



The impact of shift work schedules on PVT performance in naturalistic settings: a systematic review

Matthew Ferris^{1,3} · Kelly-Ann Bowles² · Mikaela Bray³ · Emma Bosley³ · Shantha M. W. Rajaratnam⁴ · Alexander P. Wolkow⁴

Received: 2 September 2020 / Accepted: 8 February 2021 / Published online: 11 March 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH, DE part of Springer Nature 2021

Abstract

Objective The Psychomotor Vigilance Test (PVT) is considered the gold standard for detecting sleep loss and circadian misalignment related changes in performance in laboratory and field settings. This short 3-, 5- or 10-min test appraises an individual's sustained vigilant attention on a visual stimulus through reaction time, false starts and performance lapses. The PVT has been widely used as a measure to assess vigilant attention among shift workers, but information evaluating the application and performance of this test in different naturalistic shift work settings is limited. The purpose of this review is to synthesise and evaluate existing literature which has used the PVT to assess and monitor psychomotor performance in response to shift work schedules and rosters performed in real-world settings.

Methods A systematic search of studies examining PVT performance in response to 24/7 shift work schedules (e.g., day, afternoon, evening and night shifts) performed under naturalistic conditions was conducted. Articles were identified by searching Medline, Embase, CINHAL and PsycINFO databases in April 2020.

Results The search yielded 135 results, of which 16 publications were suitable to be included in this review. Articles were grouped according to when the PVT was applied to a research cohort, which included (a) multiple instances per shift, (b) commencement and cessation of shift and (c) other varying times.

Conclusions This review suggests PVT performance is typically congruent across studies when the test is applied at generally consistent time intervals. The lack of research concerning the use of the PVT during extended duty shifts (e.g., shifts and on call work > 30 h) is an area for future research.

Keywords Sleep · Fatigue · Psychomotor vigilance test

Background

Sleepiness and sleep deprivation are more prevalent now than they have ever been before with the shift towards a 24/7 society (Bonfond et al. 2006). As a result, diurnal working patterns are no longer considered normal

(Ganesan et al. 2019). Industries, such as manufacturing, defence and healthcare are heavily reliant on this highly functioning, nonstop workforce. Circadian mismatch can result from poor sleeping habits, rapidly rotating shifts, nonconventional business hours and a poor sleeping environment (Bonfond et al. 2006; Ganesan et al. 2019). As a result, fatigue, decreased alertness and lowered employee morale can prevail throughout a workforce. This poses many challenges within the 24/7 industries; however, many contemporary approaches exist to combat fatigue, such as regular rostering reform, fatigue related education and fatigue monitoring (Sadeghniai-Haghighi and Yazdi 2015; Wolkow et al. 2019). Techniques to monitor fatigue have been deemed essential in some industries to ensure employee safety and optimise productivity and performance (Wolkow et al. 2019).

✉ Matthew Ferris
matthew.ferris@monash.edu

¹ Monash University, Wellington Rd, Clayton, Victoria 3800, Australia

² Department of Paramedicine, Monash University, Melbourne, Australia

³ Queensland Ambulance Service, Brisbane, QLD, Australia

⁴ Turner Institute for Brain and Mental Health, Monash University, Clayton, Victoria, Australia

Fatigue can be evaluated through various objective or subjective measures. Employers often implement a series of tests to monitor and assess fatigue in employees to determine if they pose an increased risk or are safe to be at work. Quantitative fatigue information can be sought through objective testing. Objective tests for measuring fatigue can include computer/electronic-based assessments, such as the psychomotor vigilance test (PVT) (Basner et al. 2011), as well as drowsiness measures assessed via biological markers, such as saliva (Harris et al. 2010), urine (Flynn-Evans et al. 2018) and blood (Wolkow et al. 2019). The latter three; however, can be considered to be more invasive and take hours-to-days to be conclusive—rendering them incompatible as a rapid measure.

The PVT is a short reaction time test used to measure objective vigilant attention (Basner et al. 2011; Basner and Dinges 2011; Arsintescu et al. 2017). The test requires an individual to press a key in response to a visual stimulus displayed on an electronic screen (Basner and Dinges 2011; Arsintescu et al. 2017). Typically, this can be in the form of a handheld electronic device, mobile telephone or tablet, or desktop computer/laptop. Through the recruitment of the prefrontal, motor and visual cortex, the PVT tests the cognitive domain of vigilant attention (Basner et al. 2011). Performance on the PVT is typically based on three elements: reaction time, performance lapses and false starts (Basner et al. 2011; Basner and Dinges 2011; Arsintescu et al. 2017). For instance, an individual's reaction time is determined as the duration from which the stimulus is first displayed, through to when it is acknowledged by pressing the key. A performance lapse occurs when the key is pressed too quickly, typically ≤ 100 ms, or too slowly, ≥ 500 ms (Basner and Dinges 2011; Arsintescu et al. 2017). Finally, a false start is registered when the key is pressed without a stimulus. The PVT tests can vary between 3, 5 and 10 min in duration; however, each of these tests have similar performance outcomes (i.e., reaction time, performance lapses and false starts) and have high levels of specificity (Basner et al. 2011; Basner and Dinges 2011; Arsintescu et al. 2017). The PVT has been described as the gold standard for fatigue observation due to its sensitivity in detecting sleepiness (Arsintescu et al. 2017).

The Karolinska Sleepiness Scale (KSS) is a subjective situational sleepiness measure used to judge attention and performance whilst undertaking routine tasks (Miley et al. 2016). It is a 9-point self-rated scale of one's propensity to fall asleep at that moment in time (i.e., 1 = extremely alert; 9 = very sleepy, great effort to keep alert, fighting sleep) (Miley et al. 2016). Previous literature has frequently collected data from both the PVT and KSS to assist fatigue analysis in response to differing shift schedules. As a result, both PVT and KSS will be examined in this review.

Fatigue is an important factor for individuals and organisations alike to monitor and review to ensure safety within a workplace. We know alertness in the workforce is essential for health and safety. Subjective fatigue measures may not necessarily align with objective measures, as results can be falsified (to meet an accepted fatigue level to be at work) or people may not comprehend each subjective criterium—resulting in distorted scores (Flynn-Evans et al. 2018). This subjectivity can result in unsafe working conditions for employees. As a result, an objective measure (e.g., PVT) is optimal for employers due to decreased ability to cheat or make fraudulent declarations. Although many studies have examined the influence of shift rosters on PVT performance in naturalistic/real-world settings, there are limited reviews which have been conducted to examine this area of the literature as a whole. This systematic review sought to evaluate PVT performance among shift workers in response to different 24/7 rostering schedules (i.e., day, afternoon, evening and night shifts) performed under naturalistic conditions.

Methods

The design, implementation and reporting of this systematic review has been conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al. 2009).

Data sources and search strategy

The online databases searched for this review included Embase, Medline via OVID, PsycINFO and CINAHL. The literature search was conducted in April 2020 with no exclusion search filters applied. The Population, Intervention and Outcome (PIO) framework (Sayers 2008) was used to construct the search terms used to search each database. These included (P) shift work* AND (I) psychomotor vigilance test OR PVT OR reaction time OR Joggle AND (O) alert* OR recovery OR fitness* OR response* OR safety OR performance OR fatigue OR health OR sleep*.

Records identified in the initial search were uploaded into Covidence[®] (Veritas Health Innovation 2017), an electronic data extraction and referencing tool, where duplicates were removed. Abstract and title screening of the records were completed by two independent authors searching for suitable articles, with conflicts resolved via phone discussion between the authors. The remaining articles' full texts were uploaded and reviewed against the predetermined inclusion and exclusion criteria, independently by the same two authors. The conflicts were addressed by a third author through the same software. Hand searching was completed by reviewing all the reference lists of the included articles; however, no further articles were identified.

Inclusion and exclusion criteria

Articles were excluded if they met any of the following criteria: (a) not written in English, (b) opinion pieces, (c) participants who did not work a rotating 24/7 roster (i.e., any combination of day, afternoon, evening or night shifts), (d) systematic reviews or meta-analyses, (e) studies that examined the influence of factors outside of the roster (e.g., caffeine or food) on PVT performance, (f) studies that took place in a simulated or laboratory settings and (g) if no full text was available.

Study quality assessment

A quality assessment tool, as published by the National Heart, Lung, and Blood Institute (NHLBI) was chosen to evaluate the risk of bias in each of the included studies (National Institute of Health 2014). The tool was called the ‘Quality assessment tool for observational cohort and cross-sectional studies’. Each article was independently reviewed by two authors and conflicts were addressed via phone correspondence. Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines were used by the lead author to assess the primary evidence (Atkins et al. 2004).

Synthesis of findings

The data extraction was performed by the lead author and separated into two main summaries. The first, an overall summary of the included studies based on the following categories: (a) study design, (b) country where the study was conducted, (c) occupational group/industry, (d) shift/roster type and (e) fatigue-related measures, (f) study aims and conclusions, (g) NHLBI rating and (h) relationship between fatigue measures (Appendix, Table 2). In the second summary, the data from the included studies were extracted based on the following categories: (a) sample size, (b) age range, (c) gender ratio, (d) sleep between day shifts, (e) KSS results, (f) PVT duration, PVT timing, PVT model, (g) PVT performance outcomes (i.e., PVT mean reaction time, PVT mean lapses) and (h) PVT application (Table 1). Many articles did not report their raw data, rather a descriptive analysis or graphical representation. Owing to large variability between studies in rostering patterns and inconsistent application of objective and subjective fatigue measures, a meta-analysis of the study results was not feasible. Articles were grouped broadly into categories associated with the application of the PVT for discussion and comparison. These included (a) multiple instances per shift, (b) commencement and cessation of shift and (c) other varying times throughout the shift.

Results

Search results

The search strategy yielded 135 results, of which 23 were duplicates and removed from the review. Abstracts and titles of the remaining 112 articles were then screened for eligibility based on the ‘Inclusion and exclusion criteria’ outlined above, and from these, 60 records were excluded (Fig. 1). Fifty-two articles were evaluated at the full text level, of which a further 35 were removed as per Fig. 1. After the screening and eligibility assessment, 16 articles were included in the review and progressed to data extraction (Fig. 1). No new studies were identified through hand searching (Fig. 1).

Study characteristics

Included articles were published between 2006 and 2019. Studies originated from the United States ($n=8$), Finland ($n=3$), Norway ($n=2$), Australia ($n=2$) and Germany ($n=1$). There was a large representation of occupational industries including healthcare (nurses, support staff and doctors) ($n=7$), airline support staff ($n=3$), mining ($n=3$), police officers ($n=1$), military ($n=1$) and airline pilots ($n=1$). The roster and shift types examined in the included studies varied considerably between randomised shifts to a progression forward or reverse rotating roster. Inclusion criteria resulted in each study using the PVT as a primary measure of fatigue. In addition, secondary fatigue-related measures included were: the KSS ($n=8$), other recognised subjective sleepiness scales ($n=5$), sleep duration assessed using actigraphy ($n=8$) and sleep/work diaries ($n=3$), biological markers of fatigue ($n=2$) and electroencephalography (EEG) recordings ($n=1$) (Appendix 1, Table 2).

Half of the included studies (50%) used the 10-min version of the PVT and the most common modality (25%) was the palm pilot version of the PVT, however many of the included studies (31.3%) did not indicate how the PVT was applied. Half (50%) of the included studies requested participants to undertake the PVT before, during and after their shift, while others administered the PVT based specifically on the schedule/roster (19%) or at a single timepoint in the morning, afternoon or night (19%).

Quality assessment

Risk of bias

As per the NHLBI Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies (National Institute of Health 2014), all included studies were of fair

Table 1 Details of included studies

References	Country	Sample size	Age range	Gender Ratio M:F	Sleep between day shifts (hours)	Sleep between night shifts (hours)	KSS ^a	PVT duration; PVT timing; PVT model	PVT ^b Mean RT ^c (ms)	PVT mean lapses (≥ 500 ms)	PVT Application
Multiple instances per shift											
Anderson et al. (2012)	United States	34	28 ± 1.83	23:11	6.3	6.3	–	10-min Start; middle and end of shift; PVT Model Unk ^d	Beginning of the shift, 284.6 ± 9.1 End of shift: 379.4 ± 23.9	Beginning of shift, 2.7 ± 0.59 End of Shift 6.7 ± 0.82	PVT completed before, during (each 6 h) and after each shift period
Bjorvatn et al. (2006)	Norway (North Sea)	17	22–55	16:01	–	–	Mean day: 2.4 mean night: 4.7 <i>p</i> < 0.05	5-min Taken; 1200, 1500, 1800 OR 0000, 0300 and 0600; palm PVT	Significant increase in RT at the end of the first night and moderate increase at the end of the second night, all other values are fairly consistent	–	KSS was administered every hour while on shift PVT administered at three time points 0000, 0300 and 0600 on nights and 1200, 1500 and 1800 on days. Each test was taken on days 1,3,6 of the week-long period
Flynn-Evans et al. (2018)	United States	44	30.8 ± 7.1	40:4	Early: 7.4 ± 0.9 Midday: 7.9 ± 1.1 Rest Day: 8.5 ± 0.9	Late: 7.8 ± 1.4	–	Unk; Duration Various times through day shifts depending on schedule; iPod	Early: 257 ± 70 Midday: 261 ± 62 Late: 266 ± 64 Rest: 249 ± 56	Early: 4.4 ± 5.4 Midday: 4.7 ± 5.1 Late: 4.7 ± 4.5 Rest: 4 ± 4.5	Sleep diary and actigraphy was daily PVT on work days: upon waking, before bed, top of descent and post-flight. Non-work day: 1–2 h of waking, 8–9 h after waking and 1–2 h before sleeping

Table 1 (continued)

References	Country	Sample size	Age range	Gender Ratio M:F	Sleep between day shifts (hours)	Sleep between night shifts (hours)	KSS ^a	PVT duration; PVT timing; PVT model	PVT ^b Mean RT ^c (ms)	PVT mean lapses (≥ 500 ms)	PVT Application
Shattuck and Mat-sangas (2016)	United States	77	25.3 ± 3.17	63: 14	6.88 ± 0.894, <i>p</i> < 0.05	–	–	3-min; Unk; Frequency Motion logger sleep watch	392 ± 111, <i>p</i> < 0.05	13.5 ± 12.2, <i>p</i> < 0.05	Actigraphy and sleep diary completed daily Frequency of PVT not reported
Wilson et al. (2019)	United States	22	20–60	1:10	7.1 ± 1.0	–	–	10-min; Start, middle and end of shift; PVT Model Unk	Significant increase in mean PVT RT as night shift progresses. PVT RT remains similar across the day shift	Day shift decreases Night shift threefold	Actigraphy and sleep diaries daily PVT, KSS at beginning, middle and end of 4 × duty shifts
Commencement and cessation of shift											
Behrens et al. (2019)	Germany	74	28–47	0:74	–	–	–	3-min; Between 0600 and 0700 post-night and 1300–1500 after a day; Hand-held device	Mean RT for whole study: 241.74 ± 23.49 <i>p</i> < 0.0001	Lapses > 355 ms 1st Day: 9.01. <i>p</i> = 0.0728 2nd Day: 9.19. <i>p</i> = 0.0003 False Starts < 100 ms 1st day: 3.25, <i>p</i> = 0.3259 2nd day: 2.91 ms, <i>p</i> = 0.4290	The PVT was administered at the study centre immediately after each day/night shift

Table 1 (continued)

References	Country	Sample size	Age range	Gender Ratio M:F	Sleep between day shifts (hours)	Sleep between night shifts (hours)	KSS ^a	PVT duration; PVT timing; PVT model	PVT ^b Mean RT ^c (ms)	PVT mean lapses (≥ 500 ms)	PVT Application
Bonnefond et al. (2006)	Finland	49	25–58	–	–	–	Night/ evening shift difference between start and finish results increased. Morning shift remained similar. Age had no significant effect on sleepiness ratings, but it did on work shift/ time of day $p < 0.0001$	10-min; Start and end of 1st and 3rd days; Pocket PC Computer	Mean RT showed a slight decrease between the start and end of shift $p < 0.0003$	Day and evening shifts remained similar, however night shifts lapses increased more than threefold in middle and senior age groups and doubled in young age group $p < 0.01$	Actigraphy recording was daily KSS was every 2 h during shift PVT start and end of shift
Ferguson et al. (2011)	Australia	35	42.6 \pm 9.7	–	–	–	$p < 0.0001$	5-min; Start and end of each work shift; palm PVT	Mean start day vs mean end day RT was about the same within 0.1 ms whereas night differences were 0.2 ms + different $p > 0.05$	Fastest response time was from the 7 \times 4 roster category and the worst was the 4 \times 4	Actigraphy and sleep/work diaries daily PVT at start and end of each shift
Harma et al. (2006)	Finland	113 /24	24–61	137:0	–	–	–	10-min; Various times on night shifts; PVT Model Unk ^f	RT was dependent on shift, but not age (namely increased RT on night shift; $p < 0.0064$) There was also an interaction effect showing a decrease in median reaction time at the end of the night shift among the older age groups ($p < 0.0001$)	PVT lapses were dependent on shift, but not age. Night shift was more prominent for lapses	Daily actigraphy and sleep diary PVT and KSS competed at start and end of each shift

Table 1 (continued)

References	Country	Sample size	Age range	Gender Ratio M:F	Sleep between day shifts (hours)	Sleep between night shifts (hours)	KSS ^a	PVT duration; PVT timing; PVT model	PVT ^b Mean RT ^c (ms)	PVT mean lapses (≥ 500 ms)	PVT Application	
Ruggiero et al. (2012)	United States	16	47.6 ± 8.1	0:16	–	Night shift nurses slept and napped significantly more than day shift nurses	–	10-min; Before leaving and returning home from work; PVT-192	Day 1, large difference between start and end of shift ($p=0.03$) Day and night workers reported increased sleepiness after day 2 on KSS	–	Continuous actigraphy and KSS was completed at home before and after each shift for 3 × 24 h continuous periods	
Thompson et al. (2016)	United States	38	31 ± 7.5 years	0:38	–	–	–	5-min; Unk; Frequency PC laptop and optical gaming mouse	Pre-Shift: 259.7 ms Post-Shift: 275.3 ms $p < 0.01$	Pre: 0.25 Post: 0.72 $p < 0.05$	PVT was conducted before and after the work period on 3 × 12-h shifts	
Other varying times												
Basner et al. (2017)	United States	137 Interns 87 Residents	27.6 ± 2.3 29.1 ± 2.7	50.4% 57.5%	6.93 ± 0.08 (95% CI)	–	–	3-min; Between 0600–0900 after a night; Laptop	Significantly slower in interns during mornings after periods of on-call ($p < 0.001$)	–	Daily actigraphy and sleep diary PVT and KSS completed daily, in the morning between 0600–0900	
Ganesan et al. (2019)	Australia	52	22–64 33.8 ± 9.7	14:38	5.83 ± 0.92 h	5.74 ± 1.3 h	Day shift remained similar Night shift went from 4–6.5 between first and final nights	5-min; End of last day and all nights; PVT Model Unk	PVT RT on day shifts improved slightly. PVT RT on night shifts increased by about 50 ms	Night shift number of lapses doubled towards end of shift	KSS and PVT were examined on day shift and first and subsequent night shift (3/4/5th) Sleep diary and actigraphy daily	

Table 1 (continued)

References	Country	Sample size	Age range	Gender Ratio M:F	Sleep between day shifts (hours)	Sleep between night shifts (hours)	KSS ^a	PVT duration; PVT timing; PVT model	PVT ^b Mean RT ^c (ms)	PVT mean lapses (≥ 500 ms)	PVT Application
Harris et al. (2010)	Norway	19	28–60	13:06	–	–	–	10-min; End of day/night 1, 7/8 and 14 of shift cycle; Hand-held computer	No statistical difference between 14 days nor swing shift, however on the first night of 14 nights, the PVT scores were significantly higher ($p < 0.05$)	–	Saliva was collected at multiple stages during the shift PVT completed on day/night 1, 7 (or 8 when on swing shift) and 14 of the shift cycle
Vanttola et al. (2019)	Finland	31	41 \pm 8	77% men	–	–	Morning shifts remain the same Night shift worsens by 1–2 points	10-min; Before and after shift and between 1000–1330 on shift; HP ipaq 514 mobile phone	–	Pre-morning shift: $0.7 \pm 1.3, p < 0.27$ Post-morning shift: $0.5 \pm 1.3, p < 0.19$ Pre-night shift: $0.5 \pm 0.6, p < 0.04$ Post-Night shift: $1.0 \pm 4.4, p < 0.61$	Actigraphy and sleep diary completed daily KSS and 10-min PVT completed on 6 pre-selected days. 2 on RDOs ^d , 2 during morning shifts and 2 on evening shifts. Each PVT completed within first and last 1–2 h of shift
Waggoner et al. (2012)	United States	–	–	–	–	–	Almost doubled post-series of night shifts $p < 0.005$	10-min; Between 0700–0930 after the shift rotation; Smart phone application	–	More than doubled the number of lapses post-night shifts. $p < 0.001$	At completion of the shift cycle, a 2-h testing session was completed including the KSS and PVT, which was then compared to the testing session completed after 3 RDO

KSS Karolinska sleepiness scale^a, PVT psychomotor vigilance test^b, mRT mean reaction time^c, measured in milliseconds, Unk Unknown^d, RDO rostered day off^e

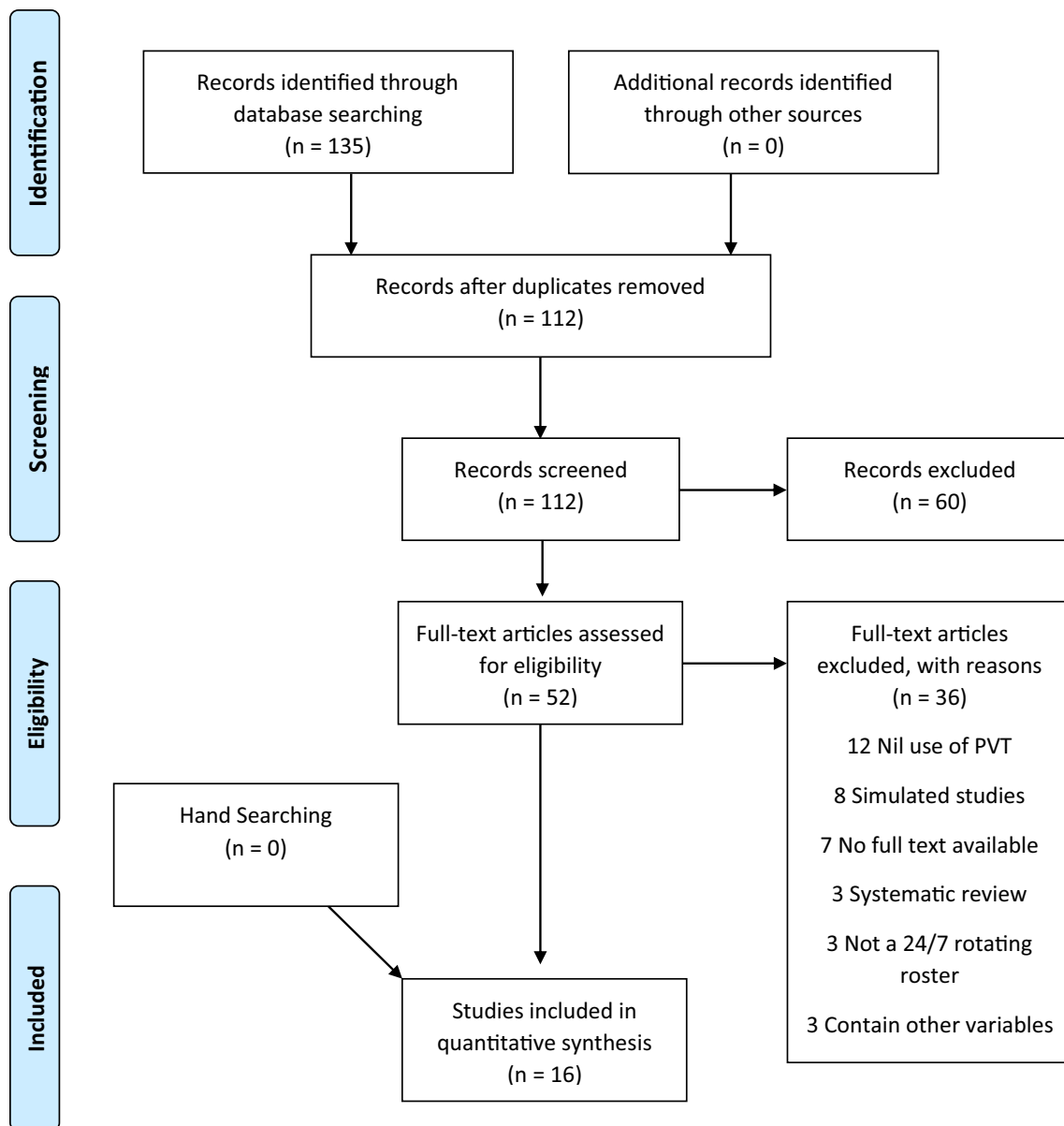


Fig. 1 PRISMA flow chart

methodological quality. None of the 16 studies used blinding which consequently may have introduced an increased risk of detection bias from outcome assessors. Three of the included studies did not have a participation rate of greater than 50% of eligible individuals and thus their results may not be generalisable to the wider population. Five of the included studies had no sample size justification and four had a loss to follow-up after baseline of greater than 20%, namely due to participants being no longer eligible to participate or unable to be re-contacted. All studies used similar ANOVA based statistical analyses approaches to examine data.

GRADE assessment

As per the GRADE recommendations, the initial body of evidence in this review is ‘low’ quality due to the majority of included studies being observational (Atkins et al. 2004). The quality of the overall data is low due to lack of blinding allocation process. A number of studies had limited follow-up assessments and/or experienced a large dropout of participants due to a variety of reported reasons (e.g., researchers could no longer get in contact with participants, participants no longer wished to remain a part of the study, complex methodologies that required extensive and timely participant interactions). Although three-quarters of the included studies

had consistent findings in relation to PVT performance, four studies had mixed results (Bonfond et al. 2006; Ganesan et al. 2019; Waggoner et al. 2012; Vanttola et al. 2019). Therefore, the overall body of evidence in this review was determined to be of ‘very low’ quality.

Narrative synthesis

The PVT was administered to participants in each study at varying timepoint intervals. The positioning of these timepoints were based on the roster design being examined in the study or the potential for the timepoint to identify key changes in workers fatigue levels, hypothesized by study authors. In combination with the PVT, other subjective and objective fatigue-related measures were applied. The KSS was the most popular subjective measure used in the included studies and this was frequently coupled with the PVT. In addition to the PVT, the most commonly used objective measure was actigraphy, which was used to track movement and sleep patterns. Actigraphy was often used alongside sleep and work diaries, which provided subjective information on sleep and work hours.

There appeared to be a relationship between objective and subjective fatigue measures in several studies which included both of these types of measures. For example, Bjorvatn et al. (Bjorvatn et al. 2006) found that if the mean reaction time (mRT) in the PVT increased towards the end of a shift, then KSS levels also showed an increase and this finding was consistent across several other studies included in this review (Bonfond et al. 2006; Basner and Dinges 2011; Bjorvatn et al. 2006; Harma et al. 2006; Basner et al. 2017; Waggoner et al. 2012; Vanttola et al. 2019). Conversely, Wilson et al. found poor congruence between PVT performance and the KSS among nurses (Wilson et al. 2019). Furthermore, a number of included studies investigated relationships between fatigue and actigraphy, with findings indicating that shorter sleep durations were related to increased objective fatigue levels (e.g., slower PVT response speed) (Bonfond et al. 2006; Flynn-Evans et al. 2018; Ferguson et al. 2011; Basner et al. 2017). Finally, Shattuck et al. found a direct correlation between decreased mRT and worsened mood (Shattuck and Matsangas 2016).

To facilitate the synthesis of the literature, the below section has been split into three categories based on the application of the PVT in the included studies (i.e., ‘Multiple instances per shift’, ‘Commencement and cessation of shift’ and ‘Other varying times’; Table 1).

Multiple instances per shift

The PVT was applied at multiple times (i.e., ≥ 3 administrations) throughout each shift in five of the included studies (Flynn-Evans et al. 2018; Shattuck and Matsangas

2016; Anderson et al. 2012; Bjorvatn et al. 2006; Wilson et al. 2019). Anderson et al. administered the PVT every 6 h through a 24–30-h shift for resident interns, which showed a slow and steady decline in cognitive attention ($p < 0.05$) (Anderson et al. 2012). Similarly, Flynn-Evans et al. tested pilots at key stages (start of shift, end of each ascent, prior to each descent and end of shift) throughout their shift, but also before and after major sleep periods, yielding a similar decline in cognitive attention measured on the PVT ($p < 0.05$) (Flynn-Evans et al. 2018). Research by Bjorvatn et al. in oil rig workers and Wilson et al., in hospital-based nurses, however, had set times the PVT was to be completed throughout their 12-h shift (Bjorvatn et al. 2006; Wilson et al. 2019). This typically occurred at 0000, 0300 and 0600 on night shifts or 1200, 1500 and 1800 on day shifts (Bjorvatn et al. 2006; Wilson et al. 2019). These two studies observed comparable decreases in PVT mRT on the first night shift. However, Wilson and colleagues found reaction time performance slowly steadied the longer a nurse was on consecutive night shifts ($p < 0.0001$) (Wilson et al. 2019). Furthermore, Wilson et al. found mean lapses on the PVT tripled on night shifts compared to days ($p < 0.006$) (Wilson et al. 2019). Conversely, Bjorvatn et al. found lapse data in oil rig workers remained consistent across the shifts (Bjorvatn et al. 2006). Overall, the use of the PVT at multiple instances throughout a shift show a decline in cognitive attention.

Commencement and cessation of shift

The PVT was administered to participants at the commencement and cessation of each shift in six of the studies in the included analysis (Bonfond et al. 2006; Ruggiero et al. 2012; Behrens et al. 2019; Ferguson et al. 2011; Thompson et al. 2016; Harma et al. 2006). Behrens et al., Ferguson et al., Ruggiero et al. and Thompson et al. all examined rosters that progressed forward (i.e., days to nights to rostered days off) and included shifts that were 12–13 h in length (Ruggiero et al. 2012; Behrens et al. 2019; Ferguson et al. 2011; Thompson et al. 2016). The studies each reported significantly lower mRT on the PVT from the beginning to the end of the first shift. After this first shift; however, mRT typically remained a constant for the rest of the roster duration, suggesting participants were accustomed to the test process and results did not decrease further (Ruggiero et al. 2012; Behrens et al. 2019; Ferguson et al. 2011; Thompson et al. 2016). This was consistent with the other studies which examined PVT performance at *multiple instances across a shift*.

Bonfond et al. and Harma et al. both found correlations between ages and shift type on fatigue assessed using the PVT (Bonfond et al. 2006; Harma et al. 2006). The middle and older age groups performed significantly worse

(threefold) in response time and lapses on night shifts than younger participants ($p < 0.01$) (Bonnefond et al. 2006). This was significant among maintenance workers performing a backwards rotating roster (mornings to nights to evenings) ($p < 0.0003$) (Bonnefond et al. 2006). Conversely, Harma et al. found a weaker correlation between age and PVT performance, however in this study, participants performed a forward-facing roster which resulted in decreased performance across all age groups in the later night shifts ($p < 0.0064$) (Harma et al. 2006).

Other varying times

In the remaining five studies, the timing of the PVT was more targeted towards specific rosters or industry designs (Ganesan et al. 2019; Vanttola et al. 2019; Basner et al. 2011; Harris et al. 2010; Waggoner et al. 2012). Basner et al. examined sleep patterns and sleep inertia of physician interns and residents (Basner et al. 2017). In this study, the PVT was administered to these workers between 0600 and 0900 each morning after either a night on-call or at home (Basner et al. 2017). The findings from this study revealed reaction time was significantly slower during the mornings after being on-call ($p < 0.001$) compared to regular shifts (Basner et al. 2017).

Waggoner et al., administered the PVT in police officers to examine vigilant attention at the end of their rostered work ‘week’, where rotations always finished with three consecutive 12-h nights (Waggoner et al. 2012). In this study, police were tested on their driving performance, cognition and vigilant attention at a research facility after the third and final 12-h night shift and these results were compared to the same test battery completed three days later after rostered days off (Waggoner et al. 2012). The KSS results showed officers reported double the level of subjective sleepiness post-night shift ($p < 0.005$) as compared to when well rested (Waggoner et al. 2012). Furthermore, the number of performance lapses on the PVT more than doubled ($p < 0.001$) post-night shift as compared to when rested (Waggoner et al. 2012).

The remaining three studies tested participants typically on their first day and subsequent (i.e., 3rd, 5th and 7th shift) shifts or twice during random allocation of shifts (Vanttola et al. 2019; Harris et al. 2010; Ganesan et al. 2019). Ganesan et al., who looked at healthcare workers, found day shift mRT on the PVT slowly increased on consecutive days of work, whereas mRT measured across consecutive night shifts was found to increase by 50 ms in comparison (Ganesan et al. 2019). These results indicate a faster decrease in

healthcare workers vigilant attention when working nights ($p < 0.001$) (Ganesan et al. 2019). Similarly, when examining performance lapses, workers performed the worst on night duty and the number of lapses almost doubled at the end of the roster rotation ($p < 0.001$) (Ganesan et al. 2019). Harris et al., examined PVT performance among offshore oil rig personnel who were working either a 12-h day shift, 12-h night shift or a swing shift (7-night shifts followed by 7-day shifts, each 12-h in duration) schedule (Harris et al. 2010). The PVT was administered to personnel on their first, seventh (or eighth if on swing shift) and fourteenth shift (Harris et al. 2010). There was no statistical difference in mRT between either of the 14-h day or swing shift schedule. Interestingly, the PVT results showed a significant increase in mRT on the first night on the 14-h night schedule, however this promptly normalised on subsequent tests ($p < 0.05$) (Harris et al. 2010).

Discussion

The aim of this systematic review was to evaluate participants PVT performance in response to 24/7 shift rosters performed in different occupational settings under naturalistic conditions. It is clear the PVT has been applied through varying study protocols, which has yielded consistently significant results. There is a reasonable quantity of objective data that demonstrates decreased alertness and increased sleepiness throughout the majority of included studies.

Among the included studies, half used the 10-min PVT and the most common PVT modality was the palm pilot. Despite this variation in the duration and modality of PVT applied across the included studies, general conclusions can still be drawn from the cumulative results. For instance, there is evidence to support the notion that the varying length PVTs (i.e., 3-, 5- and 10-min variations) produce comparable results and one version or modality does not necessarily produce stronger or more accurate results (Basner et al. 2011; Basner and Dinges 2011; Arsintescu et al. 2017). It is, however, pertinent to note that in some prospective and/or in-field studies, the 10-min PVT may be impractical due to time, budget and/or logistical constraints. Thus, a 3- or 5-min PVT may be more appropriate and yield similar results to a 10-min version. Further, in the majority of studies that included multiple fatigue measures, there appeared to be a positive relationship between subjective and objective fatigue measures.

Multiple instances per shift

Although the studies by Anderson et al. and Flynn-Evans et al. both have similar study designs in the application of the PVT (i.e., administered 6-hourly or > 3 times per shift), they interestingly had quite different results that may be explained through differences in roster design (Flynn-Evans et al. 2018; Anderson et al. 2012). The medical interns examined by Anderson et al. had a mRT of 284.6 ms at the start of the shift, which slowly increased by 100 ms by the end of their duty period ($p < 0.05$), up to 30 h later (Anderson et al. 2012). When compared with these medical interns (Anderson et al. 2012), the pilots examined by Flynn-Evans et al. had a similar mRT at the start of their shift ($p < 0.05$) (Flynn-Evans et al. 2018). However, the duty period examined among pilots by Flynn-Evans et al. was substantially shorter than the shift duration examined in interns by Anderson and colleagues, which may have explained the increase in mRT (i.e., slower to react to the stimulus) at the cessation of shift due to prolonged working hours (Anderson et al. 2012; Flynn-Evans et al. 2018). Both studies had a moderate number of PVT performance lapses in the early phases of their respective shifts, possibly explained by participant learning error; however, the lapses tripled for the medical interns at the end of their duty period ($p < 0.05$) (Flynn-Evans et al. 2018; Anderson et al. 2012). Thus, this finding could indicate interns are at significant risk of impaired performance towards the end of their shift due to extended duty hours. There is limited research that has examined how this level of impairment on PVT performance relates to real-life/real-world medical errors. This is a research area for further evaluation in the future.

Wilson et al. and Bjorvatn et al. both demonstrated a significant decrease through the first shift (typically nights) in PVT mRT and then slower decreases on consecutive shifts afterwards ($p < 0.05$) (Bjorvatn et al. 2006; Wilson et al. 2019). This is indicative of circadian rhythm attempts at adaption (Wilson et al. 2019). However, as demonstrated in the study by Wilson et al., this cannot be maintained for extended periods of time (i.e., > 4 shifts) (Wilson et al. 2019). Finally, both Wilson et al. and Bjorvatn et al. found key PVT performance challenges at the change of shift, including changes between day-to-night and night-to-day (i.e., swing shift) ($p < 0.05$) (Bjorvatn et al. 2006; Wilson et al. 2019).

The regular use of the PVT in the above-mentioned studies has provided a clear demonstration of cognitive vigilant

attention changes throughout shift work. This is, however, not without research limitations. Components of data collection had to be removed as some participants undertook the tests too early/late which could have skewed the data (Flynn-Evans et al. 2018). Other studies reported missing individual assessments or full-days of data collection, requiring the application of missing data imputation methods (Bjorvatn et al. 2006). This resulted in incomplete data sets and thus the findings may not necessarily be a true representation of the sample. Future research could benefit by collecting data from participants at a central location, for example a research room in a hospital, to minimise incomplete data sets.

Commencement and cessation of shift

Analysis of studies examining PVT performance at the start and end of a shift was impeded by 33% of raw data not being reported through the studies, but descriptive techniques had been used instead. PVT performance examined at the beginning and end of each shift showed a decrease in cognitive vigilant attention, as expected; however, this was more extreme on night shifts (Ruggiero et al. 2012; Behrens et al. 2019; Ferguson et al. 2011; Thompson et al. 2016). Thompson et al. also supports this notion through the testing of nurses' vigilance after three 12-h shifts within a four-day time period (Thompson et al. 2016). In this study, cumulative fatigue resulted in a 21.2% increase in the standard deviation of mRT in this cohort of nurses (Thompson et al. 2016).

Two studies found that work hours are seldom the primary factor contributing to fatigue (Ruggiero et al. 2012; Ferguson et al. 2011). This adds to the growing literature that reduced sleep quantity is a major determinant of mRT later in the shift (Ferguson et al. 2011). For instance, Ferguson et al. reported that miners performing the PVT with less than 6 h of sleep in the preceding 24 h had slower response times than those who had 7+ h of sleep (Ferguson et al. 2011). Laboratory studies; however, indicated PVT performance could be maintained in the short term (i.e., 1–2 days) with 6 h of sleep (Ferguson et al. 2011). This finding is consistent with research by Ruggiero et al. who found sufficient rest (7+ hours sleep) is required for improved PVT performance in critical care registered nurses (Ruggiero et al. 2012). This can, and often is, supplemented by shift workers undertaking sanctioned napping (Basner et al. 2011, 2017; Harris et al. 2010; Flynn-Evans et al. 2018; Miley et al. 2016; Ruggiero et al. 2012;

Shattuck and Matsangas 2016; Basner and Dinges 2011; Arsintescu et al. 2017; Moher et al. 2009; Sayers 2008; Veritas Health Innovation 2017; National institute of Health 2014; Atkins et al. 2004; Anderson et al. 2012; Bjorvatn et al. 2006; Wilson et al. 2019; Tanaka et al. 2011; Behrens et al. 2019; Ferguson et al. 2011; Thompson et al. 2016; Harma et al. 2006; Waggoner et al. 2012; Vanttola et al. 2019; Wyatt et al. 1999; Richter et al. 2021) during the night shift (Ruggiero et al. 2012; Wilson et al. 2019), which has been found to improve reaction times and reduce the effects of sleepiness (Hilditch et al. 2016).

The time of day that the PVT test is completed and the number of hours of wakefulness are also known to have an impact on reaction time (Wyatt et al. 1999). For instance, laboratory-based studies have found that when the PVT is performed later in the day or during the night in workers, this can lead to a mismatch between the clock time and the natural biological clock of participants, resulting in significantly decreased neurobehavioral outputs (e.g., slower PVT mRT) (Wyatt et al. 1999). Indeed, several of the included studies found that PVT reaction time was remarkably slower at times not congruent with a natural biological sleep/wake cycle, such as on night shifts, very early mornings or late afternoons (Bonfond et al. 2006; Flynn-Evans et al. 2018; Anderson et al. 2012; Bjorvatn et al. 2006). This is undoubtedly a significant finding, however few of these studies also measured each participant's individual sleep patterns and prior hours of wakefulness (Anderson et al. 2012; Bonfond et al. 2006; Flynn-Evans et al. 2018). Thus, like prior laboratory-based studies (Wyatt et al. 1999), it is important future field studies take time of day as well as the number of hours of awakening into consideration when examining PVT performance in naturalistic shift work settings.

Although the effects of age on PVT performance was not a focus of this review, it was a noteworthy factor in the included studies. For instance, Bonfond et al. and Harma et al. both found that alertness and fatigue levels among older adults were affected the least on single night shifts and rapidly forward rotating rosters as compared to multiple successive nights (Bonfond et al. 2006; Harma et al. 2006). This not only promoted employment longevity, but also decreased recovery time to transition back

to a normal circadian rhythm (Bonfond et al. 2006). Notwithstanding this, PVT mRT and mean lapses were still markedly higher in older adults when compared with younger and middle-aged adults (Bonfond et al. 2006). Together, these studies suggest that to reduce fatigue-related impairment, older shift workers should have rapidly forward progressing rosters, increased recovery time, increased flexibility and lower frequencies of night duty (Harma et al. 2006).

Although the data appear to have reasonable academic rigour, the limitations of these studies include single institution research (Behrens et al. 2019), small to medium sample sizes (Ruggiero et al. 2012), low response rates (Ruggiero et al. 2012) and high participant drop-out in some studies (Bjorvatn et al. 2006), which may impede the generalisability of findings. However, in combination, these studies contribute to a growing evidence base that indicates fatigue levels can be maintained through sanctioned napping and rapidly forward rotating rosters. Further, the sample size in each age cohort was small ($n = 13$ to 19), which limits the generalisation of the age-related findings to the wider population (Bonfond et al. 2006). Finally, Harma et al. lowered the chance of behavioural abnormalities present in their study results by designing a 2-week roster intervention (Harma et al. 2006). In addition, alcohol and sleeping medicines were restricted and a stringent research protocol was provided (Harma et al. 2006). However, this may not be reflective of usual behaviours of shift workers in naturalistic settings who may have access to alcohol and sleeping medications (Richter et al. 2021).

Other varying times

Analysis for this domain was challenged by a 60% lack of raw data; however, descriptive statistics have been used. Basner et al. concluded that post on-call fatigue levels could cause clinical concerns due to significantly lower levels in cognitive vigilant attention (Basner et al. 2017). It has been reported fewer than 7 h of sleep in a 24-h period can increase errors and create safety challenges with the therapeutic patient–clinician relationship (Basner et al. 2017). Waggoner et al. agreed with Basner's (Basner et al. 2017) study in that extended periods of duty can decrease performance,

alertness and impair clinical and personal decision making (i.e., driving home) (Waggoner et al. 2012).

Ganesan et al. found that subjective fatigue was perceived to be lowest on the first night shift in a series of consecutive night shifts; however, objectively measured PVT performance indicated impairment was similar across all subsequent shifts (Ganesan et al. 2019). This finding was in contrast to other studies whose results suggested the contrary (Bonfond et al. 2006; Ganesan et al. 2019; Waggoner et al. 2012; Vanttola et al. 2019). Furthermore, Harris et al. found circadian adaptation to night duty occurred within a week (when working greater than 7-night shifts); however, recovery back to normality took longer with no adverse outcomes to PVT mRT (Harris et al. 2010).

Overall summary

The majority of included articles suggest subjective fatigue measures are consistent with objective measures in response to different rostering schedules (Basner et al. 2017; Waggoner et al. 2012). There was a consistent correlation between increased mRT measured on the PVT and increased KSS levels, particularly during evening/night shifts (Ganesan et al. 2019; Basner et al. 2011; Waggoner et al. 2012; Vanttola et al. 2019). Harris et al. (Harris et al. 2010) also found prolonged number of night shifts resulted in longer re-adaptation times in participants cortisol levels, which is a biological marker found in prior simulator-based research to be associated with increased levels of subjective fatigue in shift workers (Wolkow et al. 2016). Finally, Bjorvatn et al. (Bjorvatn et al. 2006) found objective sleep duration measured using actigraphy was similar to subjective sleep, which is a finding that was consistent with other studies included in this review (Bonfond et al. 2006; Flynn-Evans et al. 2018; Ferguson et al. 2011; Harma et al. 2006). However, in a small number of included studies there are divergent findings in this area (Ganesan et al. 2019; Vanttola et al. 2019). Future research should be directed towards exploring the potential human and experimental factors that may explain these deviations in the literature, especially in naturalistic settings.

Limitations

Methodological limitations in this review include the inclusion/exclusion search criteria containing only articles in the English language, which unfortunately could not be avoided as none of the research team were bilingual. The requirement for articles to be peer-reviewed increases the academic rigor of the review; however, it also means some primary, contemporary research is excluded. The influence of timed caffeine or food consumption on fatigue-related impairment would be an interesting comparison; however, these factors were not within the scope of this review and could be considered in future evaluations. Further, evaluating simulated studies against the real-world research examined in the current review could also evoke further investigation into study design or roster review.

Conclusion

Through this review, impairments in PVT performance were consistently demonstrated in response to varying roster patterns and shift types (night and day shifts), indicating increased fatigue levels and reduced alertness. There is, however, a distinct lack of evidence examining vigilant attention changes during sustained periods of work and on-call work (beyond 30+ h). Healthcare providers, defence personnel, pilots and communication specialists (translators and interpreters) spend extended periods on-call and could experience more pronounced changes in fatigue related impairments in performance. Future research addressing this area is required for improved workplace safety, productivity and performance.

Appendix 1

See Table 2.

Table 2 Summary of studies

References	Study design	Country	Occupational group	Shift type	Fatigue-related measures	Study aims and conclusions	NHLBI rating score (range 0–14)	Relationships between fatigue measures
Multiple instances per shift								
Anderson et al. (2012)	Prospective, observational	United States	Resident Physicians (Medical and Cardiac Intensive Care Units)	24–30 h work shift. Shift cycle starts every 3 days, (alternating between 24 and 30 h shifts)	PVT Sleep/work diaries	Residents within a local hospital intensive care unit were assessed on their neurobehavioral performance during extended duration work shifts (EDWS). It was shown that over 6 shifts, their PVT results significantly decreased and participants were also experiencing chronic sleep deficiency	11	PVT and sleep diaries are congruent
Bjorvatn et al. (2006)	Randomised, Placebo-controlled crossover field study	Norway (North Sea)	Oil rig workers	7 × night shifts (1830–0630) and 7 × day shifts (0630–1830)	PVT KSS Actigraphy	Adaption and re-adaption between varying shift work schedules were the key focus of this study, particularly the transition from nights to days. The workers reported the return to day work after seven-night shifts created a clear increase in subjective sleepiness and worsening of sleep parameters. It was also reported that participants performance on day shifts were worse than night shifts, and there was a clear reduction in total sleep time between days. However, all subjective and objective measures slowly improved as days progressed through each set of days/night shifts	12	KSS and PVT improved across days
Flynn-Evans et al. (2018)	Prospective observational	United States	Airline pilots	5 × mid-morning shifts, 5 × early morning shifts, 5 × high workload midday shifts followed by 5 × late shifts. Each block separated by 3–4 days off	PVT Sleep diaries Actigraphy Urine collection	This study examined the effect of circadian misalignment and sleep disruption on people working in the transportation industry. It was found that a large contributor to sleep deficiency in the modern society could be associated with non-traditional shifts and one's ability to obtain sufficient sleep	12	Actigraphy and Sleep diary are congruent

Table 2 (continued)

References	Study design	Country	Occupational group	Shift type	Fatigue-related measures	Study aims and conclusions	NHLBI rating score (range 0–14)	Relationships between fatigue measures
Shattuck-Haghighi and Yazdi (2015)	Prospective observational	United States	Volunteer navy sailors	5 h on, 10 h off work schedule	PVT Epworth Sleepiness Scale Profile of mood state	Crew members that worked in the nuclear reactor department who worked the 5 h on, 10 h off watching schedule were eligible to participate in this study. It was found by day 2, crew were already in significant sleep debt and reported lower moods, more PVT performance lapses and dissatisfaction with the work schedule. It was also found that crew experienced periodic bouts of sustained wakefulness and circadian rhythm misalignment	12	Slower PVT times correlated to lower moods and decreased levels of sleep hygiene
Wilson et al. (2019)	Prospective observational	United States	Hospital based nurses	6 × 12-h day shifts (0700–1900) or night shifts (1900–0700) per fortnight	PVT KSS	Full-time nurses at a regional centre hospital were asked to perform a fatigue test battery at the beginning, middle and end of four out of six shifts, in addition to wearing activity monitor wrist watches. Objective data remained steady for day shift workers, however gradually declined through testing phases for the night shift workers. Night shift workers also displayed greater performance variability in PVT and reported higher subjective sleepiness levels on KSS	12	KSS and PVT were not congruent
Commencement and cessation of shift								
Behrens et al. (2019)	Prospective observational	Germany	Female registered nurses	2 × day shifts (0600–1400), 3 × night shifts (2100–0600)	PVT	This study aimed to assess the psychomotor vigilance of female nurses during both day and night shifts. It was found that reaction times were increased towards the end of night shifts but no significant difference in false starts between days and nights	12	N/A

Table 2 (continued)

References	Study design	Country	Occupational group	Shift type	Fatigue-related measures	Study aims and conclusions	NHLBI rating score (range 0–14)	Relationships between fatigue measures
Bonnefond et al. (2007)	Cross-sectional field study	Finland	Aircraft maintenance workers	3 × day shifts (0700–1500), 3 night shifts (2300–0700), 3 × late shifts (1500–2300) and 2 × days off	Psychomotor Vigilance Test KSS Actigraphy	This study explored three different age cohorts of workers, comparing their reaction time and performance lapses on the PVT. While there were no significant differences between days and late shifts between the age groups in PVT reaction times or lapses, there was a trend towards a decrease in the older population cohort	12	Age had no effects on KSS but it did on actigraphy PVT lapses and KSS scores both increased throughout ones shift
Ferguson et al. (2011)	Prospective observational	Australia	Miners	12-h day shifts (0600–1800) and night shifts 4 × day shifts, 4 × rest days, 4 × night shifts, 4 × rest days or 7 × day shifts, 4 × rest days, 7 × night shifts, 4 × rest days or 7 × day shifts, 7 × night shifts, 7 × rest days	PVT Actigraphy Sleep/work diaries	Workers in the mining sector work long and progressive hours. This study explored how response times was associated with performance and alertness before and after individual shifts. As expected, it was found responses were markedly slower towards the end of night shifts compared to day shifts, but responses were also sensitive to immediate sleep history, with lower amounts of recent sleep resulting in slower response times	12	PVT and Actigraphy are congruent
Harma et al. (2006)	Controlled intervention study	Finland	Maintenance Unit of a major airline company: Finnair	Backwards rotating 3 shift roster (EEE, RR, MMM, RR, NNN, RR) or forward rotating 3 shift roster (MEN, RR, MEN, RR, MEN, RR) ^a	PVT KSS Actigraphy	Maintenance workers of this Finnish airline were asked to try two different shift work rosters to determine which would provide the best employer and employee benefits; either backward (old roster system) or forward rotating roster (new). The new roster improved main sleep length after nights and overall sleep complaints decreased. Workers subjectively reported increased benefits of health, well-being, social/ family life and alertness on the new roster system. Overall, it was found that the forward rotating roster had increased benefits when compared to the old backwards rotating roster	12	KSS and PVT are congruent

Table 2 (continued)

References	Study design	Country	Occupational group	Shift type	Fatigue-related measures	Study aims and conclusions	NHLBI rating score (range 0–14)	Relationships between fatigue measures
Ruggiero et al. (2012)	Prospective observational	United States	Critical care registered nurses	12 or 13 h day shifts (0700–1930 or 0600–1830) and 12.5 h night shifts (1900–0730) 3 shifts in a row followed by RDO	PVT Stanford Sleepiness Scale Actigraphy	Evaluation of sleep and psychomotor vigilance in female shift workers was the focus in this study. The researchers found that nurses working night shifts napped more frequently while also having greater changes to the length of their main sleep during the day. Night shift nurses also had increased fatigue at the culmination of their shift compared to day shift nurses	13	Stanford Sleepiness Scale differed from PVT
Thompson et al. (2016)	Repeated measures study	United States	Female, Hospital based Registered Nurses, Nurses Aids or licensed vocational nurses	12-h day/nights	PVT Balance and neuro-muscular control	This study aimed to explore how varying rosters would impact female nurse's reaction time and balance performance. The work period found impairments in reaction time, increase errors/lapses, and poor balance performance. Further, it was found that reaction time and balance tests may be sensitive to fatigue for nurses	12	Balance Test and PVT were congruent, i.e., increased reaction time resulted in poorer balance
Other varying times								
Basner et al. (2017)	Randomised Controlled Clinical Trial	United States	Interns and resident physicians	30 h shift (0700 until 1300 the next day)	PVT KSS Actigraphy	This study investigated whether protracted sleep periods (3–5 h) during extended periods of work resulted in increased alertness or psychomotor vigilance. It was found that extended overnight shifts increase the likelihood of chronic sleep restrictions in interns. Further, alertness levels were significantly lower on mornings after on-call	11	PVT, KSS and Actigraphy were all congruent
Ganesan et al. (2019)	Prospective observational	Melbourne, Australia	Registered nurses and doctors in an intensive care unit at a hospital	Irregular combinations of day (0700–1530), evening (1300–2130) and night (2100–0730)	PVT Actigraphy KSS	This study looked at the sleepiness, alertness and performance during day and nights shifts of intensive care unit nurses and doctors. All employees have rapidly rotating rosters with the exception of occasional continuous night shifts, and 1 to 5 consecutive shifts. It was found that a lack of circadian adaptation resulted in lower performance and alertness on night shifts	11	KSS and PVT were congruent

Table 2 (continued)

References	Study design	Country	Occupational group	Shift type	Fatigue-related measures	Study aims and conclusions	NHLBI rating score (range 0–14)	Relationships between fatigue measures
Harris et al. (2010)	Prospective observational	Norway	Oil rig workers	12-h day/night shifts Fixed shift schedule (14 × 12 hr day or night shifts) or swing shift schedule (7 × 12-h night shifts followed by 7 × 12-h day shifts)	KSS Salivary cortisol levels	This study in oil rig workers was to explore whether there would be a negative effect on overall health, cortisol levels or reaction time between fixed shifts (14 × days or nights) or a swing shift (7 × night shifts and then 7 × day shifts). The evidence suggests there is no negative outcomes from swing shifts, however it was noted that recovery took longer—despite adapting to night shift regardless of circumstances within a week	11	No significant difference between PVT and cortisol levels between the two shift types
Vanttola et al. (2019)	Observational field study	Finland	Finnish airline maintenance, customer service and catering team members	Varying shifts	PVT KSS EEG-based sleep recordings	Workers for a major Finnish airline company underwent questionnaire-based shift work disorder (SWD) testing and compared sleep and alertness to employees who did not match the SWD criteria. Employees with SWD were found to have disturbed sleep and alertness on morning and night shifts but also less compensatory sleep required on rostered days off	12	KSS and PVT congruent
Waggoner et al. (2012)	Prospective observational, quantitative	United States	Active-duty police officers	Fixed rotating schedules but finished in 5 × 10.7 h night shifts	PVT Driving simulator—lane deviations KSS	Police officers performed a driving simulation after the completion of a rotation of night shifts and after a period of time off to examine vigilant attention, psychomotor vigilance and subjective sleepiness. The results show a marked increase in metres of lane deviation, psychomotor vigilance test lapses and subjective sleepiness compared to their well-rested control	12	PVT and KSS congruent

PVT, psychomotor vigilance test, NHLBI National heart, lung and blood institute study quality assessment rating, KSS Karolinska Sleepiness Scale, RDO roster days off, EEG Electroencephalogram

^aM = morning shift (0700–1530), E = evening shift (1500–2330), N = night shift (2300–0730), R = Rest day

Acknowledgements The authors are grateful to Mr. Ric (Fredric) Bouvier and his family, for the support and inspiration in completing this research. Further thanks are extended to Ms. Anthea Ingram and Ms. Mary White for their proof reading of the manuscript.

Funding Matthew Ferris is supported by Monash Universities Ric (Fredric) Bouvier Scholarship and Monash Universities Jubilee Scholarship to complete this review. The author(s) received no other financial support for the research, authorship and/or publication of this article.

Compliance with ethical standards

Conflict of interest The author(s) declare no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

References

- Anderson C, Sullivan JP, Flynn-Evans EE et al (2012) Deterioration of neurobehavioral performance in resident physicians during repeated exposure to extended duration work shifts. *Sleep* 35(8):1137–1146
- Arsintescu L, Mulligan J, Flynn-Evans E (2017) Evaluation of a psychomotor vigilance task for touch screen devices. *Sage* 59(4):661–670
- Atkins D, Best D, Briss PA et al (2004) Grading quality of evidence and strength of recommendations. *MBJ* 328(7454):1490
- Basner M, Dinges D (2011) Maximising sensitivity of the psychomotor vigilance test (PVT) to sleep loss. *Sleep* 34(5):581–591
- Basner M, Mollicone D, Dinges D (2011) Validity and sensitivity to of a brief psychomotor vigilance test (PVT-B) to total and partial sleep deprivation. *Sci Direct* 69(11):949–959
- Basner M, Dinges DF, Shea JA et al (2017) Sleep and alertness in medical interns and residents: an observational study on the role of extended shifts. *Sleep* 40(4):1–8
- Behrens T, Burek K, Pallapies D et al (2019) Decreased psychomotor vigilance of female shift workers after working night shifts. *PLoS ONE* 14:1–17
- Bjorvatn B, Stangenes K, Oyane N et al (2006) Subjective and objective measures of adaption and readaptation to night work on an oil rig in the North Sea. *Sleep* 29(6):821–829
- Bonnefond A, Harma M, Hakola T et al (2006) Interaction of age with shift-related sleep-wakefulness, sleepiness, performance, and social life. *Exp Aging Res* 32:185–208
- Ferguson SA, Paech GM, Dorrian J et al (2011) Performance on a simple response time task: is sleep or work more important for miners? *Appl Ergon* 42:210–213
- Flynn-Evans EE, Arsintescu L, Gregory K et al (2018) Sleep and neurobehavioral performance vary by work start time during non-traditional day shifts. *Sleep Health* 4:476–484
- Ganesan S, Magee M, Stone J et al (2019) The impact of shift work on sleep, alertness and performance in healthcare workers. *Sci Rep* 9:1–13
- Harma M, Tarja H, Irja K et al (2006) A controlled intervention study on the effects of a very rapidly forward rotating shift system on sleep-wakefulness and well-being among young and elderly shift workers. *Int J Psychophysiol* 59:70–79
- Harris A, Waage S, Ursin H et al (2010) Cortisol, reaction time test and health among offshore shift workers. *Psychoneuroendocrinology* 35:1339–1347
- Hilditch C, Centofanti S, Dorrian J et al (2016) A 30-minute, but not a 10-minute night-time nap is associated with sleep inertia. *Sleep* 39(3):675–685
- Miley A, Kecklund G, Akerstedt T (2016) Comparing two versions of the Karolinska Sleepiness Scale (KSS). *Sleep Biol Rhythms* 14(3):257–260
- Moher D, Liberati A, Tetzlaff J et al (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 6(7):e1000097
- National institute of Health (2014) National Heart, Lung and Blood Institute. Quality assessment tool for observational cohort and cross-sectional studies, <http://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools>
- Richter P, Rodenback A, Weess HG et al (2021) Shiftwork and alcohol consumption: a systematic review of the literature. *Eur Addict Res* 27(9):9–15
- Ruggiero J, Redeker N, Fiedler N et al (2012) Sleep and psychomotor vigilance in female shift workers. *Biol Res Nurs* 14(3):225–235
- Sadeghniai-Haghighi K, Yazdi Z (2015) Fatigue management in the workplace. *Ind Psychiatry J* 24(1):12–17
- Sayers A (2008) Tips and tricks in performing a systematic review—chapter 4. *Br J Gen Pract* 58(547):136
- Shattuck NL, Matsangas P (2016) Operational assessment of the 5-h on/ 10-h of watch standing schedule on a US navy ship: sleep patterns, mood and psychomotor vigilance performance of crew members in the nuclear reactor department. *Ergonomics* 59(5):657–664
- Tanaka K, Takahashi M, Tanaka M et al (2011) Brief morning exposure to bright light improves subjective symptoms and performance in nurses with rapidly rotating shifts. *J Occup Health* 53:258–266
- Thompson B, Stock M, Banuelas V et al (2016) The impact of a rigorous multiple work shift schedule and day versus night shift work on reaction time and balance performance in female nurses: a repeated measures study. *J Occup Environ Med* 58(7):737–743
- Vanttola P, Harma M, Viitasalo K et al (2019) Sleep and alertness in shift work disorder: findings of a field study. *Int Arch Occup Environ Health* 92:523–533
- Veritas Health Innovation (2017) Covidence systematic review software. Melbourne, VIC, Australia
- Waggoner L, Grant D, Van Dongen H et al (2012) A combined field and laboratory design for assessing the impact of night shift work on police officer operational performance. *Sleep* 35(11):1575–1577
- Wilson M, Permito R, English A et al (2019) Performance and sleepiness in nurses working 12-h day shifts or night shifts in a community hospital. *Accid Anal Prev* 126:43–46
- Wolkow A, Aisbett B, Reynolds J et al (2016) Acute psychophysiological relationships between mood, inflammatory and cortisol changes in response to simulated physical fire-fighting work and sleep restriction. *Appl Psychophysiol Biofeedback* 41(2):165–180
- Wolkow AP, Rajaratnam SMW, Anderson C et al (2019) Recommendations for current and future countermeasures against sleep disorders and sleep loss to improve road safety in Australia. *Intern Med J* 49(9):1181–1184
- Wyatt J, Cecco A, Czeisler C, Dijk D (1999) Circadian temperature and melatonin rhythms, sleep and neurobehavioral function in humans living on a 20-h day. *AM J Phys* 277:1152–1163

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.