks the end of the Cold War period and as a consequence the types of military engagements have changed considerably since then. Second, modern military engagement relies heavily on activating reservist components to fill manpower demands. Because active duty and reservists components tend to be different in observable (and probably unobservable) ways deployment can be expected to have different effects on modern armies. Third, as a consequence of technological developments in body armour more military personnel now potentially become survivors of military confrontations, which would have led to death in earlier conflicts. As a result the composition of physical and psychological trauma in military populations have changed since the end of the Cold War.

Studies of deployment to military bases (abroad) not involved in active military operations will be excluded. Studies that focus on the deployment of civilian personnel to peace-keeping or war zone operations will be excluded.

All types of service members are eligible. Specifically, active duty, reserve, and national guards are eligible.

All types of military personnel are eligible. Specifically, personnel from army, air force, navy, marine corps, special forces, are all eligible for inclusion in this review.

Military personnel regardless of age, gender, ethnic background, marital status, education, military rank, and country will be included.

3.1.3   Types of Interventions

The "intervention" is deployment to an international military operation.

3.1.4   Types of Outcomes

This review will include studies that focus on outcomes for the *individual who has been deployed*. Studies that solely focus on the consequences for the spouse and/or children will not be included.

3.1.4.1 Mental Health

Of interest to this review is the effect of deployment on mental health outcomes. The mental health outcomes that are included in this review are:

- (Probable) posttraumatic stress disorder (PTSD)
- (Probable) major depression disorder (MDD)
- (Probable) common mental disorder
- (Probable) substance-related disorder

We write probable since we do not expect that the majority of studies measure mental health outcomes via structured clinical interviews, which are considered the "gold" standard in this literature. Instead we expect that questionnaires are used to screen for probable mental disorders. Engelhard et al. (2007) and Kehle et al. (2011) are examples of studies where PTSD outcome is measured following structured clinical interviews. Some studies may indirectly infer clinical diagnosis from register data. Also, in this case, the outcome will only be registered for those who were exposed to the diagnostic or sought medical assistance on their own. Such measures may therefore also be susceptible to bias.

The most common form of detection is via self-reported questionnaires for the study population. The use of different instruments of detection may therefore be an important source of variation for the incidence of measured mental health outcomes. Smith et al. (2008) uses self-reported symptoms measured by DSM-IV criteria using a 17-item PTSD checklist, PCL-C. The PTSD Checklist (PCL) (sensitivity: 1.0, specificity: .92) was used by Hoge et al. (2004) to detect (probable) PTSD. They also used a stricter version of the same questionnaire (sensitivity: .60, specificity: .99). Hoge et al. used the 9-item Patient Health Questionnaire (PHQ-9) to detect (probable) MDD. A commonly used instrument for detecting alcohol abuse is the 10-item WHO Alcohol Use Disorder Identification Test (e.g., Harvey et al., 2011).

3.1.4.2 Social Functioning

The aim of including social functioning outcomes is to provide indications of the consequences of deployment for the experience of returning to civilian life. Social functioning is a multi-dimensional concept and covers perceived social support from military and family/friends, social participation, sexual functioning, civilian work adjustment, etc. Some of these measures are clearly more relevant for reservists and guards. For example, reservists will often leave civil employment when activated for duty. The time away and increased risk of trauma may lead to loss of specific and general human capital, and make it difficult to return to the civil career track they were on prior to deployment. Other measures, for example, family functioning, are expected to affect all deployed, due to the separation from regular life that deployment entails.

This review will focus on the effect of deployment on employment and homelessness.

Due to the multi-faceted nature of the concept of social functioning we expect primary studies to use a range of outcome variables to measure aspects of social functioning. Some may come from self-reported questionnaires (either dichotomous, multi-scale or index), authorities, files and registry data.

3.1.4.3 Time of Measurement

Another important factor is the temporal aspect of particularly the measurement of mental health outcomes. The time at which surveys are administered is important; some adverse conditions take time to develop, while others may have already been treated if surveys are administered late (Bliese, Wright, Adler, Thomas, & Hoge, 2007; Castro & Hoge, 2005; Hotopf et al., 2006). Hence, studies may reveal a specific country's battery of treatments and their effectiveness, rather than the incidence of mental affliction. We shall therefore record the time at which outcomes are measured relative to the end of a deployment spell. We return to this point in Section 3.3.5.

3.2 SEARCH METHODS FOR IDENTIFICATION OF STUDIES

3.2.1   Electronic Searches

Relevant studies will be identified through electronic searches of bibliographic databases, government policy databanks and internet search engines. No language or date restrictions will be applied to the searches.

The following bibliographic databases will be searched:

Medline
Embase
PsycINFO
Socindex
ERIC
Academic Search Elite
Cochrane
SweMed+

3.2.2   Search Terms

An example of the search strategy for MEDLINE searched through the Ovid platform is listed below. This strategy will be modified for the different databases. We will report details of the modifications used for other databases in the completed review.

1   Military Personnel/
2   Veterans/
3   (Soldier* or Marine* or airm#n* or Veteran* or militarian* or officer*).ti,ab.
4   (Army or troop* or militar* or navy or ((armed or naval or air or defen#e or national) adj1 Force*)).ti,ab.
5   or/1-4
6   exp War/
7   (war or wars or warfare*).tw.
8   (combat* or deploy* or post-deploy* or postdeploy*).tw.

9   ((peace-keep* or peacekeep* or peace keep* or military) adj1 mission*).tw.

10   (active adj1 duty).tw.

11   or/6-10

12   5 and 11

13   Randomized Controlled Trial/ or random allocation/

14   (randomized controlled trial or controlled clinical trial).pt.

15   (randomized or placebo or randomly).ab.

16   (trial or groups).ab.

17   13 or 14 or 15 or 16

18   humans.sh.

19   17 and 18

20   exp case control studies/ or exp cohort studies/

21   Comparative Study.pt.

22   Cross-sectional studies/ or Epidemiolog# studies/

23   (case control or cohort or cross sectional or cross-sectional or point in time or point-in-time).tw.

24   ((epidemiolog# adj2 study) or ((follow up or followup or follow-up) adj2 study)).tw.

25   (longitudinal or observational or (prospective adj2 study) or retrospective).tw.

26   Intervention Studies/

27   (quasi-experiment* or quasiexperiment* or experiment*).tw.

28   (assign* adj5 (subject* or patient* or client*)).ab,ti.

29   (Propensity score* or heckman or difference in difference or regression discontinuity
          (compar* adj1 group*) or (match* adj1 control*) or (match* adj1 compar*) or
          assessment only or comparison samp* or propensity-matched).ab,ti.

30   (Non-random* or nonradom* or (non adj1 random*)).ab,ti.

31   rct.tw.

32   or/20-31

33   12 and 32

34   limit 33 to humans

3.2.3   Searching Other Resources

The review authors will check reference lists in relevant reviews and included primary
studies for new leads. We will contact international experts to identify unpublished and on-
going studies, and provide them with the inclusion criteria for the review along with the list
of included studies, asking for any other published, unpublished or ongoing studies relevant
for the review.

3.2.4   Grey Literature

OpenGrey will also be used to search for European grey literature (http://opengrey.eu/). Australian Centre for Posttraumatic Mental Health will also be used to search for Australian grey literature (www.acpmh.unimelb.edu.au/). Likewise we will use Rand for US grey literature (www.rand.org) and US department of defense (www.dod.gov) and Walter Reed Army Institute of research (http://wrair-www.army.mil/) for relevant US military grey literature. DTIC (http://www.dtic.mil/dtic/) will be searched as well.

Copies of relevant documents will be made recording the exact URL and date of access.

3.3 DATA COLLECTION AND ANALYSIS

3.3.1   Selection of Studies

Under the supervision of MB, two members of the review team will independently screen titles and abstracts and exclude studies that are clearly irrelevant. Studies considered relevant by at least one of the reviewers will be retrieved in full text. The full texts will then be appraised by two members of the review team. The same persons will decide whether the studies meet the inclusion criteria. Any disagreements about eligibility will be resolved by consulting a third reviewer. Reasons for exclusion will be documented for each study that is retrieved in full text. The study inclusion coding sheet will be piloted and adjusted if required by the review authors (see Appendix 6.1).

The inclusion coding questions for both the title and abstract screening for relevance and the full-text screening for eligibility will be piloted and adjusted if required (see appendix 7.1) and changes will be reported. If necessary, primary investigators will be contacted to clarify study eligibility. In the event of disagreements, a third reviewer and content specialist will be consulted to obtain a consensus. Disagreements resolved by a third reviewer will be reported. Exclusion reasons for studies that otherwise might be expected to be eligible will be documented and presented in an appendix. The overall search and screening process will be illustrated in a flow-diagram. None of the review authors will be blind to the authors, institutions, or the journals responsible for the publication of the articles.

3.3.2   Data Extraction and Management

At least two review authors will independently code and extract data from the included studies. Data and descriptive information will be extracted on: characteristics of participants (gender, age, ethnicity, rank nationality, etc.), intervention characteristics (mission location, type, command, combat exposure, duration, etc.), control conditions (non-deployed, deployed to non-combat or low intensity mission), research design, confounding factors, sample size, risk of bias, outcomes and results. Extracted data will be stored electronically. A data extraction sheet will be piloted on several studies and revised as necessary (see Appendix 6.2). Disagreements will be resolved by consulting a third review author with

extensive content and methods expertise. Disagreements resolved by a third reviewer will be reported. Analysis will be conducted in RevMan5, SAS and Stata.

3.3.3   Assessment of Risk of Bias in Included Studies

We will assess the methodological quality of studies using a risk of bias model developed by Prof. Barnaby Reeves in association with the Cochrane Non-Randomised Studies Methods Group.[9] This model is an extension of the Cochrane Collaboration's risk of bias tool and covers risk of bias in non-randomised studies that have a well-defined control group.

The extended model is organised and follows the same steps as the existing risk of bias model according to the Cochrane Handbook, chapter 8 (Higgins and Green, 2008). The extension to the model is explained in the three following points:

1) The extended model specifically incorporates a formalised and structured approach for the assessment of selection bias in non-randomised studies by adding an explicit item about confounding. This is based on a list of confounders considered to be important and defined in the protocol for the review. The assessment of confounding is made using a worksheet where, for each confounder, it is marked whether the confounder was considered by the researchers, the precision with which it was measured, the imbalance between groups, and the care with which adjustment was carried out (see appendix 6.3). This assessment will inform the final risk of bias score for confounding.

2) Another feature of non-randomised studies that make them at high risk of bias is that they need not have a protocol in advance of starting the recruitment process. The item concerning selective reporting therefore also requires assessment of the extent to which analyses (and potentially, other choices) could have been manipulated to bias the findings reported, e.g., choice of method of model fitting, potential confounders considered / included. In addition, the model includes two separate yes/no items asking reviewers whether they think the researchers had a pre-specified protocol and analysis plan.

3) Finally, the risk of bias assessment is refined, making it possible to discriminate between studies with varying degrees of risk. This refinement is achieved with the addition of a 5-point scale for certain items (see the following section, *Risk of bias judgment items* for details).

The refined assessment is pertinent when thinking of data synthesis as it operationalizes the identification of studies (especially in relation to non-randomised studies) with a very high risk of bias. The refinement increases transparency in assessment judgments.

---

[9] This risk of bias model was introduced by Prof. Reeves at a workshop on risk of bias in non-randomised studies at SFI Campbell, February 2011. The model is a further development of work carried out in the Cochrane Non-Randomised Studies Method Group (NRSMG).

3.3.3.1 Risk of Bias Judgment Items

The risk of bias model used in this review is based on nine items (see appendix 6.3). The 9 items refer to: sequence generation, allocation concealment, confounders, blinding, incomplete outcome data, selective outcome reporting, other potential threats to validity, a priori protocol, and a priori analysis plan.

3.3.3.2 Confounding

An important part of the risk of bias assessment of non-randomised studies is how the studies deal with confounding factors. Selection bias is understood as systematic baseline differences between groups and can therefore compromise comparability between groups. Baseline differences can be observable (e.g. age and gender) and unobservable (to the researcher; e.g. underlying "mental resilience"). There is no single non-randomised study design that always deals adequately with the selection problem. Different designs represent different approaches to dealing with selection problems under different assumptions and require different types of data. It especially varies how different designs deal with selection on un-observables. The "adequate" method to control for selection depends on the model generating participation, i.e., assumptions about the nature of the process by which participants are selected into a program.

For this review, we have identified the following observable confounding factors to be most relevant: mental health history, gender, age, ethnicity, military rank (enlisted, officer), branch of service (Army, Navy, Air Force, etc.), Duty/Enlistment status (Active, Reserve/Guard), and number of previous deployments. All confounding variables should be measured at their pre-deployment values.

In addition, the presence of a "healthy worker" effect (Hotopf et al., 2006; Larson et al., 2008; Li & Sung, 1999) may confound results. The military have high induction standards followed by rigorous training to test physical and mental health. While waiting for deployment, serving troops are continually observed (and sorted). Reservists and National Guards are subject to less rigorous screening. Troops can be discharged at several decision points if physical or mental health problems are noticed (and would therefore not be deployed; Hyams, 2006). As a result, the military population will be healthier than the general population, and in addition, we expect that the deployed may differ in systematic ways from the non-deployed, although we do not expect this difference to be as stark as the difference between the military population and the general population.

Review authors (at least two) will independently assess the risk of bias for each included study. Disagreements will be examined by a third reviewer with content and statistical expertise. We will report the risk of bias assessment in risk of bias tables for each included study in the completed review. We now discuss the rationale for including each of the confounding variables.

Mental Health History

Several studies indicate that prior PTSD and other mental disorders may help predict the onset of new disorders (Himmelfarb, Yaeger, & Mintz, 2006; Vogt, King, & King, 2007). It is unclear whether a prior history of mental health problems increases the risk of e.g. PTSD, or simply reveals a latent variable ("mental resilience").

Gender

In general, male and female service members are likely to experience different types combat exposure. For example, US servicewomen who deployed to Iraq and Afghanistan were officially barred from serving in direct combat situations. But, those wars were essentially without a frontline; due to the nature of fighting, even support personnel were at an increased risk of combat exposure. Nevertheless, the nature of exposure may be different due to different assignments. In addition, female service members, as compared to males, are likely to report a higher incidence of sexual trauma, including sexual assault and harassment while deployed (Murdoch et al., 2006). Tanelian and Jaycox (2008) found that deployed female service members were more likely than their male counterparts to screen for probable PTSD. Given that the military is a male-dominated occupation, there may also be differential selection procedures by gender into a military career. Finally, there is the possibility that "true" gender differences exist in the way that males and females process trauma exposure, leading to differential outcomes in the link between exposure and, for example, PTSD, although evidence is mixed (Street et al., 2009).

Age

Age has been shown to be a risk factor in some studies for the development of PTSD (Brewin, Andrews, & Valentine, 2000). In addition, age may also serve as proxy for the size social support networks.

Ethnicity

Ethnicity may partially determine how, and if, trauma is verbalised and processed. Ethnicity may also be a proxy for different networks of social support both prior to and after deployment. Tanelian and Jaycox (2008, p. 99) find that Hispanic ethnicity almost doubles the risk ratio (RR) for PTSD (adjusted RR of 1.881) in a representative sample of US servicemen deployed to Iraq.

Branch of Service

Service branch is mainly a proxy for the exposure to trauma. A very basic difference between the typical airman and the typical combat soldier is the distance to wounded combatants, smell of decaying bodies, etc. which are all known risk factors for trauma. In addition, pre-deployment preparation may also vary depending on branch of service.

Military Rank

Officers are a (self-)selected group of military personnel and can be expected to be different from enlisted personnel in the amount of training received. More importantly, officers and enlisted personnel will likely differ in the ways in which they are exposed to potentially traumatic experiences in-theatre.

Duty/Enlistment Status

Deployed Military personnel may come from three different military populations: active duty personnel, reservists, and national guard. Each of these populations differs in the frequency with which they are observed and screened (physically and mentally) by officers. For example, active duty service members who live on-base will be screened much more frequently than reservists or guards. Thus, comparisons between these sub-populations may be susceptible to the "healthy worker" effect (Li & Sung, 1999). Hotopf et al. (2006) is one primary study that found different effects for active duty personnel and reservists.

Previous Deployments

Previous deployments may be important as some studies suggest that mental resilience can be compared to a form of mental capital that depreciates and is expended under stressful situations (Atkinson, Guetz, & Wein, 2009; Street et al., 2009). When mental capital is depleted, individuals are more likely to suffer adverse consequences to mental health. Previous deployments may serve as a proxy for the amount of mental capital expended.

3.3.4   Measures of the Effect of Deployment

Continuous Data

For continuous outcomes, effect sizes with 95 % confidence intervals will be calculated if means and standard deviations are available. If means and standard deviations are not available, the review authors will request this information from principle investigators. If no information is yielded, we will use methods by Lipsey and Wilson (2001) to calculate standardized mean difference effect sizes from, e.g., F-ratios, t-values, chi-squared values and correlation coefficients. Hedges' small sample correction will be employed (i.e., Hedges' *g*). If there is a mix of studies with some reporting change scores and others reporting final values, we will contact authors and request the final values. If we do not obtain these values, we will analyse change scores and final values separately (Higgins & Green, 2008, section 9.4.5.2). Any scales related to combat social functioning, PTSD, MDD, alcohol and substance abuse, etc. are examples of relevant continuous outcomes in this review.

Discrete Data

For dichotomous outcomes, we will calculate odds ratios with 95% confidence intervals. Employment and marital status are examples of relevant dichotomous outcomes for this review.

There are statistical approaches available to re-express dichotomous and continuous data such that they can be pooled (Sánchez-Meca, Marín-Martínez, & Chacón-Moscoso, 2003). In order to calculate common metrics, odds ratios will be converted to SMD effect sizes using the Cox transformation. We will only transform dichotomous effect sizes to SMD if appropriate, e.g., as may be the case with PTSD which may be measured either as a

dichotomous outcome or on a continuous scale. We will only transform the data if a large majority of studies report continuous effect sizes, otherwise we will perform separate analysis for each measure.

When effect sizes cannot be pooled, study-level effects will be reported in as much detail as possible. Software for storing data and statistical analyses will be RevMan 5.0, Excel and Stata 12.

3.3.5   Unit of Analysis Issues

Multiple intervention and control groups

Multiple intervention groups within a study with one control group will be pooled if appropriate (if they include different individuals) and compared to that control group. Multiple controls groups will only be pooled if appropriate (if they include different individuals).

A synthetic (average) effect size will be calculated and used to avoid dependence problems. This method provides an unbiased estimate of the mean effect size parameter but overestimates the standard error. Random effects models applied when synthetic effect sizes are involved actually perform better in terms of standard errors than do fixed effects models (Hedges, 2007). However, tests of heterogeneity when synthetic effect sizes are included are rejected less often than the nominal significance level.

If pooling is not appropriate (e.g., the multiple interventions and/or control groups include the same individuals), only one intervention group will be coded and compared to the control group to avoid overlapping samples. The choice of which estimate to include will be based on our risk of bias assessment and upon theoretical relevance. We will choose the estimate that we judge to have the least risk of bias (primarily, selection bias and in case of equal scoring the incomplete data item will be used) and that is most likely to provide the best information about the outcome being assessed.

Multiple studies using the same sample of data

In some cases, several studies may have used the same sample of data. We will review all such studies, but in the meta-analysis we will only include one estimate of the effect from each sample of data. This will be done to avoid dependencies between the "observations" (i.e. the estimates of the effect) in the meta-analysis. The choice of which estimate to include will be based on our risk of bias assessment of the studies. We will choose the estimate from the study that we judge to have the least risk of bias (primarily selection bias), the number of observations, and precision of estimates.

We consider the risk of bias measure to be an important instrument when selecting among studies using the same primary data. As an example, suppose two different studies, A and B, measure the same outcomes in the same population. Both studies control for confounders

but only study A controls for an important confounder, say type of service. We believe that study A should take precedence because that study is much more likely to provide an effect estimate closer to the "truth". If studies score identically on risk of bias we will also consider number of observations and precision of estimate as important instruments when choosing which study to include.

Multiple time points

If outcomes are measured at multiple time points, each time specific outcome will be analysed first in a separate meta-analysis with other comparable studies taking measures at a reasonably similar time point. Secondly, we will explore the extent to which the multiple outcomes can be included in a meta-analysis using time of measurement as a moderating variable.

We remain flexible as to which classifications to consider. We will record the time of measurement from the end of deployment. As a guideline we will consider: short-term (0-<6 months after deployment), medium term (6- 24 months after deployment) long term (at least 24 months after deployment). We will report any changes to this classification.

3.3.6 Dealing with missing data and incomplete data

Missing data and attrition rates will be assessed in the included studies. Notably due to the nature of the field, studies of deployment typically estimate the effect on data collected from administrative registers and, even more frequently, by questionnaires. Studies using data collected from questionnaires are especially subject to missing data (non-response). For studies using questionnaire data, a sensitivity analysis will be used to assess potential bias, and the extent to which the results might be biased by missing data will be discussed.

The review authors will record attrition/response rates and (if possible) reasons for attrition from included studies.

Another potential source of bias is the procedures under which questionnaire data are collected and by whom. For example, response rates and reporting may be affected by whether questionnaires are anonymous or not. Even if collection is anonymous, but collected by, say, the Department of Defense, personnel may be worried that their responses will be used against them if filing a claim, or will affect opportunities for promotion.

3.3.7 Assessment of heterogeneity

Heterogeneity among primary outcome studies will be assessed with Chi-squared (Q) test, and the I-squared, and τ-squared statistics (Higgins, Thompson, Deeks, & Altman, 2003). Any interpretation of the Chi-squared test will be made cautiously on account of its low statistical power.

3.3.8   Assessment of reporting biases

Reporting bias refers to both publication bias and selective reporting of outcome data and results. Here, we state how we will assess publication bias.

We will use funnel plots for information about possible publication bias if we find sufficient studies (Higgins & Green, 2008). However, asymmetric funnel plots are not necessarily caused by publication bias (and publication bias does not necessarily cause asymmetry in a funnel plot). If asymmetry is present, we will consider possible reasons for this.

3.4 DATA SYNTHESIS

We will synthesize results of deployment in the included studies initially using a broad approach.

Absolute effects

Analysis of the absolute effects of deployment will involve comparing deployed military personnel to non-deployed military personnel serving the same nation in the same era.

Relative effects

The relative effects of deployment (high combat intensity mission deployment versus low combat intensity or non-combat mission deployment) will be conducted separately and will include studies that compare different types of deployment of military personnel from same nation.

Studies that have been coded with a very high risk of bias (5 on the risk of bias scale) will not be included in the data synthesis. It is important to emphasise that the risk of bias tool is an extension of the standard risk of bias tool to non-randomised studies. Included studies will be ranked relative to this scale, taking into account that the studies are non-randomised. On account of this, a score of 5 will be given cautiously, and only on the grounds that such a study is more likely to mislead than to inform. If, counter to our expectation, many studies (more than 25%) score 5, we will perform a sensitivity analysis including all studies.

We will pool included studies where appropriate, dependent on the availability of data and heterogeneity. We will pool all types of deployment and types of military assignments/jobs to an international mission, where outcome measures are adequately similar and meaningful to pool.

All follow–up durations reported in the primary studies will be recorded and we will do separate analysis using the single measure that is the most common time point (or closest to the most common time) within the following (preliminary) categories: Short term outcomes

(from end of deployment to less than 6 months), medium term outcomes (from 6 months to less than 2 years after deployment), and long term outcomes (2 years and longer after end of deployment).

Meta-analysis

Meta-analysis will be conducted when event rates or means and standard deviations are available or can be calculated and when time point and outcome measurements are adequately similar. In other cases, when these conditions are not met, results will be shown but not pooled (Higgins & Green, 2008, section 16.5.4).

Random effects meta-analysis will be used. For subsequent analyses of moderator variables in search for systematic variations we will switch to the mixed model (Hedges & Pigott, 2004; Konstantopoulos, 2006).

When meta-analysis is inappropriate, a narrative description of the study results alone will be provided, although general conclusions on the effects of deployment would not be possible in that case.

3.4.1   Moderator Analysis and Investigation of Heterogeneity

We will investigate the following (subset of) factors with the aim of explaining observed heterogeneity: Study-level summaries of participant characteristics: mental health history, gender, age, ethnicity, military rank, branch of service, and duty/enlistment status. Many of the proposed variables have a strong link to policy. For example, if age turns out to be an important factor for the incidence of PTSD, new rotational procedures can be devised which delay deployment for personnel at certain ages.

In addition, and if available in summary format, the number and duration of previous deployments, and duration of most recent deployment, will be investigated as moderating variables. These are very direct policy instruments, and it is therefore of great interest to explore these relationships, although caution must be exercised in placing a too literal interpretation on the potential moderating effects of deployment length. The longer personnel are deployed, the higher the risk that they will experience at least one trauma, all else being equal, and, as a consequence, may develop adverse mental health responses. Therefore, in a mechanical way, we would expect that longer deployment lengths are associated with higher incidence of mental health problems. It does not follow from this, however, that deployment length is suboptimal. The comparison would be accurate only if we were to compare groups of individuals deployed for the same total length of time but with one group experiencing two shorter deployments and the other experiencing a single longer deployment. If incidence of PTSD in the long deployment group is larger than the effect of total deployment, then we might conclude that deployment length is sub-optimally long. A situation where this might be the case is a situation where mental resilience builds up among the non-deployed (Atkinson et al., 2009). In this case, two shorter deployments may be

preferred to one long deployment of equal total length. Similarly, Cesur et al. (2012) show that combat exposure, not deployment length in itself, has a statistically significant impact on developing PTSD.

We will also explore how combat exposure moderates outcomes. It is well known that the degree of combat exposure increases the risk of trauma. The Combat Exposure Scale (CES) is a validated instrument of combat exposure (Keane et al., 1989). The scale includes several items related to specific combat related experiences (Fairbank et al., 2001). As an example Tanielian and Jaycox (2008) used the Combat Trauma Exposure scale (adapted from Hoge et al., 2004).

If the number of included studies is sufficient (at least 10 degrees of freedom) and given there is variation in the covariates, we will perform moderator analyses (multiple meta-regression using the mixed model) to explore how observed variables are related to heterogeneity. We will estimate the (new) residual variance component to be used in a weighted least squares analysis conditional on this variance component estimate. The residual variance component will be estimated using the method-of-moments estimator (Hartung, Knapp, & Sinha, 2008; Konstantopoulos, 2006). We will report the 95% confidence intervals for regression parameters. We will estimate the correlations between the covariates and consider the possibility of confounding. Conclusions from meta-regression analysis will be cautiously drawn and will not be based on significance tests.

Otherwise, single factor subgroup analysis will be performed. The assessment of any difference between subgroups will be based on 95% confidence intervals. No conclusions from subgroup analyses will be drawn and interpretation of relationships will be cautious, as they are based on subdivision of studies and indirect comparisons.

In general, the strength of inference regarding differences in effects among subgroups is controversial. However, making inferences about different effect sizes among subgroups on the basis of between-study differences entails a higher risk compared to inferences made on the basis of with-in study differences (Oxman & Guyatt, 1992). We will therefore use within study differences where possible.

3.4.2   Sensitivity Analysis

Sensitivity analysis will be carried out by restricting the meta-analysis to a subset of the totality of studies included in the original meta-analysis. For example if several studies are excluded due to receiving a 5 on the risk of bias scale, we will present a sensitivity analysis where those studies are included.

Sensitivity analysis will be used to evaluate whether the pooled effect sizes are robust across study design and components of methodological quality. For methodological quality, we will consider sensitivity analysis for each major component of the risk of bias checklists and restrict the analysis to studies with a low risk of bias.

Sensitivity analysis will further be used to examine the rigour of conclusions in relation to the quality of data (whether data is based on questionnaires or administrative registers). For example, if there is significant variation in how primary studies detect (probable) mental disorders, then we will perform sensitivity analysis with respect to each instrument.

4 References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC.

Andersen, H.E. (1998): "Danske FN-soldater – 2 år efter: Opfølgningsundersøgelse af DANBAT 7 & 8." [Danish UN soldiers – 2 years after: follow-up study of DANBAT 7 & 8]. *FCL Publikation* 144.

Angrist, J. D., Chen, S. H., & Frandsen, B. R. (2010). Did Vietnam veterans get sicker in the 1990s? The complicated effects of military service on self-reported health. *Journal of Public Economics, 94*(11), 824-837.

Atkinson, M.P., A. Guetz, & L.M. Wein (2009): A Dynamic Model for Posttraumatic Stress Disorder Among U.S. Troops in Operation Iraqi Freedom. *Management Science, 55*(9), 1454-1468.

Australian Department of Defence. (2011). Global Operations. Retrieved from: http://www.defence.gov.au/op/index.htm. Accessed on April 27, 2011.

Bartone, P.T. (2000). Hardiness as a resiliency factor for United States forces in the Gulf War. In Violanti, J. M., Paton, D., & Dunning, C. (Eds.), *Posttraumatic Stress Intervention: Challenges, Issues and Perspectives* (pp. 115-133). Springfield, IL: Charles C. Thomas.

Bisson, J. I. (2007). Post-traumatic stress disorder. *Occupational Medicine, 57*, 399-403.

Bliese, P. D., Wright, K. M., Adler, A. B., Thomas, J. L., & Hoge, C. W. (2007). Timing of postcombat mental health assessments. *Psychological Services, 4*(3), 141-148.

Breslau, N., Davis, G. C., Peterson, E. L., & Schultz, L. R. (2000). A second look at comorbidity in victims of trauma: The posttraumatic stress disorder-major depression connection. *Biological Psychiatry, 48*(9), 902-909.

Brewin, C. R., Andrews, B., & Valentine, J. D. (2000). Meta-analysis of risk factors for posttraumatic stress disorder in trauma-exposed adults. *Journal of Consulting and Clinical Psychology, 68*(5), 748-766.

Buckman, J.E.J, Sundin, J., Greene, T., Fear, N.T., Dandeker, C., Greenberg, N., & Wessely, S. (2011). The impact of deployment length on the health and well-being of military personnel: A systematic review of the literature. *Occupational Environment Medicine, 68,* 69-76.

Castro, C. A., & Hoge, C. W. (2005, August). *Psychological impact of modern warfare: The WRAIR Land Combat Study.* Symposium presented at the American Psychological Association Annual Convention, Washington DC (pp. 18-21).

Cesur, R., Sabia, J. J., & Tekin, E. (2012). The psychological costs of war: military combat and mental health. *Journal of Health Economics, 32*(1), 51-65.

Cherry, N., Creed, F., Silman, A., Dunn, G., Baxter, D., Smedley, J., ... & Macfarlane, G. J. (2001). Health and exposures of United Kingdom Gulf war veterans. Part II: The relation of health to exposure. *Occupational and Environmental Medicine, 58*(5), 299-306.

Christensen, A.K. (2001). SFOR, hold 4, 5, 6 og 7: Resultater fra opfølgningsundersøgelsen. [Results from follow-up study of SFOR 4,5, 6 & 7]. *IMP* Småtryk 31.

Danish Armed Forces. (2011). Internationale operationer. Retrieved from http://forsvaret.dk/FPT/Nyt%20og%20Presse/tal/intops/Pages/default.aspx. Accessed on April 27, 2011.

Danish Ministry of Defence [Forsvarsministeriet]. (2010). Redegørelse for veteranforhold [Briefing on veterans affairs]. Note

Davidson, A. C., & Mellor, D. J. (2001). The adjustment of children of Australian Vietnam veterans: Is there evidence for the transgenerational transmission of the effects of war-related trauma? *Australian and New Zealand Journal of Psychiatry, 35*(3), 345-351.

Dekel, R., & Goldblatt, H. (2008). Is there intergenerational transmission of trauma? The case of combat veterans' children. *American Journal of Orthopsychiatry, 78*, 281–289.

Department of Defense. (2008). Managing the Reserve Components as an Operational Force. Washington DC: Office of the Assistant Secretary of Defense for Reserve Affairs.

Department of Defense. (2011). Retrieved from http://policy.defense.gov/solic/index.aspx.

Dohrenwend, B. P., Turner, J. B., Turse, N. A., Adams, B. G., Koenen, K. C., & Marshall, R. (2006). The psychological risks of Vietnam for U.S. veterans: A revisit with new data and methods. *Science, 313*, 979-982.

Engelhard, I.M., van der Hout, M. A., Weerts, J. A., Arntz, J. J., Hox, C. M., & McNally, R. J. (2007). Deployment-related stress and trauma in Dutch soldiers returning from Iraq. *British Journal of Psychiatry, 191*, 140-145.

Fairbank, J. A., Ebert, L., & Caddell, J. M. (2001). Posttraumatic Stress Disorder. In Sutker, P. B. & Adams, H. E. (Eds.), *Comprehensive Handbook of Psychopathology* (3rd ed.) (pp. 183-209). New York: Kluwer Academic/Plenum Publishers.

Fear, N. T., Jones, M., Murphy, D., Hull, L., Iversen, A. C., Coker, B., Machell, L., Sundin, J., ..., & Wessely, S. (2010). What are the consequences of deployment to Iraq and Afghanistan on mental health of the UK armed forces? A cohort study. *The Lancet, 375*, 1783-1797.

Fontana, A., & Rosenheck, R. (1994). Posttraumatic stress disorder among Vietnam theatre veterans: A causal model of the etiology in a community sample. *Jounal of Nervous and Mental Disease, 182*, 677-84.

Friedman, M. J. (2006). Posttraumatic Stress Disorder Among Military Returnees From Afghanistan and Iraq. *American Journal of Psychiatry, 163*(4), 586-593.

Global Security (2011). Fundamentals of Low Intensity Conflict. Retrieved from: http://www.globalsecurity.org/military/library/policy/army/fm/100-20/10020ch1.htm#s_9.

Goss Gilroy Inc. (1998). *Health Study of Canadian Forces Personnel Involved in the 1991 Conflict in the Persian Gulf (vol.1).* Ottawa: Goss Gilroy Inc. Retrieved from http://www.dnd.ca/menu/press/Reports/Health/health _study_ e_ vol1_TOC.htm

Gray, G. C., Kaiser, K. S., Hawksworth, A. W., Hall, F. W., & Barrett-Connor, E. (1999). Increased postwar symptoms and psychological morbidity among US Navy Gulf War veterans. *American Journal of Tropical Medicine and Hygiene, 60*(5), 758-766.

Haley, R. W. (1998). Point: bias from the" healthy-warrior effect" and unequal follow-up in three government studies of health effects of the Gulf War. *American journal of epidemiology, 148*(4), 315-323.

Hartung, J., Knapp, G., & Sinha, B. K. (2008). *Statistical meta-analysis with applications.* New York: Wiley.

Harvey, S. B., Hatch, S. L., Jones, M., Hull, L., Jones, N., Greenberg, N., ... & Wessely, S. (2011). Coming Home: Social Functioning and the Mental Health of UK Reservists on Return from Deployment to Iraq or Afghanistan. *Annals of Epidemiology, 21*, 666-672.

Hedges, L. W., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods, 9*(4), 426–445.

Helzer, J. E, Robins, L. N., & McEvoy, L. (1987). Posttraumatic stress disorder in the general population: findings of the epidemiologic catchment area survey. *New England Journal of Medicine, 317*, 1630-1634.

Higgins J. P. T., & Green S. (Eds.). (2008). *Cochrane Handbook for Systematic Reviews of Interventions.* Wiley-Blackwell.

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327*(7414), 557-60.

Himmelfarb, N., Yaeger, D., & Mintz, J. (2006). Posttraumatic stress disorder in female veterans with military and civilian sexual trauma. *Journal of Traumatic Stress, 19*(6), 837-846.

Holmes, D. T., Tariot, P. N., & Cox, C. (1998). Preliminary evidence of psychological distress among reservists in the Persian Gulf War. *Journal of Nervous and Mental Disease, 186*, 166-173.

Hoge, C. W., Castro, C. A., Messer, S. C., McGurk, D., Cotting, D. I., & Koffman, R. L. (2004). Combat duty in Iraq and Afghanistan, mental health problems, and barriers to care. *The New England Journal of Medicine, 351*(1), 13-22.

Hoge, C. W., Auchterlonie, J. L., & Milliken, C. S. (2006). Mental health problems, use of mental health services, and attrition from military service after returning from

deployment to Iraq or Afghanistan. *Journal of the American Medical Association,*
295(9), 1023-1032.

Hosek, J., Kavanagh, J., & Miller, L. (2006). *How deployment affects service members.*
Rand Corporation. Retrieved from:
http://www.rand.org/content/dam/rand/pubs/monographs/2005/RAND_MG432.s
um.pdf

Hotopf, M., Hull, L., Fear, N. T., Browne, T., Horn, O., Iversen, A., ... & Wessely, S. (2006).
The health of UK military personnel who deployed to the 2003 Iraq war: a cohort
study. *The Lancet, 367*(9524), 1731-1741.

Hyams, K. C. (2006). Mental health screening before troop deployment. *British Medical*
Journal, 333, 979–80.

Hyams, K. C., Wignall, F. S., & Roswell, R. (1996). War syndromes and their evaluation:
From the U.S. Civil War to the Persian Gulf War. *Annals of Internal Medicine, 125*,
398-405.

Iowa Persian Gulf Study Group. (1997). Self-reported illness and health status among Gulf
War veterans: A population-based study. *JAMA, 277*, 238-245.

Ishøy, T., Suadicani, P., Guldager, B., Appleyard, M., Hein, H. O., & Gyntelberg, F. (1999).
State of health after deployment in the Persian Gulf: The Danish Gulf War
study. *Danish Medical Bulletin, 46*(5), 416-419.

Jacobson, I. G., Ryan M. A. K., Hooper, T. I., Smith, T. C., Amoroso, P. J., Boyko, E. J., ... &
Bell, N. S. (2008). Alcohol use and alcohol-related problems before and after military
combat deployment. *Journal of the American Medical Association, 300*(6), 663-675.

Jakupcak, M., Conybeare, D., Phelps, L., Hunt, S., Holmes, H. A., Felker, B., ... & McFall, M.
E. (2007). Anger, hostility, and aggression among Iraq and Afghanistan war veterans
reporting PTSD and subthreshold PTSD. *Journal of Traumatic Stress*, *20*(6), 945-
954.

Jones, E., Hodgins-Vermaas, R., McCartney, H., Everitt, B., Beech, C., Poynter, D., ... &
Wessely, S. (2002). Post-combat syndroms from the Boer war to the Gulf war: A
cluster analysis of their nature and attribution. *British Medical Journal, 324*(7333),
321-324.

Kang, H. K., Mahan, C. M., Lee, K. Y., Magee, C. A., & Murphy, F. M. (2000). Illnesses
among United States veterans of the Gulf War: A population-based survey of 30,000
veterans. *Journal of Occupational & Environmental Medicine, 42*(5), 491-501.

Karney, B. R., Ramchand, R., Osilla, K. C., Caldarone, L. B., & Burns, R. M. (2008). *Invisible*
Wounds. Predicting the Immediate and Long-term Consequences of Mental Health
Problems in Veterans of Operation Enduring Freedom and Operation Iraqi
Freedom. Rand Corporation.

Keane, T. M., Fairbank, J. A., Caddell, J. M., Zimering, R. T., Taylor, K. L., & Mora, C. A.
(1989). Clinical evaluation of a measure to assess combat exposure. *Psychological*
Assessment: A Journal of Consulting and Clinical Psychology, *1*(1), 53-55.

Kehle S. M., Reddy, M. K., Ferrier-Auerbach, A. G., Erbes, C. R., Arbisi, P. A., & Polusny, M. A. (2011). Psychiatric diagnoses, comorbidity, and functioning in National Guard troops deployed to Iraq. *Journal of Psychiatric Research, 45*(1), 126-32.

King, L. A., King, D., Fairbank, J. A., Keane, T. M., & Adams, G. A. (1998). Resilience/recovery factors in posttraumatic stress disorder among female and male Vietnam veterans: Hardiness, postwar social support, and additional stressful life events. *Journal of Personality and Social Psychology, 74*, 420-34.

Konstantopoulos, S. (2006). Fixed and mixed effects models in meta-analysis. *IZA Discussion Paper no. 2198.*

Larson, G. E., Highfill-McRoy, R. M., & Booth-Kewley, S. (2008). Psychiatric diagnosis in historic and contemporary military cohorts: combat deployment and the healthy warrior effect. *American Journal of Epidemiology, 167*(11), 1269-1276.

Li, C.-Y., & Sung, F.-C. (1999). A review of the healthy worker effect in occupational epidemiology. *Occupational Medicine, 49*(4), 225-229.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical Meta-Analysis.* Thousand Oaks: Sage Publications.

Mental Health Advisory Team (MHAT-IV). (2006). Operation Iraqi Freedom 05-07, MHAT-IV *Report.* U.S. Department of the Army, Office of the Surgeon, Multinational Force–Iraq and Office of the Surgeon General, U.S. Army Medical Command

Murdoch, M., Bradley, A., Mather, S. H., Klein, R. E., Turner, C. L., & Yano, E. M. (2006). Women and war. *Journal of General Internal Medicine, 21*(S3), S5-S10.

National Defence and Canadian Forces. (2011). International Operations. Retrieved from http://www.cefcom-comfec.forces.gc.ca/pa-ap/ops/index-eng.asp. Accessed on April 27, 2011

NATO. (2014). NATO Operations and Missions. Retrieved from http://www.nato.int/cps/en/natolive/topics_52060.htm Accessed on January 10, 2014.

NATO. (n.d.). NATO Training Mission – Afghanistan (NTM-A). Retrieved from: http://www.aco.nato.int/page272701224.aspx

Obama, Barack. (2010). Weekly Address: President Obama Announces Changes to Help Veterans with PTSD Receive the Benefits They Need. Retrieved from http://www.whitehouse.gov/the-press-office/weekly-address-president-obama-announces-changes-help-veterans-with-ptsd-receive-be

O'Brien, L. S., & Hughes, S. J. (1991). Symptoms of post-traumatic stress disorder in Falklands veterans five years after the conflict. *British Journal of Psychiatry, 159*, 135-141.

O'Toole B. I., Schureck, R. J., Marshall, R. P., Grayson, D. A., & Dobson, M. (1999). Combat, dissociation and posttraumatic stress disorder in Australian Vietnam veterans. *Journal of Traumatic Stress, 12*, 625-640.

O'Toole, B. I., Outram, S., Catts, S. V., & Pierse, K. R. (2010). The mental health of partners of Australian Vietnam veterans three decades after the war and its relation to Veteran

military service, combat, and PTSD. *The Journal of Nervous and Mental Disease*, *198*(11), 841-845.

Oxford University Press. (2001). The Oxford Essential Dictionary of the U.S. Military.

Oxman, A. D., & Guyatt, G. H. (1992). A consumer's guide to subgroup analyses. *Annals of Internal Medicine*, *116*(1), 78-84.

Perl, L. (2007). *Veterans and Homelessness.* Congressional Research Service, Washington DC. Retrieved from http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA543259

Perconte, S. T., Wilson, A. T., Pontius, E. B., & Dietrick, A. L. (1993). Psychological and war stress symptoms among deployed and non-deployed reservists following the Persian Gulf War. *Military Medicine,* 158, 516-521.

Pierce, P. F. (1997). Physical and emotional health of Gulf War veteran women. *Aviation, Space and Environmental Medicine, 68*, 317-321.

Proctor, S. P., Heeren, T., White, R. F., Wolfe, J., Borgos, M. S., Davis, J. D., ... & Oznoff, D. (1998). Health status of Persian Gulf War veterans: self-reported symptoms, environmental exposures and the effect of stress. *International Journal of Epidemiology*, *27*(6), 1000-1010.

Pitman, R.K., Altman, B., & Macklin, M. L. (1989). Prevalence of posttraumatic stress disorder in wounded Vietnam veterans. *American Journal of Psychiatry, 146*, 667-669.

Ramos, C., & Leal, I. (2013). Posttraumatic Growth in the Aftermath of Trauma: A Literature Review About Related Factors and Application Contexts. *Psychology, Community & Health*, *2*(1), 43-54.

Regan, T. (2004). High survival rate for US troops wounded in Iraq. *Christian Science Monitor.*

Rona, R. J., Hyams, K. C., & Wessely, S. (2005). Screening for psychological illness in military personnel. *JAMA, 293*(19), 1257-1260.

Rona, R. J., Hooper, R., Jones, M., Iversen, A. C., Hull, L., Hotopf, D. M., & Wessely, S. (2009). The contribution of prior psychological symptoms and combat exposure to post Iraq deployment mental health in the UK military. *Journal of Traumatic Stress, 22(1)*, 11-19.

Ryan, M. A., Smith, T. C., Smith, B., Amoroso, P., Boyko, E. J., Gray, G. C., ... & Hooper, T. I. (2007). Millennium Cohort: Enrollment begins a 21-year contribution to understanding the impact of military service. *Journal of Clinical Epidemiology*, *60*(2), 181-191.

Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods, 8*(4), 448-467.

Sareen, J., Cox, B. J., Afifi, T. O., Stein, M. B., Belik, S. L., Meadows, G., & Asmundson, G. J. (2007). Combat and peacekeeping operations in relation to prevalence of mental disorders and perceived need for mental health care: findings from a large

representative sample of military personnel. *Archives of General Psychiatry*, *64*(7), 843-52.

Seelig, A. D., Jacobson, I. G., Smith, B., Hooper, T. I., Boyko, E. J., Gackstetter, G. D., ... & Smith, T. C. (2010). Sleep patterns before, during, and after deployment to Iraq and Afghanistan. *Sleep*, *33*(12), 1615-1622.

Smith, T. C., Ryan, M. A., Wingard, D. L., Slymen, D. J., Sallis, J. F., & Kritz-Silverstein, D. (2008). New onset and persistent symptoms of post-traumatic stress disorder self reported after deployment and combat exposures: prospective population based US military cohort study. *British Medical Journal*, *336*(7640), 366-371.

Smith, T. C., Jacobson, I. G., Hooper, T. I., LeardMann, C. A., Boyko, E. J., Smith, B., ... & Ryan, M. A. (2011). Health impact of US military service in a large population-based military cohort: findings of the Millennium Cohort Study, 2001-2008. *BMC Public Health*, *11*(1), 69-78.

Stimpson, N. J., Thomas, H. V., Weightman, A. L., Dunstan, F., & Lewis, G. (2003). Psychiatric disorder in veterans of the Persian Gulf War of 1991 Systematic review. *The British Journal of Psychiatry*, *182*(5), 391-403.

Steele, L. (2000). Prevalence and patterns of Gulf War illness in Kansas veterans: Association of symptoms with characteristics of person, place, and time of military service. *American Journal of Epidemiology*, *152*(10), 992-1002.

Street, A. E., Vogt, D., & Dutra, L. (2009). A new generation of women veterans: Stressors faced by women deployed to Iraq and Afghanistan. *Clinical Psychology Review*, *29*(8), 685-694.

Stretch, R. H., Bliese, P. D., Marlowe, D., & Wright, K. M. (1996a). Psychological health of Gulf War-era military personnel. *Military Medicine, 161*, 257-261.

Stretch, R. H., Marlowe, D. H., Wright, K. M., & Bliese, P. D. (1996). Post-traumatic stress disorder symptoms among Gulf War veterans. *Military Medicine, 161*, 407-410.

Tanielian, T. L., & Jaycox, L. H. (Eds.). (2008). *Invisible wounds of war: Psychological and cognitive injuries, their consequences, and service to assist recovery*. RAND Corporation.

Tedeschi, R. G., & Calhoun, L. G. (1996). The Posttraumatic Growth Inventory: Measuring the positive legacy of trauma. *Journal of Traumatic Stress*, *9*(3), 455-471.

Tedeschi, R. G., & Calhoun, L. G. (2004). Posttraumatic Growth: Conceptual foundations and empirical evidence. *Psychological Inquiry*, *15*(1), 1-18.

Thomsen, Y. D., Joanssen, R., Berntsen, D., Johannesen, K. B., Hoxer, C. R., Pedersen, ... & Bertelsen, M. (2010). Undersøgelse af psykiske efterreaktioner hos soldater udsendt til Afghanistan i perioden februar-august 2009 [Study of psychological reactions in soldiers deployed to Afghanistan February – August, 2009]. *Forsvarsakademiet.*

Thoresen, S. (2006). *Suicide in peacekeepers. A study of risk factors for suicide in Norwegian veterans of international military operations* (Unpublished doctoral dissertation). Faculty of Social Sciences, Norwegian University, Oslo.

United Nations. (2014a). Post-Cold War Surge. Retrieved from: http://www.un.org/en/peacekeeping/operations/surge.shtml

United Nations [UN]. (2014b). United Nations Peacekeeping Operations. Retrieved from http://www.un.org/en/peacekeeping/documents/operationslist.pdf. Downloaded on January 10, 2014.

Unwin, C., Blatchley, N., Coker, W., Ferry, S., Hotopf, M., Hull, L., ... & Wessely, S. (1999). Health of UK servicemen who served in Persian Gulf War. *The Lancet, 353*(9148), 169-178.

US Army. (2011). Retrieved from http://www.army.mil. Downloaded on April 28, 2011.

USF-Iraq. (2011). Retrieved from http://www.usf-iraq.com. Downloaded on April 28, 2011.

Vogt, D. S., King, D. W., & King, L. A. (2007). Risk pathways for PTSD. *Handbook of PTSD: Science and practice*, 99-114.

Warden, D. (2006). Military TBI during the Iraq and Afghanistan wars. *Journal of head trauma rehabilitation, 21*(5), 398-402.

Warner, C. H., Appenzeller, G. N., Parker, J. R., Warner, C. M, Hoge, C. W. (2011). Effectiveness of mental health screening and coordination of in-theater care prior to deployment to Iraq: A cohort study. *American Journal of Psychiatry, 168*(4), 378-85.

Weisæth, L. (2002). The European History of Psychotraumatology. *Journal of Traumatic Stress, 15*, 443-452.

Wells, T. S., LeardMann, C. A., Fortuna, S. O., Smith, B., Smith, T. C., Ryan, M. A., ... & Blazer, D. (2010). A prospective study of depression following combat deployment in support of the wars in Iraq and Afghanistan. *American Journal of Public Health, 100*(1), 90-99.

Wessely, S. (2006). Twentieth-century theories on combat motivation and breakdown. *Journal of Contemporary History, 41*(2), 269–86.

Wolfe, J., Proctor, S. P., Erickson, D. J., Heeren, T., Friedman, M. J., Huang, M. T., ... & White, R. F. (1999). Relationship of psychiatric status to Gulf War veterans' health problems. *Psychosomatic Medicine*, *61*(4), 532-540.

Zatzick, D. F., Marmar, C. R., Weiss, D. S., Browner, W. S., Metzler, T. J., Golding, J. M., ... & Wells, K. B. (1997). Posttraumatic stress disorder and functioning and quality of life outcomes in a nationally representative sample of male Vietnam veterans. *American Journal of Psychiatry, 154*(12), 1690-1695.

5 Appendices

5.1 STUDY ELIGIBILITY SCREENING LEVEL ONE AND TWO

Screening level one is based on title and abstract. Second level screening is on the basis of full text.

Reference id. No.:

Study id. No.:

Reviewer's initials:

Source:

Year of Publication:

Mission/Period Covered:

Nationality:

Author:

1. **Are the participants in the study deployed military personnel?**
   Yes (include)
   No (if no stop here and exclude)
   Uncertain (include)

Guidance: Deployment to international operations is understood as the deployment of military personnel outside their country of origin. Deployed personnel may include combatants as well as non-combatants (e.g. logistics and support functions). Military personnel deployed abroad that do not directly participate in military operations are not relevant. For example being deployed to a US military base in Germany is not relevant.

2. **Is the study about the effect of deployment to international military operations after January 1st, 1989?**
   Yes (include)

No (if no stop here and exclude)

Uncertain (include)

Guidance: To be included the study should measure the consequences to military personnel of being deployed to international military operations. Only deployments which started after January 1st, 1989 should be included. The intervention is "being deployed". For instance studies that include a comparison group to the deployed are included. A study that is not included is a study that uses a sample of military personnel that has been deployed, but the intervention of interest is not the effect of deployment itself. For example evaluating the effect of treatment of PTSD in a sample of deployed military personnel is not relevant for the review.

3. **Is the study about the effect of deployment on mental health and social functioning?**

   Yes (include)

   No (if no stop here and exclude)

   Uncertain (include)

Guidance: To be included the study should evaluate the effects of deployment on measures of mental health (for instance posttraumatic stress disorder, major depression, depression, anxiety, aggression, etc.) and/or on social functioning (e.g. employment, homelessness, marital status, etc.). The measured effect should be on those that were deployed, not the effect deployment had on e.g. their next of kin, children, etc.

The report reference is excluded if one (or more) of the answers to questions 1, 2, or 3 are no. If the answers are yes or uncertain the full report is retrieved for second level screening. Questions that were answered as 'uncertain' need to be answered again based on full text. If information is not available or the report is unclear, report authors will be contacted to clarify study eligibility.

Questions for second level screening

4. **Does the study compare a sample of deployed military personnel to a sample of non-deployed military personnel?**
   Yes
   No
   Uncertain (state reason)

Guidance: To be included the study should compare a sample of military personnel to a sample of non-deployed military personnel. For instance a study that compares deployed military personnel to Operation Enduring Freedom with non-deployed military personnel is eligible. A study that compares deployed military personnel to a civilian sample is not included.

5. **Does the study compare samples of deployed military personnel who are deployed to operations with different combat intensities?**
   Yes
   No
   Uncertain (state reason)

Guidance: To be included the study should compare a sample of military personnel deployed in one combat situation to a sample of military personnel deployed in a combat situation with different intensity. For instance a study that compares two groups of military personnel who were both deployed to the same mission but in different regions is eligible.

If the answers to 4 and 5 are both no, then the study is excluded from review.

5.2 DATA EXTRACTION

Mission characteristics

1. **Mission Location**

    Specify country (and city if possible)

2. **Type of Mission**

   Peacekeeping

   Peace enforcing

   Other (specify, pg. #)

3. **Mission Command**

   NATO

   UN

   International coalition (Specify)

   Other (specify)

   Can't tell

Deployment characteristics

4. **Location of Deployment (check all that apply)**

   War zone

   Civil conflict zone

   Other, specify

   Can't tell

5. **Deployment Duration/Length of Deployment**

   Deployment start: _____

   Deployment end: _____

Participant characteristics

	Participants	Pg. # & NOTES
Gender (e.g. % male)		
Age		
Ethnicity (e.g. % white)		
Nationalities		
Branch of Service (% Army/Navy/Air Force)		
Rank (e.g. % officers)		
Duty Status (%Active/Reserve)		
Family situation (e.g. % married)		
Prior History of Mental Health (e.g. % with prior mental health treatment)		
Number of previous deployments		
Combat patrols (% of sample, # patrols)		
Experienced enemy fire (% of sample, # incidents)		
Discharged weapons (% of sample, # times)		
Experienced wounded or casualties (% of sample)		
Responsible for death of enemy combatant/civilian (% of sample)		
Handle dead bodies: personnel/civilian (% of sample)		
Experience IED/landmines/artillery/rockets (% of sample)		
Combat Exposure/intensity, overall assessment (e.g. % High/Med/Low/None)		

Outcome measures

Instructions: Please enter outcome measures in the order in which they are described in the report. Note that a single outcome measure can be completed by multiple sources and at multiple points in time (data from specific sources and time-points will be entered later).

#	Outcome & measure	Reliability & Validity	Format	Direction	Source	Mode Admin	Blind (outcome assessors)?	Pg# & notes

1		Info from:
		   Other samples
		   This sample
		   Unclear
		
		Info provided:
			Dichotomy
			Continuous
				High score or event is
				
				Positive
				Negative
				Can't tell
					Researcher
					Clinician
					Military
					Admin data
					Other
					Unclear
						Self-admin
						Interview
						Other
							Yes
							No
							Can't tell	

* Repeat as needed

OUT COME DATA

DICHOTOMOUS OUTCOME DATA

OUTCOME	TIME POINT (record exact time taken from baseline)	SOURCE	VALID Ns	CASES	NON-CASES	STATISTICS	Pg. # & NOTES
	•1st measure after baseline •1st follow-up • 2nd follow-up • 3rd follow-up •4th follow-up • other	• researcher • clinician • military • admin data • other (specify)	Deployment Comparison	Deployment Comparison	Deployment Comparison	RR (risk ratio) OR (odds ratio) SE (standard error) 95% CI DF P- value (enter exact p value if available) Chi2 Other Covariates (control variables)	

*Repeat as needed

CONTINUOUS OUTCOME DATA

Enter change and gain scores under Statistics (Other)

OUTCOME	TIME POINT (record exact time taken from baseline)	SOURCE (specify)	VALID Ns	Means	SDs	STATISTICS	Pg. # & NOTES
	•1st measure after baseline •1st follow-up • 2nd follow-up • 3rd follow-up •4th follow-up • other	• researcher • clinician •admin data • other (specify)	Deployment	Deployment	Deployment	P t F Df ES Covariates Other	
			Comparison	Comparison	Comparison		

*Repeat as needed

5.3 ASSESSMENT OF RISK OF BIAS IN INCLUDED STUDIES

Risk of bias table

Item	Judgement[a]	Description (quote from paper, or describe key information)
1. Sequence generation		
2. Allocation concealment		
3. Confounding[b,c]		
4. Blinding?[b]		
5. Incomplete outcome data addressed?[b]		
6. Free of selective reporting?[b]		
7. Free of other bias?		
8. *A priori* protocol?[d]		
9. *A priori* analysis plan?[e]		

[a] Some items on low/high risk/unclear scale (double-line border), some on 5 point scale/unclear (single line border), some on yes/no/unclear scale (dashed border). For all items, record "unclear" if inadequate reporting prevents a judgement being made.

[b] For each outcome in the study.

[c] This item is only used for NRCTs and NRSs. It is based on list of confounders considered important at the outset and defined in the protocol for the review (*assessment against worksheet*).

[d] Did the researchers write a protocol defining the study population, intervention and comparator, primary and other outcomes, data collection methods, etc. in advance of starting the study?

[e] Did the researchers have an analysis plan defining the primary and other outcomes, statistical methods, subgroup analyses, etc. in advance of starting the study?

Risk of bias tool

Studies for which RoB tool is intended

The risk of bias model is developed by Prof. Barnaby Reeves in association with the Cochrane Non-Randomised Studies Methods Group.[10] This model, an extension of the Cochrane Collaboration's risk of bias tool, covers both risk of bias in randomised controlled trials (RCTs and QRCTs), but also risk of bias in non-randomised studies (NRCTs and NRSs).

The point of departure for the risk of bias model is the Cochrane Handbook for Systematic Reviews of interventions (Higgins and Green, 2008). The existing Cochrane risk of bias tool needs elaboration when assessing non-randomised studies because, for non-randomised studies, particular attention should be paid to selection bias / risk of confounding. Additional item on confounding is used only for non-randomised studies (NRCTs and NRSs) and is not used for randomised controlled trials (RCTs and QRCTs).

Assessment of risk of bias

Issues when using modified RoB tool to assess included non-randomised studies:

- Use existing principle: score judgment and provide information (preferably direct quote) to support judgment
- Additional item on confounding used only for non-randomised studies (NRCTs and NRSs).
- 5-point scale for <u>some</u> items (distinguish "unclear" from intermediate risk of bias).
- Keep in mind the general philosophy – assessment is <u>not</u> about whether researchers could have done better but about risk of bias; the assessment tool must be used in a standard way whatever the difficulty / circumstances of investigating the research question of interest and whatever the study design used.
- Anchors: "1/No/low risk" of bias should correspond to a high quality RCT. "5/high risk" of bias should correspond to a risk of bias that means the findings should not be considered (too risky, too much bias, more likely to mislead than inform)

1. Sequence generation
- Low/high/unclear RoB item
- Always high RoB (not random) for a non-randomised study

---

[10] This risk of bias model was introduced by Prof. Reeves at a workshop on risk of bias in non-randomised studies at SFI Campbell, February 2011. The model is a further development of work carried out in the Cochrane Non-Randomised Studies Method Group (NRSMG).

- Might argue that this item redundant for NRS since always high – but important to include in RoB table ('level playing field' argument)

2. Allocation concealment
- Low/high/unclear RoB item
- Potentially <u>low</u> RoB for a <u>non-randomised study</u>, e.g. quasi-randomised (so high RoB to sequence generation) but concealed (reviewer judges that the people making decisions about including participants didn't know how allocation was being done, e.g. odd/even date of birth/hospital number)

3. RoB from confounding (<u>additional item for NRCT and NRS; assess for each outcome</u>)
- Assumes a <u>pre-specified</u> list of potential confounders defined in the protocol
- Low(1) / 2 / 3 / 4 / high(5) / unclear RoB item
- Judgment needs to factor in:
    o proportion of confounders (from pre-specified list) that were considered
    o whether most important confounders (from pre-specified list) were considered
    o resolution/precision with which confounders were measured
    o extent of imbalance between groups at baseline
    o care with which adjustment was done (typically a judgment about the statistical modeling carried out by authors)
- Low RoB requires that all important confounders are balanced at baseline (<u>not primarily/not only</u> a statistical judgment OR measured 'well' <u>and</u> 'carefully' controlled for in the analysis).

Assess against pre-specified worksheet. Reviewers will make a RoB judgment about each factor first and then 'eyeball' these for the judgment RoB table.

4. RoB from lack of blinding (<u>assess for each outcome</u>, as per existing RoB tool)
- Low(1) / 2 / 3 / 4 / high(5) / unclear RoB item
- Judgment needs to factor in:
    o nature of outcome (subjective / objective; source of information)
    o who was / was not blinded and the risk that those who were not blinded could introduce <u>performance or detection</u> bias.
    o see Ch.8

5. RoB from incomplete outcome data (<u>assess for each outcome</u>, as per existing RoB tool)

- Low(1) / 2 / 3 / 4 / high(5) / unclear RoB item
- Judgment needs to factor in:
    - reasons for missing data
    - whether amount of missing data balanced across groups, with similar reasons
    - see Ch.8

6. RoB from selective reporting (<u>assess for each outcome</u>, NB different to existing Ch.8 recommendation)
- Low(1) / 2 / 3 / 4 / high(5) /unclear RoB item
- Judgment needs to factor in:
    - existing RoB guidance on selective outcome reporting (see Ch.8)
    - also, extent to which analyses (and potentially other choices) could have been manipulated to bias the findings reported, e.g. choice of method of model fitting, potential confounders considered / included
    - look for evidence that there was a protocol in advance of doing any analysis / obtaining the data (difficult unless explicitly reported); NRS very different from RCTs. RCTs must have a protocol in advance of starting to recruit (for REC/IRB/other regulatory approval); NRS need not (especially older studies)
    - Hence, separate yes/no items asking reviewers whether they think the researchers had a pre-specified protocol and analysis plan.

7. RoB from other bias (<u>assess for each outcome</u>, NB different to existing Ch.8 recommendation)

- Low(1) / 2 / 3 / 4 / high(5) /unclear RoB item
- Judgment needs to factor in:
    - existing RoB guidance on other potential threats to validity (see Ch.8) also, assess whether suitable cluster analysis is used (e.g. cluster summary statistics, robust standard errors, the use of the design effect to adjust standard errors, multilevel models and mixture models), if assignment of units to treatment is clustered
- For this review a particular bias from lack of blinding may be present for outcomes collected from self-reporting. Depending on whether relevant authorities (DoD/VA/etc.) could view individual level results. For example answers to PTSD questionnaire might affect career path and disability claims. Important items to consider: Who collected data?/Who was able to view the data?/Were individuals aware of the purpose of the study?/Were the procedures similar for control and deployment subsamples?

Confounding Worksheet

Assessment of how researchers dealt with confounding	
Method for *identifying* relevant confounders described by researchers:　　　　yes 　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　no If yes, describe the method used:	☐ ☐
Relevant confounders described:　　　　　　　　　　　　yes 　　　　　　　　　　　　　　　　　　　　　　　　　　　　no 　　　List confounders described on next page	☐ ☐
Method used for controlling for confounding 　　　At design stage (e.g. regression discontinuity, instrumental variables): 　　　　　　　　　　　　　　　　　　　　　　.. 　　　　　　　　　　　　　　　　　　　　　　.. 　　　　　　　　　　　　　　　　　　　　　　.. 　　　At Analysis stage (e.g. matching, stratification, regression, difference-in-difference): 　　　　　　　　　　　　　　　　　　　　　　.. 　　　　　　　　　　　　　　　　　　　　　　.. 　　　　　　　　　　　　　　　　　　　　　　.. 　　　Describe confounders controlled for below	

Confounders described by researchers

Tick (yes[0]/no[1] judgment) if confounder considered by the researchers [Cons'd?]

Score (1[good precision] to 5[poor precision]) precision with which confounder measured

Score (1[balanced] to 5[major imbalance]) imbalance between groups

Score (1[very careful] to 5[not at all careful]) care with which adjustment for confounder was carried out

Confounder	Considered	Precision	Imbalance	Adjustment
Mental Health History	☐	☐	☐	☐
Previous Deployments	☐	☐	☐	☐
Duty Status	☐	☐	☐	☐
Branch	☐	☐	☐	☐
Ethnicity	☐	☐	☐	☐
Rank	☐	☐	☐	☐
Age	☐	☐	☐	☐
Gender	☐	☐	☐	☐
Unobservables[11]	☐	Irrelevant	☐	☐
Other:	☐	☐	☐	☐

---

[11] See user guide for unobservables

User guide for unobservables

Selection bias is understood as systematic baseline differences between groups and can therefore compromise comparability between groups. Baseline differences can be observable (e.g. age and gender) and unobservable (to the researcher; e.g. motivation and 'ability'). There is no single non-randomised study design that always solves the selection problem. Different designs solve the selection problem under different assumptions and require different types of data. Especially how different designs deal with selection on unobservables varies. The "right" method depends on the model generating participation, i.e. assumptions about the nature of the process by which participants are selected into a programme.

As there is no universal correct way to construct counterfactuals we will assess the extent to which the identifying assumptions (the assumption that makes it possible to identify the counterfactual) are explained and discussed (preferably the authors should make an effort to justify their choice of method). We will look for evidence that authors using e.g. (this is NOT an exhaustive list):

Natural experiments:
Discuss whether they face a truly random allocation of participants and that there is no change of behaviour in anticipation of e.g. policy rules.

Instrument variable (IV):
Explain and discuss the assumption that the instrument variable does not affect outcomes other than through their effect on participation.

Matching (including propensity scores):
Explain and discuss the assumption that there is no selection on unobservables, only selection on observables.

(Multivariate) Regression:
Explain and discuss the assumption that there is no selection on unobservables, only selection on observables. Further discuss the extent to which they compare comparable people.

Regression Discontinuity (RD):

Explain and discuss the assumption that there is a (strict!) RD treatment rule. It must not be changeable by the agent in an effort to obtain or avoid treatment. Continuity in the expected impact at the discontinuity is required.

Difference-in-difference (Treatment-control-before-after):

Explain and discuss the assumption that outcomes of participants and nonparticipants evolve over time in the same way.

SOURCES OF FUNDING

SFI Campbell

DECLARATIONS OF INTEREST

Please declare any potential conflicts of interest.

REVIEW AUTHORS

Lead review author:
The lead author is the person who develops and co-ordinates the review team, discusses and assigns roles for individual members of the review team, liaises with the editorial base and takes responsibility for the on-going updates of the review.

Name:	Martin Bøg (MB)
Title:	Researcher
Affiliation:	SFI Campbell
Address:	Herluf Trolles Gade 11
City, State, Province or County:	Copenhagen
Postal Code:	DK-1052 Copenhagen K
Country:	Denmark
Phone:	+4533697741
Mobile:	+4533697741
Email:	mbg@sfi.dk

Co-author(s): (There should be at least one co-author)

Name:	Trine Filges (TF)
Title:	Senior Researcher
Affiliation:	SFI Campbell
Address:	Herluf Trolles Gade 11
City, State, Province or County:	Copenhagen
Postal Code:	DK-1052 Copenhagen K
Country:	Denmark

Phone:	+4533480926
Mobile:	+4533480926
Email:	tif@sfi.dk

Co-author(s): (There should be at least one co-author)

Name:	**Stephanie Vincent Lyk-Jensen (SVJ)**
Title:	Senior Researcher
Affiliation:	SFI
Address:	Herluf Trolles Gade 11
City, State, Province or County:	Copenhagen
Postal Code:	DK-1052 Copenhagen K
Country:	Denmark
Phone:	+4533480934
Mobile:	+4533480934
Email:	svj@sfi.dk

Co-author(s): (There should be at least one co-author)

Name:	**Anne-Marie Klint Jørgensen (AMK)**
Title:	Information Specialist
Affiliation:	SFI Campbell
Address:	Herluf Trolles Gade 11
City, State, Province or County:	Copenhagen
Postal Code:	DK-1052 Copenhagen K
Country:	Denmark
Phone:	+4533480868
Mobile:	+4533480868
Email:	amk@sfi.dk

Co-author(s): (There should be at least one co-author)

Name:	**Maia Lindstrøm**
Title:	

Affiliation:	
Address:	
City, State, Province or County:	
Postal Code:	
Country:	
Phone:	
Mobile:	
Email:	

Duplicate the above table as necessary to include all co-authors.

ROLES AND RESPONSIBILITIES

Please give brief description of content and methodological expertise within the review team. The recommended optimal review team composition includes at least one person on the review team who has content expertise, at least one person who has methodological expertise and at least one person who has statistical expertise. It is also recommended to have one person with information retrieval expertise.

Roles	Name(s)	Background, Skills, Areas of Expertise
Content	SVJ	PhD Economics, Content specialist, Deployment of Danish Army members abroad
Systematic Review Methods	TF, AMK, ML	TF: PhD Econ, extensive authorship on Campbell reviews, content specialist on labor market. AMK: Librarian, extensive authorship on Cambell reviews, information retrieval, Reference Manager. ML: MSc Sociology, extensive authorship on Campbell reviews, project management, coordination.
Statistical Analysis	MB, TF	MB: PhD Econ; applied microeconometrics, field experiments; TF: PhD Econ; applied microeconometrics, meta analysis,
Information Retrieval	AMK	AMK: See above
Other Roles	Describe as needed	

ACKNOWLEDGEMENTS

Acknowledge all the individuals and organizations contributing to the preparation of the protocol that are not identified in prior sections.

Trondur Møller Sandoy, Ulrik Højmark Pedersen, Rasmus Henriksen Klokker, Marcel Mirzaei-Fard, Asta Breinholdt Lund, and Ida Scheel Rasmussen will act as screeners and descriptive coders.

EXPECTED TIMEFRAME

A timetable with target dates for accomplishing the key tasks required to complete the review. If the review is not submitted within 24 months of protocol approval, the review area may be opened up for other authors.

November 2014: Literature Search

December 2014-February 2015: Screening (1st and 2nd level), double screening (humans).

March – May 2015: Coding of included studies; descriptive, numerical and risk of bias coding (two coders for each study)

June – July 2015: Analysis, including meta-analysis studies permitting.August 2015: Complete draft of review manuscript

September 2015: Submit manuscript to Campbell

PLANS FOR UPDATING THE REVIEW

Describe any plan for updating the review once it is completed.

Once the review has been completed and placed in the Campbell library, we plan to update the review with a frequency of 2 years. MB will be responsible.

AUTHORS' RESPONSIBILITIES

Authors' responsibilities
By completing this protocol, you accept responsibility for preparing, maintaining, and updating the review in accordance with Campbell Collaboration policy. The Coordinating Group will provide as much support as possible to assist with the preparation of the review.

A draft review must be submitted to the Coordinating Group within two years of protocol acceptance. If drafts are not submitted before the agreed deadlines, or if we are unable to contact you for an extended period, the Coordinating Group has the right to de-register the

title or transfer the title to alternative authors. The Coordinating Group also has the right to de-register or transfer the title if it does not meet the standards of the Coordinating Group and/or the Campbell Collaboration.

You accept responsibility for maintaining the review in light of new evidence, comments and criticisms, and other developments, and updating the review every five years, when substantial new evidence becomes available, or, if requested, transferring responsibility for maintaining the review to others as agreed with the Coordinating Group.

Publication in the Campbell Library
The support of the Coordinating Group in preparing your review is conditional upon your agreement to publish the protocol, finished review, and subsequent updates in the Campbell Library. The Campbell Collaboration places no restrictions on publication of the findings of a Campbell systematic review in a more abbreviated form as a journal article either before or after the publication of the monograph version in *Campbell Systematic Reviews*. Some journals, however, have restrictions that preclude publication of findings that have been, or will be, reported elsewhere and authors considering publication in such a journal should be aware of possible conflict with publication of the monograph version in *Campbell Systematic Reviews*. Publication in a journal after publication or in press status in *Campbell Systematic Reviews* should acknowledge the Campbell version and include a citation to it. Note that systematic reviews published in *Campbell Systematic Reviews* and co-registered with the Cochrane Collaboration may have additional requirements or restrictions for co-publication. Review authors accept responsibility for meeting any co-publication requirements.
I understand the commitment required to undertake a Campbell review, and agree to publish in the Campbell Library.

Signed on behalf of the authors: Martin Bøg
Form completed by: Martin Bøg   **Date: September 18, 2014**