DIPLOME D'ÉTAT IPA

1ère année – UE Langue VIVANTE

Fascicule de lectures

Complementary Therapies in Medicine 39 (2018) 14–18

Contents lists available at ScienceDirect

Complementary Therapies in Medicine

journal homepage: www.elsevier.com/locate/ctim

Association between dietary flavonoids intake and prostate cancer risk: A case-control study in Sicily

Giulio Reale[a], Giorgio I. Russo[a,*], Marina Di Mauro[a], Federica Regis[a], Daniele Campisi[a], Arturo Lo Giudice[a], Marina Marranzano[b], Rosalia Ragusa[c], Tommaso Castelli[a], Sebastiano Cimino[a], Giuseppe Morgia[a]

[a] Urology section – University of Catania, Catania, Italy
[b] Department of Medical and Surgical Sciences and Advanced Technologies "G.F. Ingrassia", Section of Hygiene and Preventive Medicine, University of Catania, Catania, Italy
[c] Health Direction of A.O.U.-VE, Azienda Ospedaliera Universitaria Policlinico "Vittorio Emanuale", Catania, Italy

ARTICLE INFO

Keywords:
Flavonoids
Prostate cancer
Risk factors
Diet
Mediterranean diet

ABSTRACT

Objectives: The aim of this study is to test the association between dietary flavonoids intake and prostate cancer (PCa) in a sample of southern Italian individuals.
Design: A population-based case–control study on the association between PCa and dietary factors was conducted from January 2015 to December 2016, in a single institution.
Setting: Patients with elevated PSA (Prostate Specific Antigen) and/or suspicion of PCa underwent transperineal prostate biopsy (≥12 cores). A total of 118 histopathological-verified PCa cases were collected and matched with controls, which were selected from a sample of 2044 individuals randomly recruited among the same reference population. Finally, a total of 222 controls were selected.
Main outcome measures: Prevalence of PCa.
Results: Consumption of certain groups of flavonoids significantly differed between controls and cases, in particular: flavonols (63.36 vs 37.14 mg/d, P < 0.001), flavanols (107.61 vs. 74.24 mg/d, P = .016), flavanones (40.92 vs. 81.32 mg/d, P < 0.001), catechins (63.36 vs. 36.18 mg/d, P = .006). In the multivariate model, flavanols and flavones were associated with reduced risk of PCa, despite not in the highest quartile of intake. Higher flavonol and catechin intake was consistently associated with reduced risk of PCa (Odds Ratio (OR) = 0.19, 95% CI: 0.06–0.56 and OR = 0.12, 95% CI: 0.04–0.36). In contrast, the highest intake of flavanones was positively associated with PCa.
Conclusion: Flavonols and catechins have proved to be the most promising molecules for a potential protective role against PCa. Nevertheless, further research on flavanones is needed to better establish whether they are associated with PCa.

1. Introduction

Prostate cancer (PCa) is the most common cancer in men with incidental diagnosis.[1] Current statistics predicted that in the USA new cases of PCa will be more than 150,000 per annum over the next few years [2]. Although major efforts have been paid to prevent this cancer, including identification of risk factors (racial/ethnic background and family history) [3], there are aspects of etiopathogenesis still not clarified. Identifying causal factors of PCa would lead to new prevention methods. Also for PCa, as for benign prostatic hyperplasia, it has been hypothesized the possibility that the etiopathogenesis of the disorder is linked to chronic inflammation: the main pathway proposed suggests that the presence of oxidative stress associated to chronic inflammation in the cellular environment causes an increase of pro-inflammatory cytokines and growth factors, which in turn may determine an increase of the speed of cell replication, and therefore the possibility of incurring mutations [4,5]. If there is a correlation between prostate diseases, oxidative stress and chronic inflammation, the role of compounds with antioxidant action could play an important role in the prevention of PCa.

Current evidence suggests that adherence to plant-based dietary patterns, such as the Mediterranean diet, may reduce risk of PCa [6–8]. Moreover, patients exhibiting greater adherence to the Mediterranean diet after diagnosis of non-metastatic PCa were associated with lower

* Corresponding author.
E-mail address: giorgioivan.russo@unict.it (G.I. Russo).

https://doi.org/10.1016/j.ctim.2018.05.002
Received 2 March 2018; Received in revised form 11 May 2018; Accepted 14 May 2018
Available online 15 May 2018
0965-2299/

overall mortality [9]. Interestingly, pure vegetarian dietary patterns did not show the same inverse association with risk of PCa [10], suggesting that the retrieved associations are related to beneficial compounds rather than only reduction of unhealthy ones (i.e., trans-fatty acids). Among the others, important components of the Mediterranean diet that have been hypothesized to be responsible for its potential beneficial effects are polyphenols [11]. These compounds occur naturally in plant-derived foods, such as fruits and vegetables, nuts, whole-grains, olive oil, coffee and tea. Based on their biochemical structures, they are divided into different subclasses of which, the most representative, are flavonoids [12]. In turn, the six principal subclass of flavonoids are flavonols, flavones, flavanones, flavanols, and anthocyanidins [13]. The interest on studying flavonoids as anti-cancer substances depends on a variety of properties and potential mechanisms of action that may affect the risk of cancer [14–17]. In example, flavonoids were shown to modulate several molecular pathways implicated in PCa carcinogenesis process; in particular, by targeting important transcription factors, such as NF-kB (Nuclear Factor kappa-light-chain-enhancer of activated B cells) and AP-1 (activator protein-1), implicated in regulation of inflammatory response. Results on the association between dietary flavonoid intake and human health are promising, but research in relation to cancer is still ongoing and rather incomplete [18]. Some studies investigated the relation between dietary flavonoids and PCa [19,20]: however, the relation between PCa and flavonoid subclasses remains unclear. Thus, the aim of this article is to test the association between dietary flavonoids, including all major subclasses, and PCa in a sample of southern Italian individuals.

2. Material and methods

2.1. Study population

A population-based case–control study on the association between PCa and dietary factors was conducted from January 2015 to December 2016 in a single institution of the municipality of Catania, southern Italy. Patients with elevated PSA and/or suspicious PCa underwent transperineal prostate biopsy (≥12 cores). A total of 118 histopathological-verified PCa cases were collected.

Controls were selected from a sample of 2044 individuals included in a cohort study [21]: individuals were randomly selected among the same reference population of the cases, and matched by age, BMI, and smoking status with cases. A total of 222 controls were selected.

All the study procedures were carried out in accordance with the Declaration of Helsinki (1989) of the World Medical Association and participants provided written informed consent after accepting to participate. The study protocol was approved by the ethic committee of the referent health authority (Policlinico Hospital of Catania, Registration number: 41/2015).

2.2. Data collection

Demographics (including age, and educational level) and lifestyle characteristics (including physical activity, smoking and drinking habits) were collected. Educational level was categorized as (i) low (primary/secondary), (ii) medium (high school), and (iii) high (university). Physical activity level was evaluated through the International Physical Activity Questionnaires (IPAQ) [22] which comprised a set of questionnaires (5 domains) investigating the time spent being physically active in the last 7 days: based on the IPAQ guidelines, final scores allows to categorized physical activity level as (i) low, (ii) moderate, and (iii) high. Smoking status was categorized as (i) non-smoker, (ii) ex-smoker, and (iii) current smoker. Alcohol consumption was categorized as (i) none, (ii) moderate drinker (0.1–12 g/d) and (iii) regular drinker (> 12 g/d).

2.3. Dietary assessment

Dietary data was collected by using two food frequency questionnaires (FFQs) specifically developed and validated for the Sicilian population[23,24]. The long-version FFQ consisted of 110 food and drink items. Patients were specifically asked whether they changed their diet due to course of the disease and to answer to the questionnaire referring to their habitual diet before the disease. Participants were asked how often, on average, they had consumed foods and drinks included in the FFQ, with nine responses ranging from “never” to “4-5 times per day”. Intake of food items characterized by seasonality referred to consumption during the period in which the food was available and then adjusted by its proportional intake in one year.

2.4. Estimation of flavonoid intake

The methodology used to retrieve dietary flavonoids has been widely used in literature and largely described elsewhere [25]. Briefly, data on the polyphenol content in foods was obtained from the Phenol-Explorer database (www.phenol-explorer.eu). A new module of the Phenol-Explorer database containing information on the effects of cooking and food processing on polyphenol contents was used whenever possible in order to apply polyphenol-specific retention factors [26]. A total of 75 items were searched in the database after exclusion of foods that contained no polyphenols. Following the standard portion sizes used in the study, food items were converted in g or ml and then proportioned to 24-h intake. Next, a search was carried out in the Phenol-Explorer database to retrieve mean content values for flavonoid (total and major subclasses) contained in the foods obtained and their intake was then calculated by multiplying the flavonoid content by the daily consumption of each food. Finally, intake of flavonoids was adjusted for total energy intake (kcal/d) using the residual method.

2.5. Statistical analysis

Categorical variables are presented as frequency and percentage, continuous variables are presented as mean and standard deviation. Differences of frequency between groups were calculated by Chi-square test. Total flavonoid intake distribution was tested for normality distribution with the Kolmogorov-Smirnov test and it followed a slightly asymmetric normal distribution due to extreme values of the upper side. Mann-Whitney *U* test and Kruskall-Wallis test were used to compare differences in intakes between groups, as appropriate. Association between dietary intake of total and subclasses of flavonoid and PCa was calculated through logistic regression analysis adjusted for age (years, continuous), energy intake (kcal/d, continuous), weight status (normal, overweight, obese), smoking status (smokers, non-smokers), alcohol consumption (< 12 g/d, ≥12 g/d), physical activity level (low, medium, high), family history of PCa. All reported *P* values were based on two-sided tests and compared to a significance level of 5%. SPSS 17 (SPSS Inc., Chicago, IL, USA) software was used for all the statistical calculations.

3. Results

Table 1 lists the baseline characteristics of cases and controls. Besides the characteristics for which controls were matched with cases, most of other variables had different distribution between groups: specifically, among cases there was a higher prevalence of low education, low physical activity level, higher alcohol consumption and family history of PCa than controls, despite mean BMI levels were lower in the former than in the latter.

No significant differences between cases and controls have been found concerning total dietary flavonoids (Table 2). However, regarding flavonoid subclasses, differences between intake of some compounds were statistically significant: flavonols (63.36 vs. 37.14,

Table 1
Baseline characteristics of cases and controls.

	Cases (n = 118)	Controls (n = 222)	P-value
Age (y), mean (SD)	69.13 (6.60)	68.09 (8.18)	0.238
BMI, mean (SD)	26.49 (3.34)	27.49 (3.28)	< 0.001
Weight status, n (%)			0.220
Normal	42 (35.6	59 (26.6)	
Overweight	60 (50.8)	127 (57.2)	
Obese	16 (13.6)	36 (16.2)	
Smoking status, n (%)			0.220
Non-smoker	68 (57.6)	143 (64.4)	
Current smoker	50 (42.4)	79 (35.6)	
Alcohol intake, n (%)			< 0.001
< 12 g/d	55 (46.6)	153 (68.9)	
≥12 g/d	63 (53.4)	69 (31.1)	
Education, n (%)			< 0.001
Primary	96 (81.4)	49 (22.1)	
Secondary	22 (18.6)	173 (77.9)	
Physical activity level, n (%)			< 0.001
Low	38 (32.2)	49 (26.2)	
Medium	64 (54.2)	67 (35.8)	
High	16 (13.6)	71 (38.0)	
Family history of prostatic cancer, n (%)	43 (36.44)	9 (4.05)	< 0.001

Table 2
Mean differences of total and subclasses of flavonoid intake between cases and controls.

	Cases (n = 118)	Controls (n = 222)	P-value
Flavonoids, mean (SD)	268.68 (166.87)	286.04 (207.38)	0.434
Anthocyanins, mean (SD)	59.82 (51.55)	59.09 (55.57)	0.903
Flavonols, mean (SD)	37.14 (29.23)	63.36 (46.85)	< 0.001
Flavanols, mean (SD)	74.24 (71.42)	107.61 (139.97)	0.016
Flavanones, mean (SD)	81.32 (76.94)	40.92 (40.38)	< 0.001
Flavones, mean (SD)	9.86 (13.84)	8.50 (8.70)	0.266
Catechins, mean (SD)	36.18 (46.43)	63.36 (100.11)	0.006

P < 0.001), flavanols (107.61 vs. 74.24, P = 0.016), flavanones (40.92 vs. 81.32 P < 0.001), and catechins (63.36 vs. 36.18, P = 0.006). There was no statistical significant difference in the subclasses of flavones and anthocyanins between PCa group when compared to controls.

The univariate logistic regression analysis showed that intake of various compounds was associated with PCa (Table 3). When analysis was adjusted for potential confounding factors (including age, energy intake, weight status, smoking status, alcohol consumption, physical activity level, family history of prostatic cancer), only the highest intake of flavonol and catechin intake was consistently associated with less likelihood of having PCa (OR = 0.19, 95%CI: 0.07–0.50 and OR = 0.12, 95%CI: 0.04–0.36, respectively). Moreover, flavanol and flavone intake was also associated with PCa, despite the significant association was relative to third quartile of intake, while the highest showed no significant results (Table 3). In contrast, the highest quartile of flavanone intake was directly associated with higher likelihood of having PCa compared to the lowest (Table 3). No significant association between total flavonoid intake and PCa was found.

4. Discussion

In this study, we tested the association between dietary flavonoid intake and PCa in a sample of individuals living in the Mediterranean area. Although results were not consistently significant for all the various compounds tested, higher intake of nearly all flavonoids was inversely associated with PCa compared to lower intake.

Published data on flavonoid intake and risk of PCa is somehow contrasting. Some individual studies and recent meta-analyses showed

Table 3
Association between quartiles of total and subclasses of flavonoid intake and prostate cancer.

	Q1	Q2	Q3	Q4
Flavonoids				
No. of cases	30	35	19	36
OR (95% CI)[a]	Ref.	0.93 (0.49–1.76)	0.48 (0.24–0.97)	0.75 (0.39–1.45)
OR (95% CI)[b]	Ref.	1.44 (0.57–3.63)	0.51 (0.18–1.42)	1.10 (0.40–2.99)
Anthocyanins				
No. of cases	28	31	32	27
OR (95% CI)[a]	Ref.	1.09 (0.57–2.09)	0.98 (0.51–1.85)	0.62 (0.31–1.25)
OR (95% CI)[b]	Ref.	0.58 (0.23–1.45)	0.51 (0.20–1.31)	0.36 (0.12–1.06)
Flavonols				
No. of cases	64	28	15	11
OR (95% CI)[a]	Ref.	0.44 (0.24–0.79)	0.21 (0.10–0.42)	0.11 (0.53–0.26)
OR (95% CI)[b]	Ref.	0.44 (0.18–1.04)	0.19 (0.07–0.50)	0.19 (0.06–0.56)
Flavanols				
No. of cases	106	9	2	1
OR (95% CI)[a]	Ref.	0.71 (0.29–1.72)	0.03 (0.003–0.33)	0.35 (0.04–3.04)
OR (95% CI)[b]	Ref.	0.76 (0.22–2.58)	0.02 (0.002–0.31)	1.84 (0.17–19.4)
Flavanones				
No. of cases	18	10	19	71
OR (95% CI)[a]	Ref.	0.51 (0.21–1.21)	0.96 (0.45–2.06)	3.03 (1.56–5.87)
OR (95% CI)[b]	Ref.	0.26 (0.07–0.90)	1.59 (0.50–5.02)	5.76 (2.06–16.09)
Flavones				
No. of cases	37	32	10	39
OR (95% CI)[a]	Ref.	0.82 (0.44–1.51)	0.24 (0.11–0.55)	0.84 (0.46–1.54)
OR (95% CI)[b]	Ref.	0.61 (0.25–1.45)	0.33 (0.12–0.91)	0.87 (0.34–2.21)
Catechins				
No. of cases	50	22	31	15
OR (95% CI)[a]	Ref.	0.39 (0.21–0.75)	0.49 (0.27–0.89)	0.20 (0.097–0.41)
OR (95% CI)[b]	Ref.	0.21 (0.08–0.55)	0.35 (0.14–0.90)	0.12 (0.04–0.36)

[a] OR adjusted for energy intake (kcal/d, continuous).

[b] OR adjusted for age (years, continuous), energy intake (kcal/d, continuous), weight status (normal, overweight, obese), smoking status (smokers, non-smokers), alcohol consumption (< 12 g/d, ≥12 g/d), physical activity level (low, medium, high), family history of prostatic cancer.

that a direct association between specific classes of flavonoids (i.e., flavanones and anthocyanins) and likelihood of having PCa may exist, while results on other classes are largely inconsistent [19,20,27]. Thus, our findings are only partially in line with existing literature, as we found a potential protective association for certain classes, while we also reported direct association between flavanone intake and PCa.

Regarding the potential detrimental effects, it has been hypothesized that excessive antioxidants consumption may actually increase risk of cancer [28]. Despite not relative to flavanones, there is evidence that certain antioxidant vitamins, including ascorbic acid, may have detrimental effects on reactive oxygen species production processes related to PCa [29–31]; it is noteworthy that major source of ascorbic acid coincide with those of flavanones (i.e., citrus fruits), thus potentially leading to collinearity and confusing the interpretation of results on this flavonoid class.

Regarding the potential protective effects of flavonoids, there is evidence of general beneficial effects of dietary flavonoids on metabolic health [32,33] and metabolic status has been shown in a recent meta-analysis to be a mediating factor for PCa risk [34]. Flavonoids may act as activator of the transcription factor NF-kB that regulate a variety of cellular activities that include inflammation, immune response, cell growth and death, therefore resulting in a cascade of events that may lead to carcinogenesis [33]. Overexpression of NF-kB leads to the activation of several signalling pathways, among which also the activation of COX-2 (cyclooxygenase-2), which increases levels of pro-inflammatory cytokines, a perfect scenario for developing PCa cells. It has been shown, that synthetized analogs of flavanols are promising category of compounds that can inhibit cell growth and interfere with the components of the androgen receptor and PSA proteins in human-derived PCa cell line [35] or directly compromise the PCa cell vitality of several prostate cancer cell lines, including: 22Rv1, TRAMP C2, PC-3

and LNCaP [36] Moreover, flavonols may expression of several genes, such as MMP-1 (matrix metalloproteinase-1), MMP-9 (matrix metalloproteinase-9), MMP-14 (matrix metalloproteinase-14), c-Fos (proto-oncogene), c-Jun (a subunit of the transcription factor), and VEGF (vascular endothelial growth factor), that have been previously shown to be associated with PCa [36] and in particular in changing the ratio Bcl-2 (B-cell lymphoma 2)/Bax (apoptosis regulator) mRNA, which directly determines neoplastic cell sensitization for the apoptotic pathway [37,38].

Among individual compounds, higher intake of flavonols and, specifically, catechins was found to be significantly inversely associated with having PCa. Catechins are typically contained in tea, one of the most widely consumed beverage in the world [39]. The tea plant (*Camellia sinensis*) produced in Asia, China, Japan, and Thailand has been traditionally used in natural medicine[40]. The most represented individual compounds contained in tea, such as epigallocatechin (EGC), epicatechin-3-gallate (ECG), epicatechin (EC), and epigallocatechin-3-gallate (EGCG), have been studied in relation to PCa[41]. Among the main mechanisms of action, these molecules may play a role in the prevention of etiopathogeneiss of PCa by inducing cell growth arrest and apoptosis primarily via p53-dependent pathway or inhibiting COX-2 (inducible enzymatic isoform, rapidly induced by growth factors, tumor promoters, oncogenes, and carcinogens) without affecting COX-1 (cyclooxygenase-1), at both the mRNA and protein levels [15,40], Besides these biological properties, previous data found no association between PCa risk and green tea intake [42–44]. However, a recent meta-analysis showed that green tea intake might reduce the incidence of PCa with a linear dose–response effect and decrease PCa risk significantly with daily intake of over 7 cups/day [45]. There is still only limited evidence from randomized clinical trials investigating the association of green tea catechins with the risk of PCa, however showing that green tea catechins had a significant effect on the reduction of PCa risk compared to placebo [46,47].

Some limitations of the present analysis have to be addressed for a better interpretation of the results. First, the observational nature of study does not permit to assess causal relationships, rather only associations. Second, the variety of flavonoid composition is complex and use of FFQs may lead to measurement errors. Third, our sample was based on patients with no previous biopsy and PCa diagnosis could have been underestimated. In contrast, controls were selected from the general population, thus we are unaware whether undiagnosed PCa cases existed.

5. Conclusion

In conclusions, flavonoids may have different levels of protection against PCa. Among the compounds tested in our study, flavonols and catechins have proved to be the most interesting molecules for a protective value. Future randomized clinical trials are needed to strengthen the findings obtained in this study and provide adequate evidence of the potential protective effects of flavonoids toward PCa risk.

Compliance with ethical standards

All the study procedures were carried out in accordance with the Declaration of Helsinki (1989) of the World Medical Association and participants provided written informed consent after accepting to participate. The study protocol was approved by the ethic committee of the referent health authority (Policlinico Hospital of Catania, Registration number: 41/2015).

Conflict of interest

Each author declares no conflict of interest.

Author's contribution

G.R.: Manuscript writing
G.I.R.: Protocol/project development, data management, manuscript editing
F.R.: Data Collection
M.D.: Data Collection
D.C.: Data Collection
A.L.: Data Collection
M.M.: Supervision
R.R.: Supervision
V.F.: Supervision
S.C.: Supervision
G.M.: Supervision

References

1. Dy GW, Gore JL, Forouzanfar MH, Naghavi M, Fitzmaurice C. Global burden of urologic cancers, 1990–2013. *Eur Urol.* 2017;71(3):437–446.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics. *CA Cancer J Clin.* 2017;67(1):7–30.
3. Mottet N, Bellmunt J, Bolla M, et al. EAU-ESTRO-SIOG guidelines on prostate cancer. part 1: screening, diagnosis, and local treatment with curative intent. *Eur Urol.* 2017;71(April (4)):618–629. http://dx.doi.org/10.1016/j.eururo.2016.08.003 [Epub 2016 Aug 25].
4. De Nunzio C, Kramer G, Marberger M, et al. The controversial relationship between benign prostatic hyperplasia and prostate cancer: the role of inflammation. *Eur Urol.* 2011;60(1):106–117.
5. Kruslin B, Tomas D, Dzombeta T, Milkovic-Perisa M, Ulamec M. Inflammation in prostatic hyperplasia and carcinoma-basic scientific approach. *Front Oncol.* 2017;7:77.
6. Schwingshackl L, Hoffmann G. Adherence to Mediterranean diet and risk of cancer: an updated systematic review and meta-analysis of observational studies. *Cancer Med.* 2015;4(12):1933–1947.
7. Grosso G, Bella F, Godos J, et al. Possible role of diet in cancer: systematic review and multiple meta-analyses of dietary patterns, lifestyle factors, and cancer risk. *Nut Rev.* 2017;75(June (6)):405–419.
8. Russo G.I., Campisi D., Di Mauro M., et al. Dietary consumption of phenolic acids and prostate cancer: a case-control study in sicily.
9. Kenfield SA, DuPre N, Richman EL, Stampfer MJ, Chan JM, Giovannucci EL. Mediterranean diet and prostate cancer risk and mortality in the Health Professionals Follow-up Study. *Eur Urol.* 2014;65(5):887–894.
10. Godos J, Bella F, Sciacca S, Galvano F, Grosso G. Vegetarianism and breast, colorectal and prostate cancer risk: an overview and meta-analysis of cohort studies. *J Hum Nutr Diet.* 2017;30(June (3)):349–359.
11. Godos J, Rapisarda G, Marventano S, Galvano F, Mistretta A, Grosso G. Association between polyphenol intake and adherence to the Mediterranean diet in Sicily, southern Italy. *NFS J.* 2017;8:1–7.
12. Del Rio D, Costa LG, Lean ME, Crozier A. Polyphenols and health: what compounds are involved? *Nutr Metab Cardiovasc Dis.* 2010;20(1):1–6.
13. Del Rio D, Rodriguez-Mateos A, Spencer JP, Tognolini M, Borges G, Crozier A. Dietary (poly)phenolics in human health: structures, bioavailability, and evidence of protective effects against chronic diseases. *Antioxid. Redox Signal.* 2013;18(14):1818–1892.
14. Ramos S. Effects of dietary flavonoids on apoptotic pathways related to cancer chemoprevention. *J Nutr Biochem.* 2007;18(7):427–442.
15. Khan N, Adhami VM, Mukhtar H. Review: green tea polyphenols in chemoprevention of prostate cancer: preclinical and clinical studies. *Nutr Cancer.* 2009;61(6):836–841.
16. Adhami VM, Malik A, Zaman N, et al. Combined inhibitory effects of green tea polyphenols and selective cyclooxygenase-2 inhibitors on the growth of human prostate cancer cells both in vitro and in vivo. *Clin Cancer Res.* 2007;13(5):1611–1619.
17. Abbas A, Patterson 3rd W, Georgel PT. The epigenetic potentials of dietary polyphenols in prostate cancer management. *Biochem Cell Biol.* 2013;91(6):361–368.
18. Grosso G, Micek A, Godos J, et al. Dietary flavonoid and lignan intake and mortality in prospective cohort studies: systematic review and dose-response meta-analysis. *Am J Epidemiol.* 2017;185(12):1304–1316.
19. Guo K, Liang Z, Liu L, Li F, Wang H. Flavonoids intake and risk of prostate cancer: a meta-analysis of observational studies. *Andrologia.* 2016;48(10):1175–1182.
20. Grosso G, Godos J, Lamuela-Raventos R, et al. A comprehensive meta-analysis on dietary flavonoid and lignan intake and cancer risk: level of evidence and limitations. *Mol Nutr Food Res.* 2017;61(Apr (4)).
21. Grosso G, Marventano S, D'Urso M, Mistretta A, Galvano F. The Mediterranean healthy eating, ageing, and lifestyle (MEAL) study: rationale and study design. *Int J food Sci Nutr.* 2016:1–10.
22. Craig CL, Marshall AL, Sjostrom M, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc.* 2003;35(8):1381–1395.
23. Marventano S, Mistretta A, Platania A, Galvano F, Grosso G. Reliability and relative validity of a food frequency questionnaire for Italian adults living in Sicily: southern Italy. *Int J Food Sci Nutr.* 2016:1–8.

24. Buscemi S, Rosafio G, Vasto S, et al. Validation of a food frequency questionnaire for use in Italian adults living in Sicily. *Int J Food Sci Nutr.* 2015;66(4):426–438.
25. Godos J, Rapisarda G, Marventano S, Galvano F, Mistretta A, Grosso G. Association between polyphenol intake and adherence to the Mediterranean diet in Sicily, southern Italy. *NFS J.* 2017;8(Aug):1–7.
26. Rothwell JA, Perez-Jimenez J, Neveu V, et al. Phenol-Explorer 3.0: a major update of the Phenol-Explorer database to incorporate data on the effects of food processing on polyphenol content. *Database (Oxford).* 2013;2013:bat070.
27. Bosetti C, Bravi F, Talamini R, et al. Flavonoids and prostate cancer risk: a study in Italy. *Nutr Cancer.* 2006;56(2):123–127.
28. Bjelakovic G, Nikolova D, Gluud LL, Simonetti RG, Gluud C. Antioxidant supplements for prevention of mortality in healthy participants and patients with various diseases. *Cochrane Database Syst Rev.* 2012(3):D007176.
29. Fukumura H, Sato M, Kezuka K, et al. Effect of ascorbic acid on reactive oxygen species production in chemotherapy and hyperthermia in prostate cancer cells. *J Physiol Sci.* 2012;62(3):251–257.
30. Grosso G, Bei R, Mistretta A, et al. Effects of vitamin C on health: a review of evidence. *Front Biosci (Landmark Ed.).* 2013;18:1017–1029.
31. Heaney ML, Gardner JR, Karasavvas N, et al. Vitamin C antagonizes the cytotoxic effects of antineoplastic drugs. *Cancer Res.* 2008;68(19):8031–8038.
32. Grosso G, Stepaniak U, Micek A, Stefler D, Bobak M, Pajak A. Dietary polyphenols are inversely associated with metabolic syndrome in Polish adults of the HAPIEE study. *Eur J Nutr.* 2017;56(June (4)):1409–1420.
33. Mansuri ML, Parihar P, Solanki I, Parihar MS. Flavonoids in modulation of cell survival signalling pathways. *Genes Nutr.* 2014;9(3):400.
34. Gacci M, Russo GI, De Nunzio C, et al. Meta-analysis of metabolic syndrome and prostate cancer. *Prostate Cancer Prostatic Dis.* 2017;20(2):146–155.
35. Britton RG, Horner-Glister E, Pomenya OA, et al. Synthesis and biological evaluation of novel flavonols as potential anti-prostate cancer agents. *Eur J Med Chem.* 2012;54:952–958.
36. Hill CU, Saad SE, Britton RG, et al. Inhibition of prostate cancer cell growth by 3',4',5'-trimethoxyflavonol (TMFol). *Cancer Chemother Pharmacol.* 2015;76(1):179–185.
37. Lewandowska U, Szewczyk K, Owczarek K, et al. Flavanols from evening primrose (Oenothera paradoxa) defatted seeds inhibit prostate cells invasiveness and cause changes in Bcl-2/Bax mRNA ratio. *J Agric Food Chem.* 2013;61(12):2987–2998.
38. Lewandowska U, Szewczyk K, Owczarek K, et al. Flavanols from Japanese quince (Chaenomeles japonica) fruit inhibit human prostate and breast cancer cell line invasiveness and cause favorable changes in Bax/Bcl-2 mRNA ratio. *Nutr Cancer.* 2013;65(2):273–285.
39. Cimino S, Sortino G, Favilla V, et al. Polyphenols: key issues involved in chemoprevention of prostate cancer. *Oxid Med Cell Longev.* 2012;2012:632959.
40. Pandey M, Gupta S. Green tea and prostate cancer: from bench to clinic. *Front Biosci (Elite Ed.).* 2009;1:13–25.
41. Khan N, Mukhtar H. Modulation of signaling pathways in prostate cancer by green tea polyphenols. *Biochem Pharmacol.* 2013;85(5):667–672.
42. Zheng J, Yang B, Huang T, Yu Y, Yang J, Li D. Green tea and black tea consumption and prostate cancer risk: an exploratory meta-analysis of observational studies. *Nutr Cancer.* 2011;63(5):663–672.
43. Lin YW, Hu ZH, Wang X, et al. Tea consumption and prostate cancer: an updated meta-analysis. *World J Surg Oncol.* 2014;12:38.
44. Fei X, Shen Y, Li X, Guo H. The association of tea consumption and the risk and progression of prostate cancer: a meta-analysis. *Int J Clin Exp Med.* 2014;7(11):3881–3891.
45. Guo Y, Zhi F, Chen P, et al. Green tea and the risk of prostate cancer: a systematic review and meta-analysis. *Medicine (Baltimore).* 2017;96(13):e6426.
46. Brausi M, Rizzi F, Bettuzzi S. Chemoprevention of human prostate cancer by green tea catechins: two years later. *Follow-Up Update Eur Urol.* 2008;54(2):472–473.
47. Kumar NB, Pow-Sang J, Egan KM, et al. Randomized, placebo-controlled trial of green tea catechins for prostate cancer prevention. *Cancer Prev Res (Phila).* 2015;8(10):879–887.

Cancer Epidemiology 54 (2018) 71–74

Contents lists available at ScienceDirect

Cancer Epidemiology

journal homepage: www.elsevier.com/locate/canep

Adiposity at different periods of life and risk of adult glioma in a cohort of postmenopausal women

Geoffrey C. Kabat*, Thomas E. Rohan

Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY 10461, United States

ARTICLE INFO

Keywords:
adiposity
body mass index
waist circumference
glioma
glioblastoma
cohort studies

ABSTRACT

Background: Little is known about risk factors for adult glioma. Adiposity has received some attention as a possible risk factor.

Methods: We examined the association of body mass index (BMI), waist circumference (WC) and waist-to-hip ratio (WHR), measured at enrollment, as well as self-reported weight earlier in life, with risk of glioma in a large cohort of postmenopausal women. Over 18 years of follow-up, 217 glioma cases were ascertained, including 164 glioblastomas. Cox proportional hazards models were used to estimate hazard ratios and 95% confidence intervals.

Results: There was a modest, non-significant trend toward increasing risk of glioma and glioblastoma with increasing measured BMI and WHR. No trend was seen for WC. Self-reported BMI earlier in life showed no association with risk.

Conclusions: Our weak findings regarding the association of adiposity measures with risk of glioma are in agreement the results of several large cohort studies. In view of the available evidence, adiposity is unlikely to represent an important risk factor for glioma.

1. Introduction

Gliomas are the most common primary intracranial cancer and the most fatal type of brain tumor [1]. Little is known about risk factors for gliomas [2]. Established risk factors include several rare, inherited genetic syndromes and exposure to ionizing radiation [2]; however, these account for only a small proportion of gliomas.

Adiposity has received attention as a possible risk factor for glioma [3–10]. Such an association might be mediated by circulating insulin levels, since hyperinsulinemia is common among obese and sedentary individuals, and insulin has pro-mitotic properties [4,5]. Insulin crosses the blood-brain barrier, and insulin's actions within the CNS are mediated by two canonical pathways involved in carcinogenesis [11]. However, most studies have found little evidence of an association with adiposity. Of two meta-analyses of the association of adult body mass index (BMI) and glioma, one found no association in men or women [9], whereas the other reported a significant association in women but not in men [8]. Two other studies [4,5] showed positive associations of BMI at age 18 and 21, respectively, with glioma risk, raising the possibility of an etiologic role of obesity earlier in life.

We examined the association of BMI, waist circumference (WC), and waist-to-hip ratio (WHR) measured at enrollment with risk of glioma in a large cohort of postmenopausal women. Additionally, we assessed the association of BMI at age 18, 35, and 50 with risk of glioma in a subset of the study population.

2. Methods

The Women's Health Initiative is a large, multicenter study designed to advance understanding of the determinants of major chronic diseases in postmenopausal women. It is composed of a clinical trial component (CT, n = 68,132) and an observational study component (OS, n = 93,676) [12]. Women between the ages of 50 and 79 and representing the major racial/ethnic groups were recruited from the general population at 40 clinical centers throughout the US between 1993 and 1998.

At study entry, self-administered questionnaires were used to collect information on demographics, medical, reproductive, and family history, and lifestyle factors, including smoking history, alcohol consumption, diet, and recreational physical activity. All participants had their weight and height measured by trained staff at baseline. Weight was measured to the nearest 0.1 kg, and height to the nearest 0.1 cm. Body mass index was computed as weight in kilograms divided by the square of height in meters. Waist circumference and waist-to-hip ratio

* Corresponding author at: Department of Epidemiology and Population Health, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461, United States.
E-mail address: Geoffrey.kabat@einstein.yu.edu (G.C. Kabat).

https://doi.org/10.1016/j.canep.2018.03.008
Received 17 January 2018; Received in revised form 28 February 2018; Accepted 14 March 2018
Available online 11 April 2018
1877-7821/

Table 1
Association of baseline anthropometric measures with risk of glioma and glioblastoma in the Women's Health Initiative (n = 161,119[a]).

	Glioma (n cases = 217)				Glioblastoma (n cases = 164)			
BMI (kg/m^2)[b]	N cases	N non-cases	HR[d]	95% CI	N cases	N non-cases	HR[d]	95% CI
18.5– < 25.0	70	54,827	1.00	Ref.	53	54,827	1.00	Ref.
25.0– < 30.0	74	55,571	1.06	0.76–1.47	54	55,571	1.04	0.71–1.53
30.0– < 35.0	43	29,684	1.21	0.81–1.79	33	29,684	1.25	0.79–1.97
≥35.0	25	18,582	1.30	0.81–2.08	20	18,582	1.46	0.86–2.50
Missing	5	106			4	106		
P for linear trend			*0.32*				*0.18*	
Waist circumference (cm)								
< 76.0	50	40,525	1.00	Ref.	33	40,525	1.00	Ref.
76.0- < 84.5	60	40,226	1.18	0.80–1.73	48	40,226	1.44	0.92–2.25
84.5- < 95.0	57	39,395	1.24	0.84–1.83	42	39,395	1.38	0.86–2.20
≥95.0	50	40,756	1.11	0.74–1.68	41	40,756	1.41	0.87–2.29
P for linear trend			*0.66*				*0.25*	
WHR[c]								
< 0.76	47	39,855	1.00	Ref.	33	39,855	1.00	Ref.
0.76- < 0.80	59	39,846	1.22	0.83–1.80	44	39,846	1.28	0.81–2.01
0.80- < 0.86	53	40,044	1.22	0.82–1.81	41	40,044	1.30	0.82–2.06
≥0.86	56	39,725	1.34	0.90–2.00	44	39,725	1.45	0.92–2.30
Missing	2	694			2	694		
P for linear trend			*0.17*				*0.09*	

Abbreviations: HR–hazard ratio; 95% CI–95% confidence interval; BMI–body mass index; WHR–waist-to-hip ratio.
[a] Women with anthropometric measurements.
[b] 2,132 women with BMI < 18.5 or BMI missing.
[c] 740 women missing WHR measurement.
[d] Adjusted for age, smoking status, alcohol intake, physical activity, hormone therapy, years of education, ethnicity, and treatment status.

were also measured. Questions about physical activity at baseline referred to a woman's usual pattern of activity, including walking and recreational physical activity. A variable "current total leisure-time physical activity" (MET-hours/week) was computed by multiplying the number of hours per week of leisure-time physical activity by the metabolic equivalent (MET) value of the activity and summing over all types of activities [13].

BMI, which reflects overall adiposity, was categorized according to the WHO classification: 18.5- < 25.0 kg/m^2 – normal weight, 25.0- < 30.0 kg/m^2 – overweight, and ≥30.0 kg/m^2 – obese. Waist circumference (WC), a measure of central adiposity, and waist-to-hip ratio (WHR), a measure of the ratio of central to lower extremity adiposity, were categorized into quartiles based on the distribution among non-cases. Information on weight at earlier ages was available only for participants in the Observational Study (n = 92,557) and was used to compute BMI at ages 18, 35, and 50. For this analysis, owing to the reduced sample size, we created tertiles based on the distribution in non-cases.

Clinical outcomes (including new cancer diagnoses) were updated semiannually in the CT and annually in the OS using in-person, mailed, or telephone questionnaires. Self-reports of malignancy, including gliomas, were verified by central review of medical records and pathology reports by trained physician adjudicators [14]. Among 161,119 WHI participants with anthropometric measurements, 217 cases of glioma were ascertained over a median of 17.8 years of follow-up. Of these, 164 had glioblastoma. Other gliomas included: mixed glioma, ependymoma NOS, well-differentiated low-grade astrocytoma, anaplastic astrocytoma, and oligodendroglioma.

2.1. Statistical analysis

Cox proportional hazards models were used to estimate hazard ratios (HRs) and 95% confidence intervals (CI) for the association of anthropometric factors with risk of glioma and glioblastoma, using days to event as the timescale. Participants who had not developed glioma by the end of follow-up, who had died, or who withdrew from the study before the end of follow-up were censored. Cases contributed person-time to the study from their date of enrollment until the date of diagnosis, and non-cases (participants who were censored) contributed person-time from their date of enrollment until the end of follow-up, date of death, or date of withdrawal from the study, whichever came first. Hazard ratios were computed by quartile of measured anthropometric variables and by tertiles of self-reported body weight at earlier ages. Covariates were selected for inclusion in the final model if their inclusion changed the parameter estimate by > 10%. Pack-years of smoking and use of hormone therapy did not improve the model and were excluded. The final model included age (continuous), smoking status (never, former, current), alcohol intake (servings/week – continuous), physical activity (metabolic equivalent tasks [MET]-hrs/wk continuous, years of education (less than high school, high school graduate/some college, college graduate, post-college), ethnicity (white, black, other), and allocation in the clinical trial arms or observational study. A test for linear trend over quantiles of anthropometric variables was performed by assigning the median value to each category and modeling this variable as a continuous variable. In order to account for women with a prevalent cancer at the time of enrollment, we carried out a sensitivity analysis in which we excluded women who reported a history of cancer at entry into the study. We tested the proportional hazards assumption using PROC LIFETEST (SAS Institute). The formal test for non-proportional hazards was not significant and the log-log survival plots did not indicate any marked deviation from normality. All analyses were performed in SAS 9.4 (SAS Institute, Cary, NC). All P-values are 2-sided.

3. Results

Glioma cases did not differ from non-glioma cases in terms of mean age, smoking status, mean pack-years of smoking, alcohol intake, hormone therapy, or physical activity. The proportion of whites was higher among cases than among non-cases (93.5 vs. 82.7 percent, p = 0.0002).

In multivariable-adjusted analyses, HRs for glioma and, particularly glioblastoma were somewhat elevated for the highest quartiles of the anthropometric measures of interest (Table 1). There was a suggestion of a non-significant trend, particularly for the association of BMI with

Table 2
Association of self-reported weight at different ages with risk of glioma in the Observational Study component of the Women's Health Initiative (n = 91,150[a]).

	Glioma (n cases = 115)				Glioblastoma (n cases = 86)			
			HR[b]	95% CI			HR[b]	95% CI
BMI (kg/m^2) at age 18								
< 19.4	31	30,465	1.00	Ref.	25	30,465	1.00	Ref.
19.4 –21.3	45	30,822	1.28	0.81–2.03	33	30,822	1.11	0.71–1.72
≥21.3	39	29,731	1.28	0.80–2.04	28	29,731	1.01	0.63–1.61
P for linear trend			*0.33*				*0.97*	
BMI (kg/m^2) at age 35								
< 20.8	37	30,352	1.00	Ref.	28	30,352	1.00	Ref.
20.8– < 22.9	43	30,122	1.14	0.73–1.77	31	30,122	1.11	0.71–1.72
≥22.9	35	30,472	1.04	0.65-1.66	27	30,472	1.01	0.63–1.61
P for linear trend			*0.92*				*0.99*	
BMI (kg/m^2) at age 50								
< 22.1	39	30,254	1.00	Ref.	30	30,254	1.00	Ref.
22.1– < 25.1	47	30,328	1.21	0.79–1.86	35	30,328	1.17	0.72–1.92
≥25.1	28	30,454	0.84	0.51–1.39	21	30,454	0.87	0.49–1.56
P for linear trend			*0.46*				*0.60*	

Abbreviations: HR – hazard ratio; 95% CI – 95% confidence interval; BMI – body mass index.
[a] Women with self-reported weight earlier in life.
[b] Adjusted for age, smoking status, alcohol intake, physical activity, years of education, and ethnicity.

glioblastoma and for WHR with glioma and glioblastoma. WC showed no association with glioma, and HRs for glioblastoma were similar for quartiles 2 to 4, showing no trend. None of the point estimates was statistically significant, and none of the linear trends over quartiles was significant. Associations with self-reported body weight at earlier points in life (ages 18, 35, and 50), available on a subset of the population (WHI observational study), showed no suggestion of an increasing trend with either glioma or glioblastoma (Table 2). When measured BMI, WC, and WHR were reanalyzed in this subgroup, no suggestive associations or trends were seen for glioma or glioblastoma.

In the sensitivity analysis excluding women with prevalent cancer, the associations with measured BMI, WC, and WHR with glioma and glioblastoma were either unchanged or slightly attenuated (data not shown). In particular, the monotonic trend for WHR with glioblastoma was weakened (HR for 2nd to 4th quartiles: 1.32, 95% CI 0.82-2.14, 1.27, 95% CI 0.77-2.07, 1.35, 95% CI 0.82-2.23, respectively; p for linear trend 0.30).

4. Discussion

In this large prospective study of postmenopausal women, there was a suggestion of a modest and non-significant positive association of measured BMI and WHR, but not WC, with risk of glioma and glioblastoma. Self-reported BMI earlier in life showed no association with risk of adult glioma or glioblastoma.

Previous studies have found little evidence of an association of adiposity with glioma. A meta-analysis [9] of 5 studies (4 cohort, 1 case-control) with a total of 2,725 cases of glioma, showed that BMI was not associated with glioma: the summary relative risk (RR) for overweight vs. normal weight was 1.06 (95% CI 0.94-1.20) and RR for obesity was 1.11 (95% CI 0.98-1.27). The second meta-analysis [8], which included only 3 studies (2 cohort and 1 case-control) with 2,418 cases of glioma, reported a positive summary association of BMI in females (odds ratio/relative risk 1.17, 95% CI 1.03-1.32), but not in males. However, this result appears to stem from two errors, which appear in Fig. 5 of the publication [8]. First, the meta-analysis included results for "BMI at age 21," rather than "BMI in recent past" from the Little et al. study [5]. BMI at age 21 showed a borderline significant positive association, whereas BMI in recent past showed no association (see Table 2, p. 1029 [5]). It is BMI in midlife that is the focus of the meta-analyses. Second, the authors included data from an analysis of the Nurses' Health Study I by Holick et al. [15]. However, this paper presented data on intake of fruits, vegetables, and carotenoids, but not on BMI, in relation to glioma risk. (GCK contacted both the first and second authors on the paper, and they both confirmed that their data on BMI and glioma risk were not published). Therefore, the source of the data for the two entries from the Holick et al. study ("Holick – NHS I, females, 2007, 25.0-29.9" and "Holick – NHS I, females, 2007, ≥30.0") is unclear.

Since publication of the meta-analyses, results from a large Norwegian cohort study [10] with 4,382 cases of glioma showed no association of overweight or obesity with glioblastoma or any other glioma subgroup; however, information on socioeconomic status and other covariates was not available in this study.

Fewer studies have examined BMI in adolescence in relation to risk of adult glioma [4,5]. The large relative risk for obesity relative to normal BMI reported by Moore et al. [4] was based on 11 glioma cases, and there was no suggestion of an elevated risk in the overweight category. In the case-control study by Little et al. [5], none of the odds ratios for the 25-29.9 or ≥30 kg/m^2 categories was statistically significant in males or females, although the trend per unit increase in BMI was statistically significant in females but not in males.

In the present study, there was a suggestion of increased risk, particularly for glioblastoma in association with BMI and WHR, although the associations were not statistically significant due to small numbers. There was no evidence of an increasing trend for the association of WC with glioma or glioblastoma, in spite of the fact that measured WC is a reliable indicator of central adiposity [16]. When we used self-reported body weight at earlier time points, available on roughly half the study population, there were no clear trends with glioma or glioblastoma. This was also true when associations of measured BMI, WC, and WHR were examined in this subgroup. However, women in the observational study tended to have lower BMI compared to women in the clinical trials, and this could have obscured a positive association.

It should also be mentioned that gliomas are a heterogeneous group of tumors of different histopathologic types and different grades (1). Our results are driven by the results for glioblastoma, the largest single subgroup. However, the numbers of other types of glioma (diffuse astrocytomas, analplastic astrocytomas, pilocytic astrocytomas, and oligodendrogliomas), were too small to analyze separately.

Strengths of the present study include the prospective nature of the study, central adjudication of all malignancies, standardized measurement of anthropometric factors at enrollment, and the availability of self-reported weight at earlier periods of life. Limitations include the

relatively small number of cases and the fact that, due to restriction of the study to postmenopausal women, glioma cases occurring in women below age 50 were not captured.

In conclusion, we found a suggestion of a modest positive association of measured BMI and WHR, but not WC, with risk of glioma and glioblastoma. Self-reported BMI earlier in life showed no association with increased risk. Based on available evidence, it is unlikely that adiposity represents an important risk factor for glioma.

Author's contributions

Conceived the study: Geoffrey Kabat, Thomas Rohan.
Designed the study: Geoffrey Kabat, Thomas Rohan.
Acquired the data: Geoffrey Kabat.
Analyzed the data: Geoffrey Kabat.
Wrote the first draft of the manuscript: Geoffrey Kabat.
Commented on and contributed additional ideas: Thomas Rohan.
Approved the final manuscript and conclusions: Geoffrey Kabat, Thomas Rohan.

Declaration of conflicting interests/disclosure

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

[1] Q.T. Ostrom, L. Bauchet, F.G. Davis, I. Deltour, J.L. Fisher, C.E. Langer, et al., The epidemiology of glioma in adults: a "state of the science" review, Neuro-Oncology 23 (2014) 896–913.
[2] M. Wrensch, Y. Minn, T. Chew, M. Bondy, M.S. Berger, Epidemiology of primary brain tumors: current concepts and review of the literature, Neuro-Oncology 4 (2002) 278–299.
[3] V.S. Benson, K. Pirie, J. Green, D. Casabonne, V. Beral, Lifestyle factors and primary glioma and meningioma tumours in the Million Women Study cohort, Br J Cancer 99 (2008) 185–190.
[4] S.C. Moore, P. Rajaraman, R. Dubrow, A.S. Darefsky, C. Koebnick, A. Hollenbeck, et al., Height, body mass index, and physical activity in relation to glioma risk, Cancer Res 69 (2009) 8349–8355.
[5] R.B. Little, M.H. Madden, R.C. Thompson, J.J. Olson, R.V. LaRocca, E. Pan, et al., Anthropometric factors in relation to risk of glioma, Cancer Causes Control 24 (2013) 1025–1031.
[6] D.S. Michaud, G. Bové, V. Gallo, B. Shclehofer, A. Tjønneland, A. Olsen, et al., Anthropometric measures, physical activity, and risk of glioma and meningioma in a large prospective cohort study, Cancer Prevention Res 4 (2011) 1385–1392.
[7] M. Wiedmann, C. Brunborg, K. Lindmann, T.B. Johannesen, L. Vatten, E. Helseth, et al., Body mass index and the risk of meningioma, glioma and schwannoma in a large prospective cohort study (the HUNT study, Br J Cancer 109 (2013) 289–294.
[8] T.N. Sergentanis, G. Tsivgoulis, C. Perlepe, I. Ntanasis-Stathopoulos, I. Tzanninis, I.N. Sergentanis, Obesity and risk for brain/CNS tumors, gliomas and meningiomas: a meta-analysis, PLoS ONE (2015), http://dx.doi.org/10.1371/journal.pone.0136974.
[9] T. Niedermaier, G. Behrens, D. Schmid, I. Schlecht, B. Fischer, M.F. Leitzmann, Body mass index, physical activity, and risk of adult meningioma and glioma: a meta-analysis, Neurology 85 (2015) 1342–1350.
[10] M.K.H. Wiedmann, C. Brunborg, A. Di Ieva, K. Lindemann, T.B. Johannesen, L. Vatten, et al., The impact of body mass index and height on the risk for glioblastoma and other glioma subgroups: a large prospective cohort study, Neuro-Oncol 19 (2016) 976–985.
[11] W.A. Banks, J.B. Owen, M.A. Erikson, Insulin in the brain: there and back again, Pharmacol Therapeutics 136 (2012) 82–93.
[12] Women's Health Initiative Study Group, Design of the Women's Health Initiative clinical trial and observational study, Control Clin Trials 19 (1998) 61–109.
[13] A. McTiernan, C. Kooperberg, E. White, S. Wilcox, R. Coates, L.L. Adams-Campbell, et al., Recreational physical activity and the risk of breast cancer in postmenopausal women: the Women's Health Initiative cohort study, JAMA 290 (2003) 1331–1336.
[14] J.D. Curb, A. McTiernan, S.R. Heckbert, C. Kooperberg, J. Stanford, K.C. Johnson, et al., Outcomes ascertainment and adjudication methods in the Women's Health Initiative, Ann Epidemiol 13 (Suppl. 9) (2003) S122–S128.
[15] C.N. Holick, E.L. Giovannucci, B. Rosner, M.J. Stampfer, D.S. Michaud, Prospective study of intake of fruit, vegetables, and carotenoids and the risk of adult glioma, Am J Clin Nutr 85 (2007) 877–886.
[16] E.V. Bandera, S.H. Fay, E. Giovannucci, M.F. Leitzmann, R. Marklew, A. McTiernan, et al., The use and interpretation of anthropometric measures in cancer epidemiology: a perspective from the world cancer research fund international continuous update project, Int J Cancer 139 (2016) 2391–2397.

Preventive Medicine Reports 4 (2016) 453–458

Contents lists available at ScienceDirect

Preventive Medicine Reports

journal homepage: http://ees.elsevier.com/pmedr

Support or competition? How online social networks increase physical activity: A randomized controlled trial☆

Jingwen Zhang PhD, Devon Brackbill PhD, Sijia Yang MA, Joshua Becker MA, Natalie Herbert MA, Damon Centola PhD *

Annenberg School for Communication, University of Pennsylvania, 3620 Walnut Street, Philadelphia, PA 19104, United States

ARTICLE INFO

Article history:
Received 13 April 2016
Received in revised form 20 June 2016
Accepted 7 August 2016
Available online 8 August 2016

Keywords:
Exercise
Internet
Network
Online social networks
Social comparison

ABSTRACT

To identify what features of online social networks can increase physical activity, we conducted a 4-arm randomized controlled trial in 2014 in Philadelphia, PA. Students (n = 790, mean age = 25.2) at an university were randomly assigned to one of four conditions composed of either supportive or competitive relationships and either with individual or team incentives for attending exercise classes. The social comparison condition placed participants into 6-person competitive networks with individual incentives. The social support condition placed participants into 6-person teams with team incentives. The combined condition with both supportive and competitive relationships placed participants into 6-person teams, where participants could compare their team's performance to 5 other teams' performances. The control condition only allowed participants to attend classes with individual incentives. Rewards were based on the total number of classes attended by an individual, or the average number of classes attended by the members of a team. The outcome was the number of classes that participants attended. Data were analyzed using multilevel models in 2014. The mean attendance numbers per week were 35.7, 38.5, 20.3, and 16.8 in the social comparison, the combined, the control, and the social support conditions. Attendance numbers were 90% higher in the social comparison and the combined conditions (mean = 1.9, SE = 0.2) in contrast to the two conditions without comparison (mean = 1.0, SE = 0.2) ($p = 0.003$). Social comparison was more effective for increasing physical activity than social support and its effects did not depend on individual or team incentives.

1. Introduction

Physical inactivity significantly increases the risk of chronic disease (Lee et al., 2001; Sattelmair et al., 2011) and mortality (Nocon et al., 2008; Wen et al., 2011). Low levels of physical activity among young adults remains a serious nationwide problem, with 69% of Americans 18 to 24 years of age failing to meet the federal guidelines for physical activity in 2014 (National Center for Health Statistics, 2015). Among all the social and environmental factors affecting physical activity (Addy et al., 2004; Martin and Savla, 2011), interpersonal social networks are one of the most prominent targets for cost-effective interventions (Maher et al., 2015). Online social networks, in particular, have become a highly attractive target for large scale health initiatives (Centola, 2013; Cobb and Graham, 2012); however, there is insufficient knowledge about why online networks might be effective sources of social influence for improving physical activity levels. One prominent argument in the literature on networks and health suggests that online relationships improve physical activity through supportive interactions that encourage healthy behaviors (Centola, 2010, 2011). An alternative approach stresses peer competition within online networks, emphasizing the value of social comparison as a mechanism for increasing individuals' receptiveness to positive behavioral influences (Foster et al., 2010). We evaluate the effects of each of these approaches independently, and in combination, to determine how social motivations for behavior change directly impact people's exercise activity.

Social support is one of the most widely used and studied strategies for encouraging behavior change in social networks (Berkman et al., 2000). When people with similar interests interact to achieve a shared goal, social support can reduce the perceived costs of adopting a new exercise routine by providing companionship in the activity (Cavallo et al., 2014; Uchino, 2004). Further, social support reduces the uncertainty of exploring new exercises by providing access to relevant sources of peer information (Wing and Jeffery, 1999). Thus, cooperative online relationships, where people work towards the same health goals, can foster collective efficacy for improving everyone's levels of physical activity (Cohen et al., 2006).

While social support via cooperative relationships may promote physical activity, an alternative approach utilizes social comparison via

☆ This study is registered at www.clinicaltrials.gov: NCT02267369.

* Corresponding author.
E-mail address: dcentola@asc.upenn.edu (D. Centola).

http://dx.doi.org/10.1016/j.pmedr.2016.08.008

competitive social relations (Foster et al., 2010; Zhang et al., 2015). Social comparison strategies are implicit in fitness and exercise programs that use rankings, leader boards, and social status markers to increase physical activity (Festinger, 1954). In these competitive environments, people work towards their goals individually, and differences in goal attainment motivate individuals to adjust their aspirations upward. The dynamic process of comparing oneself to others increases everyone's expectation for goal attainment and eventually improves overall levels of physical activity (Leahey et al., 2012; Shakya et al., 2015).

The number of online social network health interventions has increased dramatically in recent years (Laranjo et al., 2015; Maher et al., 2014; Williams et al., 2014). However, the independent causal effects as well as the interaction effects of these two contrasting approaches have not been identified (Cavallo et al., 2012; Napolitano et al., 2013; Neiger et al., 2012). As a result, there is very little guidance as to how these approaches might be used in applied settings to maximize social resources for increased fitness. This is particularly striking in light of recent meta-analyses of online social network health interventions, which have been inconclusive both on identifying which approaches are most effective, and regarding whether there are any systematic network strategies that can reliably be used to promote physical activity (Maher et al., 2014).

These problems of identification are compounded by the fact that the vast majority of research on online social networks and behavior change supplements social motivations with non-social incentives, such as health education and behavior tracking (Korda and Itani, 2013; Williams et al., 2014). This introduces interaction effects that prevent the identification of how, or whether, social factors can directly motivate behavior change. These shortcomings raise serious theoretical difficulties for developing consistent and replicable theories of how online social networks impact physical activity. They also limit the ability to apply online social network interventions to specific behavioral settings, where clear guidelines are required in order to implement effective interventions. We addressed these problems by conducting a double-blind four-arm randomized controlled study that compared the effects of social support and social comparison on increasing physical activity.

2. Methods

2.1. Study design

An 11-week online social network-based exercise program called SHAPE-UP was conducted at a Northeastern university. The program offered 90 exercise classes. On average, eight classes were offered per week on the University's campus, and each class lasted for an hour. Class content covered both aerobic and muscle-strengthening physical activities, including running, spinning, yoga, Pilates, weight lifting, high intensity interval training, and group exercising. All classes were led by instructors from the Department of Recreation and Health Services (DRHS) at the University. Participation in all classes was restricted to the program participants. At the conclusion of the program, participants were rewarded with gift cards for their participation based on the cumulative number of exercise classes they attended.

All participants in the program received access to the SHAPE-UP website, which was the only way for participants to enroll in classes and to interact with the program. Each participant created an online profile including username, gender, age, and their University affiliation. All participants had continuous and equal access to the website. To register for an exercise class, participants selected available classes from an interactive calendar that provided a brief class description and a registration tool. Upon registering, participants immediately received a confirmation email, and a reminder email was sent 12 h before each class started. An online tracking tool provided all participants with a daily journal of their exercise classes.

Upon logging into the website for the first time, participants were randomly assigned to one of four experimental conditions. Fig. 1 illustrates the four experimental conditions and Table 1 summarizes the different intervention components of the four conditions. Participants in the control condition were given the basic website for registering for classes. The control participants were provided with no social motivations, and were rewarded at the end of the program based on their individual record of attendance at exercise classes. The top 10% of participants in the condition were rewarded $20 gift cards at the end of the program. Three different experimental manipulations supplemented the control condition by providing online peer networks with different social incentives that might increase participation.

The *social comparison* condition supplemented the basic class registration website by giving participants access to 6-person peer networks. Each participant in this condition was randomly assigned 5 peers, which comprised 5 members of the study who were connected to the participant in a program-generated social network. Participants in this condition were able to compare their performance in the program with their peers via a competitive ranking based on their peers' activity levels. As in the control condition, at the conclusion of the program, the rewards for participants were based on each participant's individual record of class attendance. The top 10% of participants in the condition were rewarded $20 gift cards at the end of the program. All peers' information was anonymous, and there was almost no possibility for direct communication between participants in this condition online or offline.

By contrast, the *social support* condition was designed to provide participants with direct peer support from other members of the program who could encourage each other to improve their levels of regular exercise. Participants in this condition were randomly assigned to 6-person teams. At the completion of the program, rewards were based on the team's collective activity levels, incentivizing team members to actively support each other's attendance at exercise classes. All members in the top 10% of teams in the condition were rewarded $20 gift cards at the end of the program. To facilitate supportive social interaction, participants in the social support condition were provided with a chatting tool that they could use to directly communicate with each other in real-time. Team members could see both each other's individual records of class attendance as well as the collective record of the team. Participants in this condition were able to register for classes individually, but they could also coordinate to register for classes collectively.

Finally, to understand if there was an interaction effect of combining the motivations of social support and social comparison, the *combined* condition randomly placed individuals on 6-person teams and provided the same team incentives and technologies as the social support condition; however this condition was supplemented with a competitive feature, in the form of an interface that allowed participants to compare their team's performance against the performances of 5 other teams. All members in the top 5% of teams in the condition were rewarded $20 gift cards at the end of the program.

In these three conditions with online networks, participants in the same network also received real-time web and email notifications about their peers' registration and attendance of classes. For instance, when a network member attended a class, all of her peers would receive a notification about her class attendance. This signaling system was identical for all online networks across conditions.

2.2. Study participants

The SHAPE-UP program was open to all graduate and professional students at the University who were 18 years or older. Participants were recruited through advertisements on the University's website, through the student email list, via advertisements from the graduate student association, and with paper flyers placed on billboards around campus. The recruitment materials specified that the purpose of the project was to improve participants' quality of life through better fitness. Eligibility for enrollment in the study was determined by a physical assessment

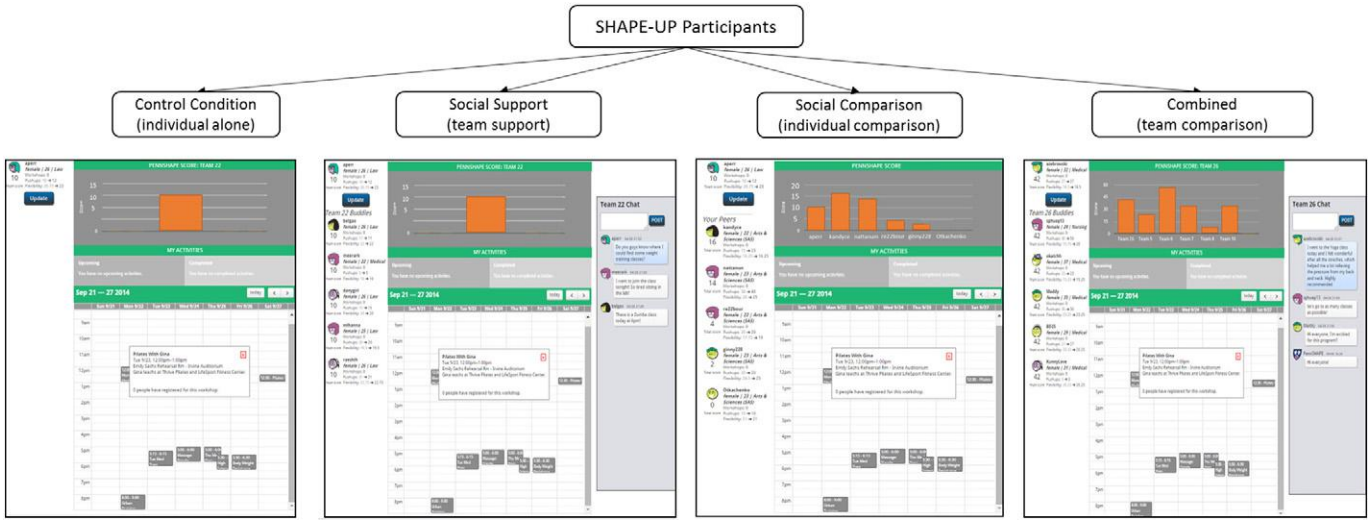

Fig. 1. Example webpage illustrations for the four experimental conditions in the trial, Philadelphia, PA, 2014.

conducted by the DRHS. Each participant completed a screening questionnaire (Canadian Society for Exercise Physiology (CSEP), 2002) designed to identify adults for whom physical activity might be inappropriate. The assessment lasted for 10 min and also measured participants' physical endurance, strength, and Body Mass Index (BMI).

Participant enrollment and assessments were conducted from August through September 2014. Eligible participants completed a baseline online survey assessing demographic and organizational information including gender, age, race, University department, and gym membership. In addition, we assessed participants' baseline physical activity with 3 items the Centers for Disease Control and Prevention (2001) developed concerning the number of days in which people participate in vigorous physical activity for at least 20 min, moderate activity for at least 30 min, and strength-building activities in the past 7 days. Participants were defined as meeting the 2008 federal physical activity guideline (Department of Health and Human Services, 2008) if they engaged in strength-building activity on at least 2 days and engaged in either 20 min of vigorous activity on at least 4 days or 30 min of moderate activity on at least 5 days.

In the RCT design, computer-generated random number sequences were used to randomly assign participants to the 4 experimental conditions after the baseline assessments. The random assignments were generated using the statistical software R, version 3.1.2.

Classes were held from September 2014 through December 2014. Participants and class instructors were blind to experimental assignments. Data collection was completed by December 2014. The study was approved by the institutional review board of the University, and all participants provided informed consent.

2.3. Outcome

The outcome of interest was the total number of exercise classes that participants attended throughout the 11-week program. Complete attendance data for all classes were provided by class instructors. Instructors collected individual attendance data with student emails and entered them into an online database. The attendance records then automatically showed up on participants' websites in real time. Attendance data were collected on site of each exercise class.

2.4. Statistical analysis

A sample size of 688 was planned because 172 participants per condition could achieve at least 90% power to detect a small to medium effect size of 0.35 (Cavallo et al., 2012; Foster et al., 2010; Zhang et al., 2015) in class attendance difference at the 5% significance level. The preliminary analysis consisted of an analysis of variance to examine the effects of social support and social comparison on the outcome, class attendance. However, it did not account for data clustering in the online networks. In each of the conditions with online networks, individuals received the treatment as members of a fully-connected network of 6 individuals, thus the primary analyses employed a multilevel regression model to account for the clustering of the treatment within these groups. The multilevel model included the social support and the social comparison factors, the support × comparison interaction, and covariates of baseline demographics. All analyses used the intention-to-treat principle and considered all participants who were randomly assigned to a condition, regardless of whether or not they received the treatment. All analyses were conducted in R, version 3.1.2.

3. Results

Of the 1007 participants who registered for the program, 790 attended the fitness evaluation and were randomly assigned to a condition. Fig. 2 shows the flow of participants. A total of 750 participants received at least one treatment exposure, as indicated by logging-in to the website. The attrition rates for participants receiving the treatment

Table 1
Intervention components in the four experimental conditions in the trial, Philadelphia, PA, 2014.

Intervention components	Control	Support	Comparison	Combined
SHAPE-UP website with an interactive calendar for class registration	x	x	x	x
Online networks with real-time web and email notifications of peer activities		x	x	x
Online chatting tool		x		x
Access to performance rankings of other peers or other teams			x	x
Rewards based on *individual* performance	x		x	
Rewards based on *team* performance		x		x

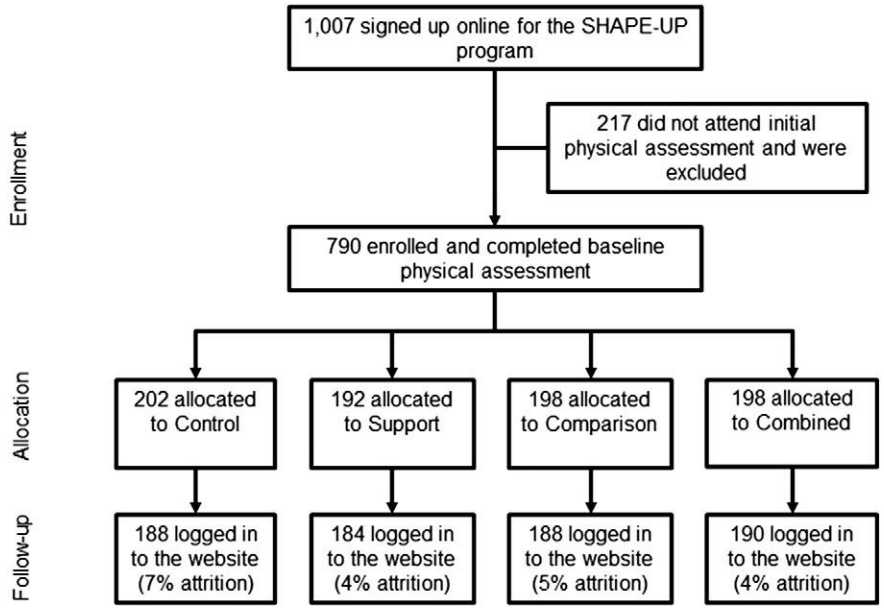

Fig. 2. Flow diagram of participants through the trial.

were statistically indistinguishable across all conditions, with 95% of all participants receiving the treatment.

Table 2 shows participants' characteristics. Participants ranged in age from 20 to 59 years (mean = 25.2, SD = 3.4), and ranged in BMI from 16.1 to 45.0 (mean = 23.0, SD = 3.8). In total, 454 (57.47%) participants did not meet the physical activity guideline. There were no significant differences in participants' characteristics at baseline across conditions.

Participants showed high levels of engagement with the website, averaging 22.8 logins (SD = 47.2) across all conditions during the program. Participants in the social support condition generated 81 online messages and participants in the combined condition generated 80 messages. The majority of the messages provided informational support regarding participants' plans for attending classes, their experiences, and opinions about the classes.

The number of exercise classes that each participant attended ranged from 0 to 39 classes. Attendance rates were 90% higher in the social comparison and the combined conditions (mean = 1.9, SE = 0.2) in contrast to the two conditions without social comparison (mean = 1.0, SE = 0.2). Both the social comparison and the combined conditions had significantly higher mean attendance rates (mean = 1.9, SE = 0.3 and mean = 1.9 SE = 0.2, respectively) than the control (mean = 1.1, SE = 0.3), while social support surprisingly performed worse (mean = 0.9, SE = 0.2). An analysis of variance shows that the presence of social comparison significantly increased activity levels ($F = 8.96$, $p = 0.003$, Cohen's D = 0.21). In contrast, the presence of social support did not significantly affect participants' exercise levels ($F = 0.04$, $p = 0.85$, Cohen's D = 0.01). There was no interaction between the two factors ($F = 0.18$, $p = 0.67$).

Table 3 presents results of the multilevel models that accounted for network-level influences within each condition. On average, social comparison increased attendance by 82% [95% CI: −5%, 168%], or 0.97 classes per participant ($p = 0.07$). After adjusting for baseline covariates, social comparison increased attendances by 62% [95% CI: 5% to 119%], or 1.06 classes per participant ($p = 0.03$). In contrast, social support had no significant effect ($p = 0.68$). Additionally, the non-significant interaction between support and comparison suggests that social support did not contribute to the increased attendance rates in the combined condition. The success of the combined condition can be thus attributed to the effects of team-based social comparison.

The mean class attendances per week were 35.7, 38.5, 20.3, and 16.8 in the social comparison, the combined, the control, and the social

Table 2
Baseline demographic characteristics of participates per experimental condition, Philadelphia, PA, 2014.

Participants (N)	Total 790	Control 202	Support 192	Comparison 198	Combined 198	*p* values
Age (years; M [SD])	25.2 [3.4]	25.0 [2.7]	25.4 [3.5]	25.3 [3.8]	25.1 [3.5]	0.539[c]
Male sex (N [%])	25.9	27.2	26.5	28.3	21.7	0.455[d]
Body Mass Index (kg/m^2; M [SD])	23.0 [3.8]	22.9 [4.0]	22.8 [3.6]	23.4 [3.8]	22.9 [3.7]	0.259[c]
Overweight (BMI [25.0–29.9]; N [%])	124 [15.7]	26 [12.9]	30 [15.6]	37 [18.7]	31 [15.7]	0.465[d]
Obese (BMI ≥ 30; N [%])	42 [5.3]	14 [6.9]	9 [4.7]	10 [5.1]	9 [4.5]	0.547[d]
Met physical activity guideline (N [%])[a]	336 [42.5]	87 [43.1]	80 [41.7]	83 [41.9]	85 [42.9]	0.998[d]
Race (N [%])[b]						0.448[d]
White	352 [44.6]	88 [43.5]	85 [44.3]	100 [50.5]	79 [39.9]	
Black	58 [7.3]	14 [7.9]	12 [6.2]	13 [6.6]	19 [9.6]	
Hispanic	62 [7.8]	15 [7.4]	12 [6.3]	12 [6.1]	23 [11.6]	
Asian	287 [36.3]	73 [36.1]	76 [39.6]	61 [30.8]	77 [38.9]	

[a] Participants met the guideline if they engaged in strength-building activity on at least 2 days and engaged in either 20 min of vigorous activity on at least 4 days or 30 min of moderate activity on at least 5 days.
[b] The omitted race category is "Other."
[c] The *p* values were based on one-way analyses of variance.
[d] The *p* values were based on chi-squared tests.

Table 3
Multilevel models for the effects of experimental conditions on exercise class attendance, Philadelphia, PA, 2014.

	Class attendance	
	Unadjusted for covariates Estimate (95% CI)	Adjusted for covariates Estimate (95% CI)
Comparison	**0.97*** (−0.06, 2.00)	**1.06**** (0.08, 2.04)
Support	−0.17 (−1.21, 0.88)	−0.21 (−1.21, 0.79)
Comparison × support	0.18 (−1.38, 1.74)	0.06 (−1.43, 1.55)
Constant	**1.19***** (0.56, 1.82)	1.72 (−1.38, 4.82)
Observations (N)	790	789
Log likelihood	−2,309.55	−2,271.03
Akaike Inf. Crit.	4,631.10	4,594.05
Bayesian Inf. Crit.	4,659.14	4,715.49

Notes: Covariates included age, gender, race, department, and having a gym membership in the previous semester. The difference in sample sizes arises from one missing data point for age.
* Boldface indicates statistical significance at $p < 0.1$.
** Boldface indicates statistical significance at $p < 0.05$.
*** Boldface indicates statistical significance at $p < 0.01$.

support conditions. Fig. 3 plots cumulative class attendance in each condition. Both of the conditions with social comparison (i.e., comparison and combined) showed significantly higher levels of attendance each day, averaging 5.1 and 5.5 people attending exercise classes per day, respectively. Both were significantly higher than the average of 2.9 attendances per day ($p < 0.001$) in the control condition. Class attendance in the support condition grew at a significantly slower rate than in the control condition, averaging only 2.4 attendances per day ($p < 0.001$), suggesting that social support reduced daily exercise rates as compared to the control condition.

4. Discussion

Despite the proliferation of online social network health interventions, social influence processes underlying social network effects have remained poorly understood. In this study, we found that social comparison in online networks provided a significantly greater source of social incentives for increasing physical activity than social support. Exposing individuals to relevant reference points, whether those reference points were other individuals or other teams, increased responsiveness to the physical activity of their peers. As a result, attendance at exercise classes was greatest in conditions where individuals and teams were motivated to exercise through competitive social relationships.

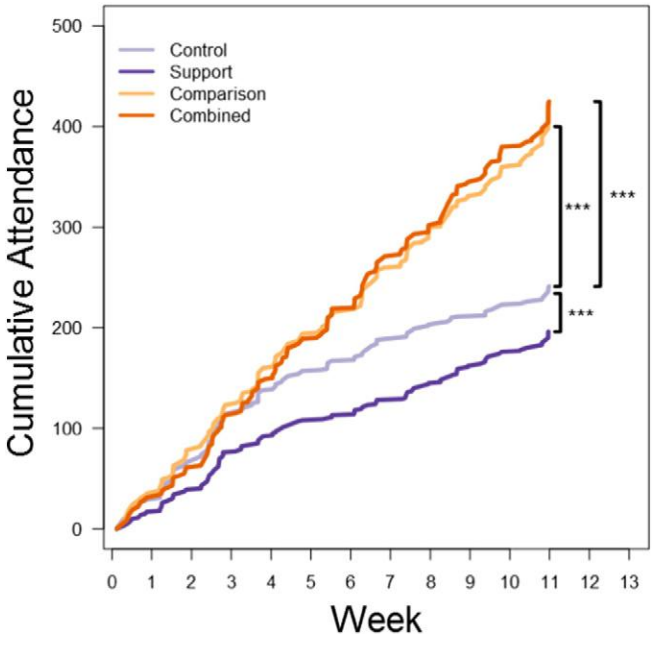

Fig. 3. Cumulative attendance at exercise classes in each of the four conditions. ***$p < 0.001$.

Physical activity in the social comparison and control conditions was incentivized based on individual rewards for exercising, while activity in the support condition was incentivized based on group rewards. The differences between the social comparison and social support conditions suggest that the observed reduction in activity in the support condition may be due to the ineffectiveness of collective rewards for motivating exercise. However, exercise levels were greatest in the combined condition, which also used collective rewards to incentivize participation. Moreover, the comparison condition and the combined condition generated nearly identical levels of exercise activity. This suggests that differences in incentives did not affect participants' exercise levels (Halpern et al., 2015; Jeffery et al., 1983; Leahey et al., 2012). Both individual and team rewards were equally effective for motivating activity.

A significant advantage of the four-arm experimental design is the ability to identify the independent and combined effects of social support and social comparison. A two-arm experiment that contrasted, e.g., the social support and social comparison conditions, or the combined condition with the control condition, would have been likely to produce incorrect inferences. For instance, in a two-arm experimental test between the combined condition and the control condition, it would appear as if social support (via team memberships) was effective at motivating increased participation. The four-way comparison between the control condition, social support, social comparison and the combined condition shows the opposite to be true – team memberships were effective for social comparison, but not for social support. It is only when the combined condition is compared independently with the social comparison condition that team memberships are shown not to produce any additional support-based motivations to exercise.

The strengths of this experimental design also entail limitations on the scope of the study. Most notably, while online networks typically evolve through endogenous social selection based on similar interests or characteristics, we artificially constructed the online networks used in this study. While this allowed us to ensure that there were no confounding effects from exogenous social information, it raises the potential concern that endogenous tie selection might create networks that were stronger sources for more effective generation of social support. In other words, our design might underestimate the effects of social support networks. Another limitation is that, although we created different social incentives in the experimental conditions, we did not measure psychological variables after the intervention, including perceived social support and social comparison. Thus, we were not able to conduct mediation analyses to test whether the intervention effects were mediated through hypothesized mechanisms. It was possible that the comparison incentive also increased participants' perceived descriptive social norms regarding physical activity as participants might pay more attention to other peers' efforts under the competitive mindset. Similarly we were not able to quantify the extent of class registration coordination among team members. Future research can extend our design with mediation surveys to identify the hypothesized psychological variables and with process evaluations to quantify participants' engagement with different components of the program. Finally, findings from this study sample are based on young adults from one university and may not be generalizable to other segments of the broader population. Future research can apply this design to study other population samples in other social contexts, as well as to study additional outcomes measures (e.g., body weight and muscle strength) that might be studied over longer periods of time.

Our results suggest that networks that emphasize social comparison among members can be surprisingly effective for motivating desirable behaviors. The results from the combined condition, where adding team performances to a supportive environment significantly increased exercise levels, suggest that the introduction of a minimal competitive reference point into an otherwise support-based environment can change ineffective health networks into highly motivating social resources. Healthcare providers, online fitness programs, and peer-to-peer communities for improving patient health all seek ways to structure social interactions among their members to provide the greatest incentives for adopting and maintaining health behaviors. Social comparison might be harnessed to address a variety of other health issues such as medication compliance, diabetes control, smoking cessation, flu vaccinations, weight loss, and preventative screenings (Centola, 2013). Future research can extend our approach to test whether the strong effect of social comparison through constructed online networks can be realized in health promotion domains other than physical activity.

Conflict of interest statement

The authors declare that there are no conflicts of interest.

Acknowledgments

This study was funded by the National Cancer Institute of the National Institutes of Health under Award Number P20CA095856.

References

Addy, C.L., Wilson, D.K., Kirtland, K.A., Ainsworth, B.E., Sharpe, P., Kimsey, D., 2004. Associations of perceived social and physical environmental supports with physical activity and walking behavior. Am. J. Public Health 94, 440–443.
Berkman, L.F., Glass, T., Brissette, I., Seeman, T.E., 2000. From social integration to health: Durkheim in the new millennium. Soc. Sci. Med. 51, 843–857.
Canadian Society for Exercise Physiology (CSEP), 2002. Physical Activity Readiness Questionnaire: PAR-Q, Canadian Society for Exercise Physiology, Ottawa, ON.
Cavallo, D.N., Tate, D.F., Ries, A.V., Brown, J.D., DeVellis, R.F., Ammerman, A.S., 2012. A social media-based physical activity intervention: a randomized controlled trial. Am. J. Prev. Med. 43, 527–532.
Cavallo, D.N., Brown, J.D., Tate, D.F., DeVellis, R.F., Zimmer, C., Ammerman, A.S., 2014. The role of companionship, esteem, and informational support in explaining physical activity among young women in an online social network intervention. J. Behav. Med. 37, 955–966.
Centers for Disease Control and Prevention, 2001. 2001 National School-based Youth Risk Behavior Survey. Public-use Data Documentation. Division of Adolescent and School Health, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, Atlanta, GA.
Centola, D., 2010. The spread of behavior in an online social network experiment. Science 329, 1194–1197.
Centola, D., 2011. An experimental study of homophily in the adoption of health behavior. Science 334, 1269–1272.
Centola, D., 2013. Social media and the science of health behavior. Circulation 127, 2135–2144.
Cobb, N.K., Graham, A.L., 2012. Health behavior interventions in the age of Facebook. Am. J. Prev. Med. 43, 571–572.
Cohen, D.A., Finch, B.K., Bower, A., Sastry, N., 2006. Collective efficacy and obesity: the potential influence of social factors on health. Soc. Sci. Med. 62, 769–778.
Department of Health and Human Services, 2008. 2008 Physical Activity Guidelines for Americans. Department of Health and Human Services, Washington, DC.
Festinger, L., 1954. A theory of social comparison processes. Hum. Relat. 7, 117–140.
Foster, D., Linehan, C., Kirman, B., Lawson, S., James, G., 2010. Motivating physical activity at work: using persuasive social media for competitive step counting. ACM 111–116.
Halpern, S.D., French, B., Small, D.S., et al., 2015. Randomized trial of four financial-incentive programs for smoking cessation. New Engl. J. Med. 372, 2108–2117.
Jeffery, R.W., Gerber, W.M., Rosenthal, B.S., Lindquist, R.A., 1983. Monetary contracts in weight control: effectiveness of group and individual contracts of varying size. J. Consult. Clin. Psychol. 51, 242–248.
Korda, H., Itani, Z., 2013. Harnessing social media for health promotion and behavior change. Health Promot. Pract. 14, 15–23.
Laranjo, L., Arguel, A., Neves, A.L., et al., 2015. The influence of social networking sites on health behavior change: a systematic review and meta-analysis. J. Am. Med. Inform. Assoc. 22, 243–256.
Leahey, T.M., Kumar, R., Weinberg, B.M., Wing, R.R., 2012. Teammates and social influence affect weight loss outcomes in a team-based weight loss competition. Obesity 20, 1413–1418.
Lee, I.M., Rexrode, K.M., Cook, N.R., Manson, J.E., Buring, J.E., 2001. Physical activity and coronary heart disease in women - is "no pain, no gain" passe? J. Am. Med. Assoc. 285, 1447–1454.
Maher, C.A., Lewis, L.K., Ferrar, K., Marshall, S., De Bourdeaudhuij, I., Vandelanotte, C., 2014. Are health behavior change interventions that use online social networks effective? A systematic review. J. Med. Internet Res. 16, e40.
Maher, C., Ferguson, M., Vandelanotte, C., et al., 2015. A web-based, social networking physical activity intervention for insufficiently active adults delivered via Facebook app: randomized controlled trial. J. Med. Internet Res. 17, e174.
Martin, E., Savla, T., 2011. A social ecological perspective for predicting physical activity outcomes: a structural equation model. Ann. Behav. Med. 41, S33.
Napolitano, M.A., Hayes, S., Bennett, G.G., Ives, A.K., Foster, G.D., 2013. Using Facebook and text messaging to deliver a weight loss program to college students. Obesity 21, 25–31.
National Center for Health Statistics, 2015. Early Release of Selected Estimates Based on Data from the National Health Interview Survey, 2014. National Center for Health Statistics, Hyattsville, MD.
Neiger, B.L., Thackeray, R., Van Wagenen, S.A., et al., 2012. Use of social media in health promotion purposes, key performance indicators, and evaluation metrics. Health Promot. Pract. 13, 159–164.
Nocon, M., Hiemann, T., Mueller-Riemenschneider, F., Thalau, F., Roll, S., Willich, S.N., 2008. Association of physical activity with all-cause and cardiovascular mortality: a systematic review and meta-analysis. Eur. J. Cardiovasc. Prev. Rehabil. 15, 239–246.
Sattelmair, J., Pertman, J., Ding, E.L., Kohl, H.W., Haskell, W., Lee, I.M., 2011. Dose response between physical activity and risk of coronary heart disease: a meta-analysis. Circulation 124, 789–795.
Shakya, H.B., Christakis, N.A., Fowler, J.H., 2015. Self-comparisons as motivators for healthy behavior. Obesity 23, 2477–2484.
Uchino, B.N., 2004. Social Support and Physical Health: Understanding the Health Consequences of Relationships. Yale University Press, New Haven.
Wen, C.P., Wai, J.P., Tsai, M.K., et al., 2011. Minimum amount of physical activity for reduced mortality and extended life expectancy: a prospective cohort study. Lancet 378, 1244–1253.
Williams, G., Hamm, M.P., Shulhan, J., Vandermeer, B., Hartling, L., 2014. Social media interventions for diet and exercise behaviours: a systematic review and meta-analysis of randomised controlled trials. BMJ Open 4, e003926.
Wing, R.R., Jeffery, R.W., 1999. Benefits of recruiting participants with friends and increasing social support for weight loss and maintenance. J. Consult. Clin. Psychol. 67, 132–138.
Zhang, J., Brackbill, D., Yang, S., Centola, D., 2015. Efficacy and causal mechanism of an online social media intervention to increase physical activity: results of a randomized controlled trial. Prev. Med. Rep. 2, 651–657.

Frontiers in Neuroendocrinology 52 (2019) 22–28

Contents lists available at ScienceDirect

Frontiers in Neuroendocrinology

journal homepage: www.elsevier.com/locate/yfrne

Exogenous melatonin as a treatment for secondary sleep disorders: A systematic review and meta-analysis

Tian Li[a,b,1], Shuai Jiang[c,1], Mengzhen Han[a], Zhi Yang[b], Jianjun Lv[b], Chao Deng[d], Russel J. Reiter[e,*], Yang Yang[a,b,*]

[a] Key Laboratory of Resource Biology and Biotechnology in Western China, Ministry of Education, Faculty of Life Sciences, Northwest University, 229 Taibai North Road, Xi'an 710069, China
[b] Department of Biomedical Engineering, The Fourth Military Medical University, 169 Changle West Road, Xi'an 710032, China
[c] Department of Aerospace Medicine, The Fourth Military Medical University, 169 Changle West Road, Xi'an 710032, China
[d] Department of Cardiovascular Surgery, The First Affiliated Hospital of Xi'an Jiaotong University, 277 Yanta West Road, Xi'an 710061, Shaanxi, China
[e] Department of Cellular and Structural Biology, UT Health Science Center, 7703 Floyd Curl Drive, San Antonio, TX 78229, USA

ARTICLE INFO

Keywords:
Melatonin
Secondary sleep disorders
Meta-analysis
Sleep onset latency
Total sleep time
Sleep efficiency

ABSTRACT

Melatonin is a physiological indoleamine involved in circadian rhythm regulation and it is currently used for secondary sleep disorders supported by empirical evidence. A small amount of evidence and some controversial results have been obtained in some randomized controlled trials (RCT). The objective of this meta-analysis is to determine the efficacy of exogenous melatonin versus placebo in managing secondary sleep disorders. Literature retrieval of eligible RCT was performed in 5 databases (PubMed, Embase, Cochrane Library, ClinicalTrials.gov, and Web of Science). In total, 7 studies of 205 patients were included. Pooled data demonstrate that exogenous melatonin lowers sleep onset latency and increases total sleep time, whereas it has little if any effect on sleep efficiency. Although, the efficacy of melatonin still requires further confirmation, this meta-analysis clearly supports the use of melatonin as a management for patients with secondary sleep disorders.

1. Introduction

Sleep disorders, or somnipathy, are serious public health problems that cause a prominent economic burden worldwide. More than 10% individuals in Western societies have sleep disorders (Auld et al., 2017; Gervais et al., 2017). Conventional drugs usually have a short half-life and a hangover effect that might contribute to a poor compliance (Walters and Lader, 1971). Implementation of non-pharmacotherapy such as cognitive and relaxation therapy is usually a complex process influenced by multiple factors. Thus, an exogenously-administered agent that mimics the actions of an endogenous molecule might serve to cure or improve sleep disorders.

Melatonin, discovered by Aaron Lerner, is an chronobiotic that modulates circadian rhythms (Suhner et al., 2001; Keijzer et al., 2014) and has a wide variety of other functions (Carrier et al., 2017; Reiter et al., 2017, 2014; Yang et al., 2014; Yu et al., 2018). In humans, melatonin enhances darkness-related behavior and induces soporific effects. For example, melatonin has been tested in secondary sleep disorders (namely secondary insomnia) caused by sleep restriction (e.g., shiftwork, jet lag) without organic diseases (Buscemi et al., 2006; Liira et al., 2014). Previous studies have demonstrated that melatonin has a hypnotic action on secondary sleep disorders (Sadeghniiat-Haghighi et al., 2016; Suhner et al., 2001; Yoon and Song, 2002) with no known adverse effects (Suhner et al., 2001; van Geijlswijk et al., 2010).

Primary sleep disorders are rarely markedly improved by exogenous drug treatment, as a result of the existence of protopathy (Zhang et al., 2016). A previous meta-analysis by Buscemi reported that melatonin is ineffective in relieving sleep problems. Buscemi et al. (2006) explored sleep disorders resulting from organic and non-organic factors and concluded that melatonin has no effects, which contradicts the conclusions of some randomized controlled trials (RCT) (Sadeghniiat-Haghighi et al., 2016; Suhner et al., 2001; Wright et al., 1998). Sample sizes of previously published studies were usually small and the results were inconsistent (Beaumont et al., 2004; Folkard et al., 1993; Suhner et al., 2001). However, some of the recent findings published suggest that melatonin is a potent drug candidate for secondary sleep disorders (Sadeghniiat-Haghighi et al., 2016; Suhner et al., 2001; Yoon and Song, 2002). Thus, an updated meta-analysis with a different focus is needed

* Corresponding authors at: Key Laboratory of Resource Biology and Biotechnology in Western China, Ministry of Education, Faculty of Life Sciences, Northwest University, 229 Taibai North Road, Xi'an 710069, China (Y. Yang).
E-mail addresses: reiter@uthscsa.edu (R.J. Reiter), yang200214yy@nwu.edu.cn (Y. Yang).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.yfrne.2018.06.004
Received 17 February 2018; Received in revised form 7 June 2018; Accepted 13 June 2018
Available online 15 June 2018
0091-3022/

to provide the latest evidence for clinical psychiatrists and neurologists. This systematic review summarizes the current data that investigated the roles of exogenous melatonin, versus placebo, in the treatment of secondary sleep disorders.

2. Methods

2.1. Study protocol

This meta-analysis was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Moher et al., 2009).

2.2. Search strategy

A systematic search was performed on PubMed, Embase, Cochrane Register for Systematic Reviews databases, ClinicalTrials.gov, and Web of Science with no language limitations (published between Jan 1, 1950 and Sep 19, 2017). Search terms included “melatonin”, or “5-methoxytryptamine”, or “ramelteon” in combination with “sleep disorders”, or “sleep disturbance”, or “sleep dysfunction”. For example, the search strategy in PubMed was as follows: “((melatonin[MeSH Terms] OR methoxytryptamine[Text Word])) AND (sleep disorder[MeSH Terms] OR secondary sleep disorder[Text Word] OR sleep dysfunction[Text Word] OR sleep disturbance[Text Word] OR sleep[Text Word])) AND (random allocation[MeSH Terms] OR randomized[Text Word] OR randomly[Text Word] OR placebo[Text Word] OR crossover[Text Word] OR cross-over[Text Word])”. Moreover, the references of relevant studies, reviews, editorials, letters, and conference abstracts were also searched. The research work was done independently and in duplicate.

2.3. Eligibility criteria

Studies meeting the following criteria were included: (a) study design is double-blind RCT (crossover or parallel); (b) populations were adult patients with secondary sleep disorders secondary accompanying sleep restriction; (c) intervention comparisons of melatonin versus placebo; (d) outcomes analyzed for nocturnal sleep without diurnal disturbance; (e) enrolled studies report any or all of the following nocturnal sleep outcomes: I. sleep onset latency (the length of time between full wakefulness to sleep onset), II. total sleep time (total time spent asleep in bed), III. sleep efficiency (total sleep time/total time spent in bed). Sleep onset latency was the primary outcome. Total sleep time and sleep efficiency were secondary outcomes.

The exclusion criteria were: (a) articles not peer-reviewed or published; (b) populations that included children (< 18 years); (c) studies that were repeatedly published or had qualitative outcomes; (d) outcomes that analyzed daytime instead of the night sleep. There was no limit on sample size, trial duration, et al. Preliminarily, retrieved results were subjected to a title and abstract screening for inclusion by two independent reviewers (Li & Jiang). Full-text retrieval was performed to determine eligibility using a standardized data abstraction form by the two independent reviewers. Disagreements regarding the inclusion of studies were discussed between the two reviewers (Li & Jiang) and ultimately decided by the third reviewer (Han).

2.4. Data extraction

Two reviewers (Li & Jiang) independently extracted relevant data with a standardized form, including study characteristics and main outcomes. Disagreement was resolved through discussion with a third author (Han). Characteristics of patients such as age, research site, sample size, study design, duration, dosage, and clinical outcomes were collected. Clinical outcomes included continuous variables of sleep onset latency (mean [SD]), total sleep time (mean [SD]), and sleep efficiency (mean [SD]). Two reviewers also independently assessed the risk of bias of enrolled studies using The Cochrane Collaboration’s tool for assessing risk of bias (Higgins et al., 2011). Three types of risk (low risk of bias, unclear risk of bias, and high risk of bias) were identified depending on the following domains: (a) random sequence generation (selection bias), (b) allocation concealment (selection bias), (c) blinding of participants and personnel (performance bias), (d) blinding of outcome assessment (detection bias), (e) incomplete outcome data (attrition bias), (f) selective reporting (reporting bias), (g) other biases. Disagreement was resolved by a third reviewer (Han).

2.5. Statistical analysis

The efficacy of exogenous melatonin versus placebo was evaluated on three continuous outcomes: sleep onset latency (mean [SD]), total sleep time (mean [SD]), and sleep efficiency (mean [SD]). We calculated results for continuous outcomes as mean differences (MD) with 95% confidence intervals (CI), comparing the change from baseline for both melatonin and placebo. All tests were two tailed and a P value of less than 0.05 was deemed statistically significant. Data were analyzed by the latest Cochrane collaboration Review Manager analysis software version 5.3 (The Nordic Cochrane Center, Copenhagen, Denmark). According to the knowledge from evidence-based medicine and Cochrane Handbook for Systematic Review of Interventions (version 5.1), the weight of enrolled studies depends on the value of mean, [SD], and total sample size. The Z-test determined an overall significance of therapeutic effect versus placebo with values of $P < 0.05$. Heterogeneity was assessed using chi-squared test and I^2 test in accordance with Cochrane collaboration's guidance for assessing heterogeneity in meta-analyses. Data were considered heterogeneous if chi-squared test yielded $P < 0.10$ and $I > 50\%$. Random effects model (REM) was utilized when heterogeneity was absent; otherwise, the fixed effects model (FEM) was chosen. If we identified sufficient trials ($N \geq 10$), a funnel plot was utilized to test potential publication bias.

3. Results

3.1. Study characteristics

The initial search by two reviewers (Li & Jiang) identified 1223 database records and 13 additional records (Fig. 1). 900 records remained after removing 336 duplicates. Then, the title and abstract of remained literature were screened, and 812 records were excluded due to review article, meta-analyses/systematic review, case-control studies, cross sectional studies, or unrelated topics. Thereafter, 88 full-text articles were assessed for eligibility and 81 records were excluded with reasons: children (n = 21), insufficient end points (n = 32), non-randomized studies (n = 10), and irrelevant reports (n = 18). Eventually, 7 RCT (Beaumont et al., 2004; Folkard et al., 1993; James et al., 1998; Sadeghniiat-Haghighi et al., 2016; Suhner et al., 2001; Wright et al., 1998; Yoon and Song, 2002) were included, their characteristics were listed in Table 1.

3.2. Systematic review

These 7 studies of 205 participants were conducted in USA (n = 4) (Beaumont et al., 2004; James et al., 1998; Suhner et al., 2001; Wright et al., 1998), England (n = 1) (Folkard et al., 1993), Republic of Korea (n = 1) (Yoon and Song, 2002), and Iran (n = 1) (Sadeghniiat-Haghighi et al., 2016), published between 1993 and 2016 (Table 1). All studies were double-blind RCT that used standard experimental/control groups (melatonin/placebo). The mean age of participants ranged from 29 (Folkard et al., 1993; James et al., 1998; Yoon and Song, 2002) to 41.3 (Suhner et al., 2001). The sample size ranged from 12 (Yoon and Song, 2002) to 74 (Suhner et al., 2001). Participants include medical staffs (James et al., 1998; Wright et al., 1998; Yoon and Song, 2002), police

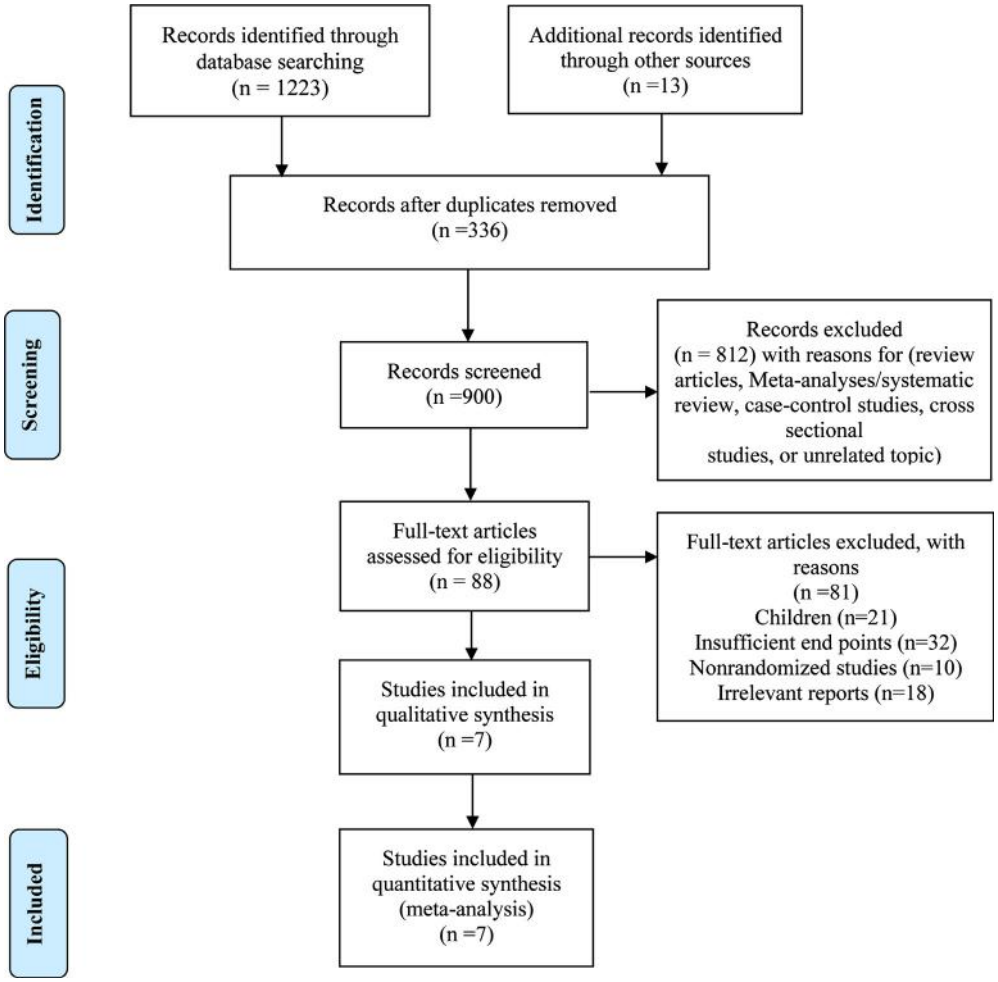

Fig. 1. Flowchart of study screening.

(Folkard et al., 1993), workers (Sadeghniiat-Haghighi et al., 2016) after shift-work, and pilots (Beaumont et al., 2004), travelers (Suhner et al., 2001) after jet lag.

All studies reported the efficacy of melatonin in sleep onset latency. 3 studies reported the efficacy of melatonin in total sleep time (Beaumont et al., 2004; Sadeghniiat-Haghighi et al., 2016; Yoon and Song, 2002). 3 studies reported the efficacy of melatonin in sleep efficiency (Beaumont et al., 2004; Sadeghniiat-Haghighi et al., 2016; Yoon and Song, 2002). The dosage of melatonin ranged from 3 (Sadeghniiat-Haghighi et al., 2016) to 6 mg (James et al., 1998; Yoon and Song, 2002), among which 5 mg was the most frequently used (Beaumont et al., 2004; Folkard et al., 1993; Suhner et al., 2001; Wright et al., 1998). The duration of treatment ranged from 3 (Sadeghniiat-Haghighi et al., 2016) to 9 days (Beaumont et al., 2004).

3.3. Sleep onset latency

Sleep onset latency is the primary outcome in this meta-analysis. All studies assessed the efficacy of melatonin in sleep onset latency. Pooled analysis of 7 studies (N = 154) demonstrate that exogenous administration of melatonin lowers sleep onset latency (Total mean difference: −2.48 min, 95% CI: −4.56, −0.40, Fig. 2), versus placebo. The overall estimated score of melatonin treatment is significant (Z = 2.33, P = 0.02). No significant heterogeneity was present across these studies (Chi^2 = 7.54, P = 0.27, I^2 = 20%).

3.4. Total sleep time

Total sleep time was the secondary outcome in this meta-analysis. 3 studies reported the efficacy of melatonin in total sleep time (Beaumont et al., 2004; Sadeghniiat-Haghighi et al., 2016; Yoon and Song, 2002). Analysis of the 3 studies (N = 71) suggest that exogenous administration of melatonin increases total sleep time (Total mean difference: 29.27 min, 95% CI: 6.68, 51.86, Fig. 3). The overall estimated action of melatonin is significant (Z = 2.54, P = 0.01). No significant heterogeneity was present across the 3 studies (Chi^2 = 2.71, P = 0.26, I^2 = 26%).

3.5. Sleep efficiency

Sleep efficiency was also the secondary outcome in this meta-analysis. 3 studies reported the efficacy of melatonin in sleep efficiency (Beaumont et al., 2004; Sadeghniiat-Haghighi et al., 2016; Yoon and Song, 2002). Analysis of the 3 studies (N = 71) suggested that exogenous melatonin has no meaningful actions on sleep efficiency (Total mean difference: 1.46, 95% CI: −0.43, 3.35, Fig. 4). The Z value of melatonin is 1.52 (P = 0.13). No significant heterogeneity was present across the 3 studies (Chi^2 = 2.03, P = 0.36, I^2 = 1%).

3.6. Risk of bias

Figs. 5 and 6 show the risk of bias across 7 trials. 100% studies have a low risk of bias on random sequence generation, blinding of participants and personnel, and selective reporting. 3 (Beaumont et al., 2004;

Table 1
Characteristic of included studies.

Study	Groups	Participants	Total number	Number (exp)	Age (years)	Research site	Study design	Duration of treatment (day)	Dosage (mg)	Outcomes	Reference
Beaumont et al, 2004	Melatonin/ Placebo	Pilots after a seven-time zone eastbound flight	18	9/9	35.3 ± 8.1 (19–47)	USA	RCT, parallel	9	5	1–3	Beaumont et al. (2004)
Folkard et al, 1993	Melatonin/ Placebo	Police officers after night shift	14	7/7	29 ± 7 (21–48)	England	RCT, parallel	7	5	1	Folkard et al. (1993)
James et al, 1998	Melatonin/ Placebo	Emergency medical staffs after night shift	22	22/22	29 ± 7 (20–41)	USA	RCT, crossover design	4	6	1	James et al. (1998)
Sadeghniiat-Haghighi et al, 2016	Melatonin/ Placebo	Shift workers	50	50/50	32.9 ± 8 (24–52)	Iran	RCT, crossover design	3	3	1–3	Sadeghniiat-Haghighi et al. (2016)
Suhner et al, 2001	Melatonin/ Placebo	Travelers after a jet lag	74	35/39	41.3 (18–68)	USA	RCT, parallel	4	5	1	Suhner et al. (2001)
Wright et al, 1998	Melatonin/ Placebo	Emergency physicians after night-shift work	15	15/15	38.6 (32–45)	USA	RCT, crossover design	4	5	1	Wright et al. (1998)
Yoon et al., 2002	Melatonin/ Placebo	Nurses after night-shift work	12	12/12	29 ± 7 (21–48)	Republic of Korea	RCT, crossover design	4	6	1–3	Yoon and Song (2002)

Suhner et al., 2001; Wright et al., 1998) and 4 (Folkard et al., 1993; James et al., 1998; Sadeghniiat-Haghighi et al., 2016; Yoon and Song, 2002) studies have a low and unclear risk of bias on allocation concealment, respectively. 1 (Wright et al., 1998) and 6 (Beaumont et al., 2004; Folkard et al., 1993; James et al., 1998; Sadeghniiat-Haghighi et al., 2016; Suhner et al., 2001; Yoon and Song, 2002) studies have a low and unclear risk of bias on blinding of outcome assessment, respectively. As for the incomplete outcome data, 2 (Beaumont et al., 2004; Yoon and Song, 2002), 2 (Folkard et al., 1993; Wright et al., 1998), and 3 (James et al., 1998; Sadeghniiat-Haghighi et al., 2016; Suhner et al., 2001) studies have a low, unclear, and high risk of bias on incomplete outcome data, respectively. 4 (Beaumont et al., 2004; Sadeghniiat-Haghighi et al., 2016; Suhner et al., 2001; Yoon and Song, 2002), 2 (Folkard et al., 1993; James et al., 1998), and 1 (Wright et al., 1998) studies have a low, unclear, and high risk of bias on other bias, respectively. Other bias (Higgins et al., 2011) might include the following situations: (1) There was no description of the ingested drug monitoring by physicians, which could result in performance bias. (2) The authors state any important concerns about bias not covered in the other domains in the tool. (3) Bias due to problems not covered elsewhere exists.

4. Discussion

4.1. Main findings

In this meta-analysis of 7 studies that analyzed exogenous melatonin versus placebo in the treatment of secondary sleep disorders, we found that melatonin reduces sleep onset latency and increases total sleep time (Figs. 2 and 3). However, based on current data, melatonin has no actions on sleep efficiency of patients with secondary sleep disorders (Fig. 4). Overall, these data demonstrate that melatonin improves sleep quality with respect to sleep onset latency and total sleep time, which lends support to melatonin as a potential approach to secondary sleep disorders.

4.2. Interpretation

Secondary sleep disorders, namely secondary insomnia caused by sleep restriction, have caused a huge economic and social burden in the world (Liira et al., 2014). The management of secondary sleep disorders remains a big issue in clinical psychiatry and sleep medicine. So far, there are no effective treatments for secondary sleep disorders. Pharmacotherapies usually have a short half-life and a hangover effect. Non-pharmacotherapies (e.g., cognitive therapy and relaxation therapy) are complex and patients usually have poor compliance with them. Therefore, the low toxicity may facilitate melatonin as a therapeutic candidate for secondary sleep disorders (Jan et al., 2009; Li et al., 2017).

This meta-analysis was conducted under the guidance of PRISMA (Moher et al., 2009). The search strategy of this meta-analysis is thorough and the inclusion criteria are broad. Both MeSH Terms and Text Word were utilized in databases from PubMed, Embase, Cochrane Library, et al., which includes a wide variety of publications from 1950 to 2017 (Robinson, 2005). Enrolled studies were analyzed by two authors (Li & Jiang) independently using unitive criteria and disagreement was solved through discussion with a third author (Han). Based on previous works (Chan et al., 2011; Nathan et al., 2017), this review also provides a general introduction for all enrolled studies in Section Systematic review and has mentioned Buscemi's work in Section Introduction. In Buscemi's work, secondary sleep disorders were clearly distinguished from those accompanying sleep restriction. In this review, secondary sleep disorders are referred to as sleep disorders accompanying sleep restriction (e.g. shift work and jet lag). Compared to Buscemi's work in 2006, this meta-analysis was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)

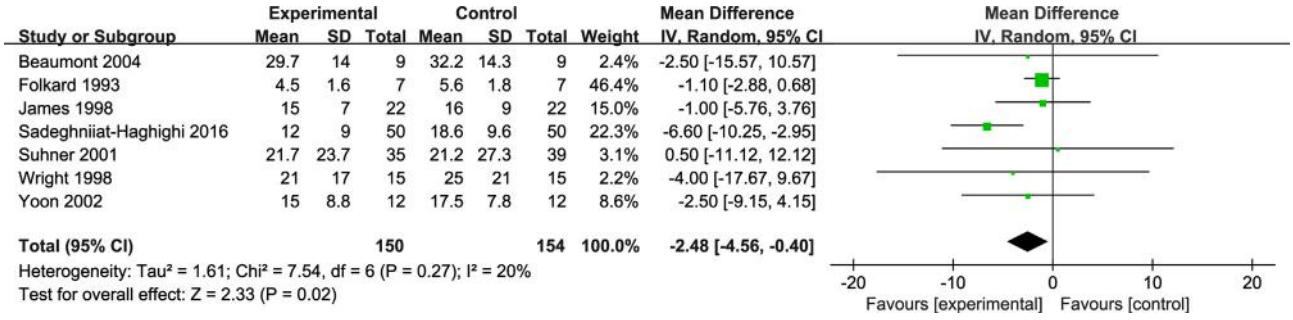

Fig. 2. Effects of melatonin on SOL. This forest plot demonstrates that exogenous administration of melatonin lowers sleep onset latency. SOL, sleep onset latency.

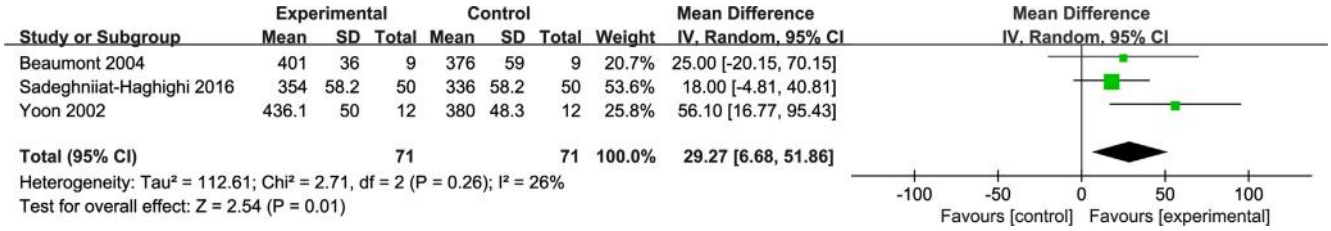

Fig. 3. Effects of melatonin on TST. This forest plot suggests that exogenous administration of melatonin increases total sleep time. TST, total sleep time.

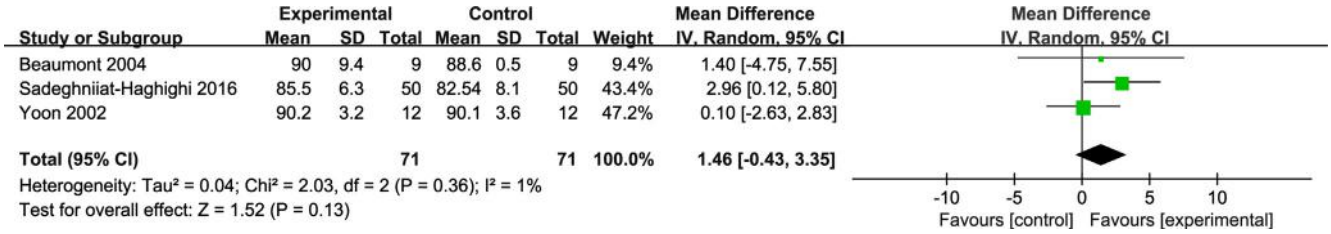

Fig. 4. Effects of melatonin on SE. This forest plot reveals that melatonin has no significant effects on sleep efficiency. SE, sleep efficiency.

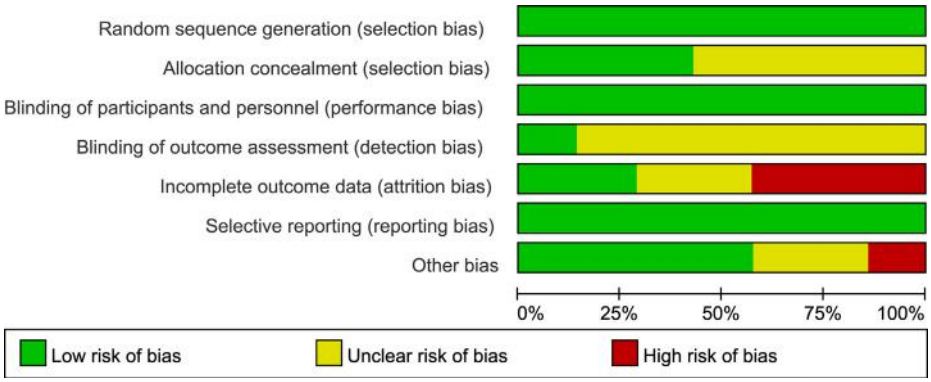

Fig. 5. Risk of bias graph. This figure shows review authors' judgements about each risk of bias item presented as percentages across all included studies.

and aimed to explore the new literature and provide a different emphasis. Thus, there is a difference of included studies between this review and other meta-analysis. So far, the safety issues of melatonin in pediatrics remain a big concern (Kennaway, 2015) and there is a lack of RCTs on long-term usage of melatonin in pediatrics. Thus, children (< 18 y) are excluded in this article, which is in accordance with previous meta-analysis (Brzezinski et al., 2005). Risk of bias of enrolled studies were assessed using The Cochrane Collaboration's tool (Figs. 5 and 6) (Higgins et al., 2011) and the overall risk of bias is low or unclear. Besides, we also searched for unpublished and commercially-sponsored work during searching. These data are excluded due to the absence of peer review and potential interacted interest.

Strikingly, the strong effect of melatonin seems to be opposed to the results of the present study. Based on previous studies and hypothesis, the following explanations may address this question: (1) The studies on the sleep induction effect of melatonin are mainly conducted in animals (Fisher et al., 2008; Fisher and Sugden, 2009), instead of human beings. (2) The meta-analysis articles are based on RCTs, some of which demonstrated that the effects of melatonin are not strong (Folkard et al., 1993; James et al., 1998; Yoon and Song, 2002), (mean difference < 5 min compared to the controls (Wade et al., 2007; Zhdanova et al., 2001)). (3) Different eligibility criteria in those meta-analysis articles contribute to different enrolled population.

As shown in Fig. 2, pooled data reveal that compared to placebo, exogenous melatonin lowers sleep onset latency (Total mean difference: −2.48 min, 95% CI: −4.56, −0.40, I^2 = 20%). The heterogeneity is less than 50% and is not significant, suggesting a good homogeneity among the enrolled studies. Auld's work evaluates the effects of

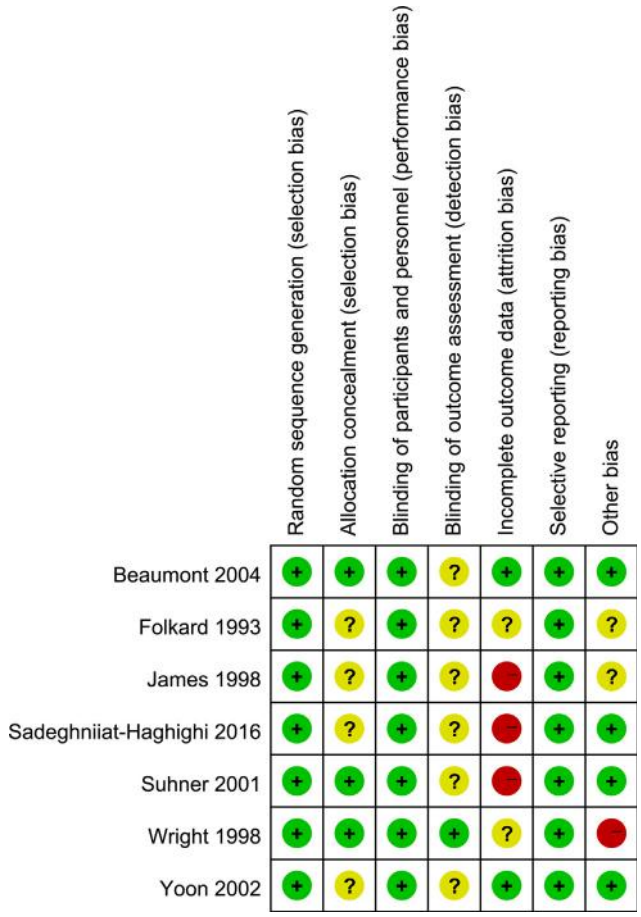

Fig. 6. Risk of bias summary. This figure shows review authors' judgements about each risk of bias item for each included study.

melatonin on primary sleep disorders and concluded a total mean difference of −5.05 min (95% CI: −8.51, −1.59) (Auld et al., 2017). Brzezinski's study assessed the effects of exogenous melatonin on sleep (including primary and secondary outcomes). The total mean difference of sleep onset latency are −4.0 min (95% CI: −2.5, −5.4) (Brzezinski et al., 2005). The result of this meta-analysis is similar to previous results (Auld et al., 2017; Brzezinski et al., 2005). Overall, despite the statistically significant mean differences, 2.48 min for secondary insomnia patients is not clinically relevant, which requires further studies to evaluate its efficacy.

Fig. 3 shows that melatonin increases total sleep time (Total mean difference: 29.27 min, 95% CI: 6.68, 51.56, $I^2 = 26\%$). The heterogeneity is also non-significant among the included studies. Fig. 4 shows that melatonin has no significant actions on sleep efficiency (Total mean difference: 1.46, 95% CI: −0.43, 3.35, $I^2 = 1\%$). It is clear that total sleep time is increased and sleep onset latency is reduced. Notably, the sleep efficiency does not increase. The following explanations may account for the observations reported. On one hand, the enrolled studies of Figs. 2 and 4 are different. Seven studies are included in Fig. 2, whereas three are inlcuded in Fig. 4. Sleep efficiency is not provided by the Folkard's study (Folkard et al., 1993), or in 3 other studies (James et al., 1998; Suhner et al., 2001; Wright et al., 1998). Thus, only 3 studies are included in Fig. 4. This difference might contribute to the insignificance of sleep efficiency. Conversely, the enrolled studies and sample size in Figs. 2 and 4, as well as the total number of individuals is small. Risk of bias of enrolled 7 studies are summarized in Figs. 5 and 6, according to The Cochrane Collaboration's tool (Higgins et al., 2011), which bears a high risk of bias on attrition bias and other biases across the 7 studies. Of note, the researches by James, Sadeghniiat-Haghighi, Suhner, and Wright have a high risk of bias (James et al., 1998; Sadeghniiat-Haghighi et al., 2016; Suhner et al., 2001; Wright et al., 1998), which also contribute to the suspicion of literature quality, which needs further study.

4.3. Limitations

There are some limitations to this review. Firstly, unpublished studies were excluded in this meta-analysis, which may increase the publication bias (Kicinski et al., 2015). Secondly, although data analysis revealed that melatonin reduces sleep onset latency and increases total sleep time, these results need additional support from additional RCT. The reduced time of sleep onset latency is not clinically significant. Third, each study does not adjust for the same confounders, such as duration of administration and dosage, therapeutic period, and gender, which requires further well-designed RCT where these issues are taken into consideration.

4.4. Conclusion

This meta-analysis is dedicated to elucidating the effects of exogenous melatonin, compared to the placebo, on the nocturnal outcomes of secondary sleep disorders. Meta-analysis of the data from a series of studies with small sample size demonstrates that exogenous melatonin improves the sleep quality of secondary sleep disorders. Based on the current advantages of melatonin in the management of secondary sleep disorders, it is hoped that there will be a tremendous growth in the use of melatonin application worldwide. Besides, little evidence is available regarding the adverse effects of long-term use of melatonin (Brzezinski, 1997; Zisapel, 2018). Clinicians should be alert to these shortcomings but also aware of the potential role of melatonin in clinical psychiatry and sleep medicine. Although further studies are needed to establish the optimal approach to this treatment in clinic, this meta-analysis clearly supports the use of melatonin as a management for patients with secondary sleep disorders as a complementary therapy.

5. Authors' contributions

Yang Y designed the study. Li T and Jiang S searched the literature and wrote the manuscript. Li T and Yang Z verified the data and participated in the resolution of disagreements. Han MZ, Lv JJ, and Deng C extracted and analyzed the data. Yang Z draw the picture. Reiter RJ revised the manuscript. All authors read the manuscript with critical revision

6. Disclosures

All authors declare no competing interests. The National Natural Science Foundation of China and China Postdoctoral Science Foundation have no roles in the design, data collection and analysis, writing of the report, or approval of the manuscript.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (81500263 and 81673578) and China Postdoctoral Science Foundation (2016T90973 and 2015M572681). We are indebted to the authors of the primary studies. Furthermore, we sincerely appreciate the hard work from the distinguished editors and reviewers.

References

Auld, F., Maschauer, E.L., Morrison, I., Skene, D.J., Riha, R.L., 2017. Evidence for the efficacy of melatonin in the treatment of primary adult sleep disorders. Sleep Med. Rev. 34, 10–22.

Beaumont, M., Batejat, D., Pierard, C., Van Beers, P., Denis, J.B., Coste, O., Doireau, P.,

Chauffard, F., French, J., Lagarde, D., 2004. Caffeine or melatonin effects on sleep and sleepiness after rapid eastward transmeridian travel. J. Appl. Physiol. 96, 50–58 (1985).
Brzezinski, A., 1997. Melatonin in humans. N. Engl. J. Med. 336, 186–195.
Brzezinski, A., Vangel, M.G., Wurtman, R.J., Norrie, G., Zhdanova, I., Ben-Shushan, A., Ford, I., 2005. Effects of exogenous melatonin on sleep: a meta-analysis. Sleep Med. Rev. 9, 41–50.
Buscemi, N., Vandermeer, B., Hooton, N., Pandya, R., Tjosvold, L., Hartling, L., Vohra, S., Klassen, T.P., Baker, G., 2006. Efficacy and safety of exogenous melatonin for secondary sleep disorders and sleep disorders accompanying sleep restriction: meta-analysis. BMJ 332, 385–393.
Carrier, J., Semba, K., Deurveilher, S., Drogos, L., Cyr-Cronier, J., Lord, C., Sekerovick, Z., 2017. Sex differences in age-related changes in the sleep-wake cycle. Front. Neuroendocrinol. 47, 66–85.
Chan, V., Chen, L., Mesana, L., Mesana, T.G., Ruel, M., 2011. Heart valve prosthesis selection in patients with end-stage renal disease requiring dialysis: a systematic review and meta-analysis. Heart 97, 2033–2037.
Fisher, S.P., Davidson, K., Kulla, A., Sugden, D., 2008. Acute sleep-promoting action of the melatonin agonist, ramelteon, in the rat. J. Pineal Res. 45, 125–132.
Fisher, S.P., Sugden, D., 2009. Sleep-promoting action of IIK7, a selective MT2 melatonin receptor agonist in the rat. Neurosci. Lett. 457, 93–96.
Folkard, S., Arendt, J., Clark, M., 1993. Can melatonin improve shift workers' tolerance of the night shift? Some preliminary findings. Chronobiol. Int. 10, 315–320.
Gervais, N.J., Mong, J.A., Lacreuse, A., 2017. Ovarian hormones, sleep and cognition across the adult female lifespan: an integrated perspective. Front. Neuroendocrinol. 47, 134–153.
Higgins, J.P., Altman, D.G., Gotzsche, P.C., Juni, P., Moher, D., Oxman, A.D., Savovic, J., Schulz, K.F., Weeks, L., Sterne, J.A., Cochrane Bias Methods Group, Cochrane Statistical Methods Group, 2011. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 343, d5928.
James, M., Tremea, M.O., Jones, J.S., Krohmer, J.R., 1998. Can melatonin improve adaptation to night shift? Am. J. Emerg. Med. 16, 367–370.
Jan, J.E., Reiter, R.J., Wasdell, M.B., Bax, M., 2009. The role of the thalamus in sleep, pineal melatonin production, and circadian rhythm sleep disorders. J. Pineal Res. 46, 1–7.
Keijzer, H., Smits, M.G., Duffy, J.F., Curfs, L.M., 2014. Why the dim light melatonin onset (DLMO) should be measured before treatment of patients with circadian rhythm sleep disorders. Sleep Med. Rev. 18, 333–339.
Kennaway, D.J., 2015. Potential safety issues in the use of the hormone melatonin in paediatrics. J. Paediatr. Child Health 51, 584–589.
Kicinski, M., Springate, D.A., Kontopantelis, E., 2015. Publication bias in meta-analyses from the Cochrane Database of Systematic Reviews. Stat. Med. 34, 2781–2793.
Li, T., Yang, Z., Jiang, S., Di, W., Ma, Z., Hu, W., Chen, F., Reiter, R.J., Yang, Y., 2017. Melatonin: does it have utility in the treatment of haematological neoplasms? Br. J. Pharmacol.
Liira, J., Verbeek, J.H., Costa, G., Driscoll, T.R., Sallinen, M., Isotalo, L.K., Ruotsalainen, J.H., 2014. Pharmacological interventions for sleepiness and sleep disturbances caused by shift work. Cochrane Database Syst. Rev CD009776.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., Prisma Group, 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 6, e1000097.
Nathan, S., Goodarzi, Z., Jette, N., Gallagher, C., Holroyd-Leduc, J., 2017. Anticoagulant and antiplatelet use in seniors with chronic subdural hematoma: systematic review. Neurology 88, 1889–1893.
Reiter, R.J., Rosales-Corral, S., Tan, D.X., Jou, M.J., Galano, A., Xu, B., 2017. Melatonin as a mitochondria-targeted antioxidant: one of evolution's best ideas. Cell Mol. Life Sci. 74, 3863–3881.
Reiter, R.J., Tan, D.X., Galano, A., 2014. Melatonin: exceeding expectations. Physiology (Bethesda) 29, 325–333.
Robinson, K.A., 2005. Should meta-analysts search Embase in addition to Medline? J. Clin. Epidemiol. 58, 320 author reply 321.
Sadeghniiat-Haghighi, K., Bahrami, H., Aminian, O., Meysami, A., Khajeh-Mehrizi, A., 2016. Melatonin therapy in shift workers with difficulty falling asleep: a randomized, double-blind, placebo-controlled crossover field study. Work 55, 225–230.
Suhner, A., Schlagenhauf, P., Hofer, I., Johnson, R., Tschopp, A., Steffen, R., 2001. Effectiveness and tolerability of melatonin and zolpidem for the alleviation of jet lag. Aviat. Space Environ. Med. 72, 638–646.
van Geijlswijk, I.M., Korzilius, H.P., Smits, M.G., 2010. The use of exogenous melatonin in delayed sleep phase disorder: a meta-analysis. Sleep 33, 1605–1614.
Wade, A.G., Ford, I., Crawford, G., McMahon, A.D., Nir, T., Laudon, M., Zisapel, N., 2007. Efficacy of prolonged release melatonin in insomnia patients aged 55–80 years: quality of sleep and next-day alertness outcomes. Curr. Med. Res. Opin. 23, 2597–2605.
Walters, A.J., Lader, M.H., 1971. Hangover effect of hypnotics in man. Nature 229, 637–638.
Wright, S.W., Lawrence, L.M., Wrenn, K.D., Haynes, M.L., Welch, L.W., Schlack, H.M., 1998. Randomized clinical trial of melatonin after night-shift work: efficacy and neuropsychologic effects. Ann. Emerg. Med. 32, 334–340.
Yang, Y., Sun, Y., Yi, W., Li, Y., Fan, C., Xin, Z., Jiang, S., Di, S., Qu, Y., Reiter, R.J., Yi, D., 2014. A review of melatonin as a suitable antioxidant against myocardial ischemia-reperfusion injury and clinical heart diseases. J. Pineal Res. 57, 357–366.
Yoon, I.Y., Song, B.G., 2002. Role of morning melatonin administration and attenuation of sunlight exposure in improving adaptation of night-shift workers. Chronobiol. Int. 19, 903–913.
Yu, L.M., Di, W.C., Dong, X., Li, Z., Zhang, Y., Xue, X.D., Xu, Y.L., Zhang, J., Xiao, X., Han, J.S., Liu, Y., Yang, Y., Wang, H.S., 2018. Melatonin protects diabetic heart against ischemia-reperfusion injury, role of membrane receptor-dependent cGMP-PKG activation. Biochim. Biophys. Acta 1864, 563–578.
Zhang, W., Chen, X.Y., Su, S.W., Jia, Q.Z., Ding, T., Zhu, Z.N., Zhang, T., 2016. Exogenous melatonin for sleep disorders in neurodegenerative diseases: a meta-analysis of randomized clinical trials. Neurol. Sci. 37, 57–65.
Zhdanova, I.V., Wurtman, R.J., Regan, M.M., Taylor, J.A., Shi, J.P., Leclair, O.U., 2001. Melatonin treatment for age-related insomnia. J. Clin. Endocrinol. Metab. 86, 4727–4730.
Zisapel, N., 2018. New perspectives on the role of melatonin in human sleep, circadian rhythms and their regulation. Br. J. Pharmacol.

Nurse Education Today 39 (2016) 16–21

Contents lists available at ScienceDirect

Nurse Education Today

journal homepage: www.elsevier.com/nedt

Nursing students' perception of high-fidelity simulation activity instead of clinical placement: A qualitative study

Mio Leng Au [1,6], Man Sao Lo [2], Wan Cheong [3], Si Chen Wang [4], Iat Kio Van [5]

Kiang Wu Nursing College of Macau, Est. Repouso No. 35, R/C, SAR Macau, China

ARTICLE INFO

Article history:
Accepted 20 January 2016

Keywords:
Clinical education
High-fidelity simulation
Nursing students
Perception
Simulation

SUMMARY

Background: The high-fidelity simulation (HFS) has been utilized in nursing education for more than 20 years. Advantages of the use of high-fidelity simulation in nursing education have been documented in the literature. Based on the advantages, it has been arranged as a part of the clinical study course of the first year baccalaureate nursing program in one of the nursing colleges in Macau recently.
Objective: The aim of this study is to explore undergraduate nursing students' perception of using high-fidelity simulation as part of their clinical study course in Macau.
Design: This is a qualitative study using open-ended questionnaire.
Setting: This study was implemented at the nursing laboratory between 1 April and 17 April 2015, which was the period of preliminary clinical study course of year one nursing students.
Participants: A purposive sample was sought from the voluntary year one undergraduate nursing students who participated in the clinical study course.
Methods: Students received two high-fidelity simulation sections during the course, while a self-administered open-ended questionnaire was allocated afterward. Qualitative content analysis was performed after data collection.
Results: Two themes emerged in this study, which included "appreciation" and "misunderstanding". They were further divided into five categories; as "positive feelings", "gaining a suitable atmosphere for learning", "assist of adequate emergency preparation: resourceful ability", "contempt", and "rote learning".
Conclusion: This was the first time to utilize HFS activities as a part of the clinical study course in one nursing college in Macau. These HFS activities instead of a part of real clinical placement were appreciated by nursing students. And it mainly contributed to the resourceful ability in students' view. During the HFS activities, nursing educators should consider the misunderstanding of HFS activities of students that a few nursing students despised simulator's life and got rote learning method.

Introduction

Simulation is recognized as an essential part of nursing clinical education by nursing institute (California Board of Registered Nursing, 2015; Colorado Department of Regulatory Agencies Board of Nursing, 2015) since it can provide a safe environment for students to learn clinical practice (Arthur et al., 2013). There has been an increasing utilization of high-fidelity simulation (HFS) in nursing education since the 1990s (Crytzer, 2011). HFS refers to the use of a computer-controlled full size manikin to demonstrate realistic clinical manifestations and clinical scenarios. It can also communicate and interact with the learners (Arthur et al., 2013; Gates et al., 2012). HFS has been arranged as part of the first year clinical study course of the baccalaureate nursing program in one of the nursing colleges in Macau recently. The aim of this study is to explore undergraduate nursing students' perception of using HFS as part of their clinical study course in Macau. It aims at providing insights into the future improvement of HFS-based clinical teaching in the nursing students.

Literature Review

HFS can be used to train learners' management of imitated life-like clinical events in nursing laboratory (Levett-Jones et al., 2011). There is emerging evidence in the effectiveness of HFS in nursing education. Shin et al. (2015) conducted a meta-analysis on the effectiveness of

E-mail addresses: aml@kwnc.edu.mo (M.L. Au), mandylo@kwnc.edu.mo (M.S. Lo), ch_wan@kwnc.edu.mo (W. Cheong), sichen@kwnc.edu.mo (S.C. Wang), van@kwnc.edu.mo (I.K. Van).
[1] Tel.: +853 66605322; fax: +853 28365204.
[2] Tel: +853 66827689; fax: +853 28365204.
[3] Tel: +853 82956230; fax: +853 28365204.
[4] Tel.: +853 82956271; fax: +853 28365204.
[5] Tel.: +853 82956202; fax: +853 28365204.
[6] Department of Nursing National Taiwan University.

http://dx.doi.org/10.1016/j.nedt.2016.01.015

simulation in nursing education by pooling 20 experimental and quasi-experimental quantitative studies. They found that HFS had a large size effect (0.81), which seemed to be larger than that in the overall simulation interventions (0.71) (Shin et al., 2015). Several systematic reviews further examined the impact of HFS in different aspects of nursing education. It was revealed in a systematic review that the mean scores of knowledge and skill exams were increased by 0.53 point and 1.15 points respectively after using HFS, but a mixed result was found in the score of objective structured clinical examination (OSCE) (Yuan et al., 2012b). Moreover, a mixed contribution of HFS was also seen in the confidence and competence in another systematic review (Yuan et al., 2012a).

The findings from some qualitative studies tend to be positive when exploring the perceptions of HFS in both faculty members and students. Silvia (2013) conducted a qualitative case study by interviewing and allocating qualitative questionnaires to 14 nursing faculty to explore how HFS can influence students' safe clinical practice. It was demonstrated that 78.5% of the faculty thought HFS activities could enhance learning outcome by providing safer patient care, while same percentage of them agreed HFS could increase nursing students' critical thinking skill (Silvia, 2013). Around 64% of them also thought that it enhanced the acquisition of skills (Silvia, 2013). Similar results were indicated in another study using semi-structured interview. Students reported an increase in knowledge, skill and confidence level in clinical placement after HFS (Ogilvie et al., 2011). Furthermore, students believed that HFS was very useful and should be set as a requirement before clinical study (Darcy Mahoney et al., 2013). This is also supported by the result of another study. It was reported that 80% of the nursing students agreed HFS as an authentic learning experience, while about 95% of them thought it could enhance patient safety, and the same proportion of them planned to apply the skills learnt in simulation to clinical practice (McCaughey and Traynor, 2010). Additionally, HFS activities could lead to students' sense of confidence, preparedness, and satisfaction with the clinical performance (Crytzer, 2011). Despite the advantages mentioned in these qualitative studies, some studies argued that HFS can also lead to anxiety in participants. A lack of communication skill during the simulated interaction was reported (Pike and O'Donnell, 2010). Students might also feel anxious as it made them think of the future placement and transition to a staff nurse (McCaughey and Traynor, 2010). There are some suggestions related to the possible improvement in HFS activities in the literature. Students suggested that it was necessary for them to have more time to be familiar with the functions of HFS (Wotton et al., 2010). The study of Ogilvie et al. (2011) showed that, students agreed that a positive HFS learning experience could be enhanced by a realistic clinical scenario experience under an appropriate facilitation, with a debriefing section. The importance of realism and facilitator was also emphasized in another study; it was argued that, students' learning experience could be influenced by the ability of engaging participants with the character and taking the activity seriously, while the knowledge and skills of facilitators were also reported to be significant during the process (Reid-Searl et al., 2011). Similar finding was also described in Pike and O'Donnell's (2010) study.

These suggest the essential elements in designing and implementing HFS activities. A simulation model on HFS activities was proposed in 2005 by Jeffries. It was widely used in guiding simulation activities; it has been applied to the simulation of end-of-life-care, self-confidence promoting etc. (Fabro et al., 2014; Samawi et al., 2014). Jeffries' simulation model comprises five elements, which are the best practices in education, student factors, teacher factors, simulation design characteristics and outcomes. Educational practices include active learning, immediate feedback, student/faculty interaction, collaborative learning, and high expectations, allowing diverse learning styles and time on task. Student factors mean that students should respond to their roles (actors and observers) during the simulation activity, while teacher factors involve the teaching and evaluating roles. The design of the simulation should be tailored to these three mentioned factors, and at the same time, be able to support course goals, skill competencies, learning outcomes and include debriefing section. The outcomes should be associated with the goals, and can be divided into knowledge, skill performance, learner satisfaction, etc. (Jeffries, 2005).

In view of the mentioned advantages, HFS activities guided by Jeffries' simulation model has been arranged as part of the first year clinical study of the baccalaureate nursing (BSN) program in one nursing college in Macau recently. Nevertheless, to date, there is no qualitative data that explores the perception of BSN students of utilizing HFS as a clinical placement either directly or indirectly. Hence, the aim of this study is to explore undergraduate nursing students' perception of HFS activities instead of clinical placement. It aims at providing insights into the future improvement of HFS-based clinical teaching in the nursing students in Macau.

Methods

This is a qualitative study using an open-ended questionnaire. Students received two four-hour HFS sections at nursing laboratory during the period of clinical study course, while a self-administered questionnaire was then allocated and a qualitative content analysis was performed. Jeffries' simulation model was employed in the design of the activity (Fig. 1), while Laerdal SimMan™ patient simulator was used during HFS activities. Several optional sections were held to introduce the simulator's function and the concept of HFS. A meeting was held by the subject teachers before class to discuss the design of the activity, while a mock section was also run among the teachers to ensure the maximization of learning experience and to seek improvement. Before the sections, a brief scenario introduction and relevant materials were uploaded onto the student learning platform.

The participants were divided into groups of 16 to 23 during each HFS activity. Three of them were assigned with different roles, mostly nurses, while the others were observers. On the other hand, as the HFS activities in this study were for learning purpose, the teachers' role was to provide support throughout the section.

The four-hour simulation class was divided into four sections for two groups respectively, which were briefing (half hour), preparation (half hour), running (half hour) and debriefing (half hour). During the briefing section, the simulated environment and technology involved were oriented, while the objectives, activity, amount of time given, role specifications and outcome expectancies were explained. Preparation time was allowed for initial discussion and being familiar with the simulated environment. The scenario was then run under a provided time frame. The role-players were required to practice using the think-aloud technique. Cues and help could be provided by both the observers and the teachers to offer ideas when required during the activity to enhance the idea of collaborative learning. Observers were also asked to take note of the clinical presentation, missing data, and given and required intervention. The whole process was recorded for debriefing reference. Students were encouraged to review and discuss the scenario after the running section under the guidance of the facilitators. Strengths and weaknesses were also discussed for future practice and improvement.

Participants

A purposive sample was recruited from the year one undergraduate nursing students in one of the nursing colleges in Macau (there are two nursing colleges in Macau). The students who were willing to answer the open-ended questionnaire voluntarily after the HFS activities were included in this study. The targeted nursing students had already finished English I, Chinese, Psychology, Fundamentals of Nursing I, Anatomy, Physiology and Biochemistry. They were studying Sociology, Health Assessment, Pharmacology, Pathophysiology, Microbiology-Immunology, Fundamentals of Nursing II and English II in the semester.

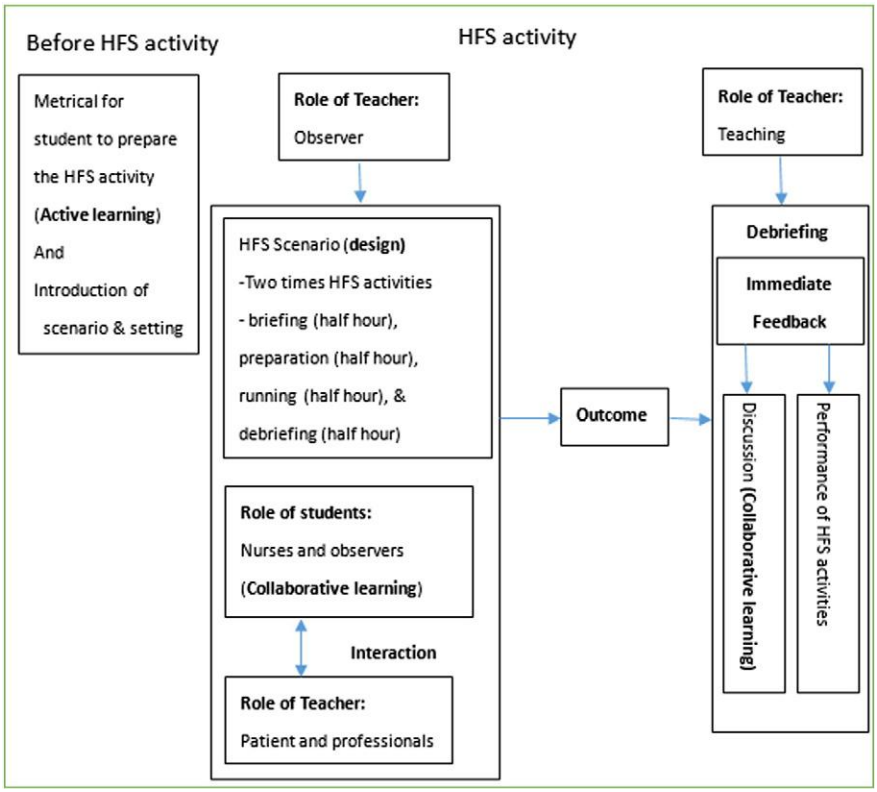

Fig. 1. Design of simulation activity with Jeffries' simulation model.

Setting

This study was implemented at a nursing laboratory between 1 April and 17 April 2015, which was the period of the participants' preliminary clinical study course. The preliminary clinical study course was associated with two subjects, which were "Health Assessment" and "Fundamentals of Nursing".

Data Collection and Analysis

An open-ended questionnaire was developed to explore the perception of using HFS activities instead of clinical placement. It was assigned to the participants immediately after the HFS activities. The questions were:

1. What is your opinion on the high-fidelity simulation activity used during clinical study course (instead of real clinical placement)?
2. What are the advantages of using the high-fidelity simulation activity?
3. What are the disadvantages of using the high-fidelity simulation activity?
4. What did you gain from the high-fidelity simulation activity?
5. What did you lose from the high-fidelity simulation activity?
6. What is your suggestion for the high-fidelity simulation activity?

Finally, the data was analyzed using qualitative content analysis. Raw data was read by three authors separately to obtain code ("based on the content representation") from each line ("breaking down data into smaller units"); themes ("grouping coded material based on shared concepts") were then retrieved by coding and categorizing (Polit and Beck, 2012, p. 564) (Table 1). To ensure the trustworthiness, the framework of Lincoln and Guba was applied which encompassed several dimensions — credibility, transferability, confirmability, and dependability. Credibility engendered confidence in the truth of the data and researchers' interpretations; transferability was that the qualitative findings could be transferred to other settings; confirmability meant that study results were derived from characteristics of the participants and the study context; and dependability referred to the evidence that

Table 1
Example of content analysis.

Original sentence (examples)	Code	Category	Theme
"*HFS activities were very good and vivid. The authentic response of SimMan could train my resourceful ability*" (13)	Good and vivid	Positive feelings	Appreciation
"*In short, it is very well. The HFS activities instead of real clinical placement could increase the experience of facing simulated patient, it could improve the way of dealing with patient*" (74)	Very well		
"*Good. It was more relaxing than real clinical setting as we can learn different health assessment methods under little stress*" (04)	More relaxing	Gaining a suitable atmosphere for learning	
"*The atmosphere of HFS activities was exciting. It could let me develop the resourceful ability for the urgent condition of patient which could be rarely encountered in real clinical setting*" (20)	Exciting		
"*I thought the arrangement of HFS activities during clinical study course could let me become familiar with the real hospital environment and develop the resourceful ability. It was good for me to be ready for the real clinical setting very soon*" (01)	Resourceful ability	Assist of adequate emergency preparation: resourceful ability	
"*It enhanced my resourceful ability with the emergency scenario through HFS activities; I will keep calm in any emergency situation in the future*" (05)			

was consistent and stable (Polit and Beck, 2012, p. 175, 180). Based on the framework, the participants' honesty was encouraged before they answered the questionnaire. After the questionnaire had been completed, the themes and categories were given back to the participants to confirm the findings (credibility and confirmability). The process of the HFS activities, data collection, participants and setting were described in this study to assure the dependability and transferability of the study. Moreover, data analysis was implemented by three authors independently in order to promote confirmability (Polit and Beck, 2012; Streubert, 2011).

Ethical Considerations

This study was conducted after receiving the agreement from the education department of the college. The project was introduced and explained to all targeted nursing students before the HFS activities, and the inform consent was obtained from those who were willing to join and answer the anonymous questionnaire in this study. The participants were informed that this subject would not be graded. Data access was limited only to the researchers of this study, and was only allowed for the use of teaching improvement and academic exchange. The completed questionnaire will be destroyed 5 years after the study has been completed according to the informed consent.

Results

There were 80 first year nursing students participating in this study, including 6 males and 74 females aged from 18 to 22. Three of them retook the clinical study course this year. All 80 participants had simulation learning experience before. 41 participants had five-day real clinical placement after the HFS activities (class A) and 39 had the clinical placement before the HFS activities (class B). From the qualitative content analysis of the participants' script (focused on question 1 for this study), 2 themes covering 5 categories emerged as below.

Appreciation

For these year one students, there was two-week clinical study course to let them put the knowledge into practice at the end of the first academic year. They had desired the clinical study course very much. Despite that our college arranged HFS activities instead of the partial real clinical placement this year, the feelings of the participants of these HFS activities were positive, although HFS activities was not new for them. In addition, the HFS activities were carried out in a simulated ward of nursing laboratory; and the interaction was arranged between participants, teachers and SimMan through several scenarios which combined three subjects ("Health Assessment", "Fundamentals of Nursing", and "Pharmacology"). The participants (no matter class A or class B) thought it was an adequate preparation for caring real patients in suitable atmosphere. Both group A and group B focused on acquiring resourceful ability and health assessment ability. Only two participants, who wanted to go to real clinical setting, thought that the HFS activities should not be a part of clinical study course. All of the mentioned showed that the participants appreciated the HFS activities. Hence, the "positive feelings", "gaining a suitable atmosphere for learning", and "assist of adequate preparation" were the three categories under the theme "appreciation".

Positive Feelings

Over 70% of the participants emphasized that they had positive feelings toward HFS activities; no matter they joined the activities before or after real clinical placement. Only one participant said it was boring. And others (24%) did not express their feelings. Among the 74.8% of all participants, around 35% used the word "good, very good, very well or not bad", 14% described it as "interesting" and around 6% used "vivid" to describe the activity, while 3.8% of them thought the activity was essential. Nearly 16% portrayed it as "practical", "fresh", "a rare opportunity" and "a good opportunity". These positive feelings indicate that students are satisfied with the HFS activities.

> "*HFS activities were very good and vivid. The authentic response of SimMan could train my resourceful ability*" (13).

> "*In short, it is very well. The HFS activities instead of real clinical placement could increase the experience of facing simulated patient; it could improve the way of dealing with patient*" (74).

Gaining a Suitable Atmosphere for Learning

85% of the participants considered that the HFS activities, instead of the real clinical placement, provided a suitable atmosphere for learning, as "relaxing", "exciting", "funny" etc. The 85% of participants reflected that the atmosphere let them acquire resourceful ability (55%), health assessment skill (28%), communication (9%), and knowledge application (8%). Related to the resourceful ability, 33% of the participants in group A and 22% of the participants in group B mentioned it. For health assessment ability, 12% of the participants in group A and 16% of the participants in group B expressed it. On the other hand, two more participants in group B than group A mentioned the learning of communication skill, while three more participants in group A than group B mentioned knowledge application. These results show that HFS activities could provide suitable atmosphere for learning, no matter the participants joined the HFS activities before or after the real clinical placement.

> "*Good. It was more relaxing than real clinical setting as we can learn different health assessmentmethods under little stress*" (04).

> "*The atmosphere of HFS activities was exciting. It could let me develop the resourceful ability for the urgent condition of patient which could be rarely encountered in real clinical setting*" (20).

Assist of Adequate Emergency Preparation: Resourceful Ability

One of the scenarios was designed for the subject "Fundamentals of Nursing — Basic Life Support"; the participants should save the life of the SimMan in an emergency condition. 55% of the 85% of participants thought that these HFS activities could train their resourceful ability. Hence, they had learnt how to find fast and smart ways to overcome difficulties. It was very useful for participants in terms of the future clinical placement.

> "*I thought the arrangement of HFS activities during clinical study course could let me become familiar with the real hospital environment and develop the resourceful ability. It was good for me to be ready for the real clinical setting very soon*" (01).

> "*It enhanced my resourceful ability with the emergency scenario through HFS activities; I will keep calm in any emergency situation in the future*" (05).

Misunderstanding

As mentioned, the HFS activities were carried out through SimMan (which was not a real human being) and scenario. A few participants misunderstood the purpose of HFS activities. They thought that SimMan's life was not an important issue; and some thought that the demonstration of scenario let them remember the situation for future. Hence, the "contempt", and "rote learning" were the two categories under the theme "misunderstanding".

Contempt

5% (2 in group A and group B respectively) of all the participants thought that making mistakes was not an issue when providing care for SimMan. They also mentioned that they might keep this attitude when caring real-life patients. It was a negative influence of the HFS activities. It might affect the participants' attitude and action in the real clinical setting.

> "*Although it is a high-fidelity simulator, it is still different from human beings in reality. Students may get used to the way of treating these simulators and keep the same attitude when treating real patients*" (33).

> "*SimMan is not a real human being, I did not concern about the mistakes I made because I will still be forgiven anyway*" (58).

Rote Learning

14% of the participants (9 in group A and 2 in group B) thought the HFS activities were used to promote their rote learning; they emphasized the importance of remembering the manifestations of SimMan. It was related to the learning style of the participants.

> "*Remembering the patient's changes throughout the scenario was what I gained from the HFS activities*" (40).

> "*The HFS activities could help us to remember the clinical manifestations in a short time*" (45).

Discussion

This was the first time to utilize HFS activities as a part of the clinical course in our nursing college. This arrangement was the transition to the modified BSN program curriculum for the future. And the amount of hours of clinical study course would be reduced to less than 1840 h in the future curriculum to match other BSN nursing program all over the world. Hence, the HFS activities would be the replacement of the real clinical placement that follows the contents of subject lecture in our college in the future. And this study tries to explore the perception of nursing students related to the HFS activities instead of real clinical placement. 85% of nursing students appreciated the HFS activities instead of real clinical placement for learning. It emerged through "positive feelings", "gaining a suitable atmosphere for learning", and "assist of adequate emergency preparation: resourceful ability". And it was the first study to find out the "resourceful ability" as category or item of result (55% of the students mentioned it) which is not the same as other qualitative or quantitative literatures that showed the result as "acquisition of skills", "clinical performance", "confidence", "critical thinking", "knowledge", "preparedness", "safe patient care", and "satisfaction" (Crytzer, 2011; McCaughey and Traynor, 2010; Ogilvie et al., 2011; Silvia, 2013; Yuan et al., 2012a, 2012b). "Resourceful" is defined as "skilled in devising expedients or in meeting difficulties; full of practical ingenuity" (Oxford University Press, 2015). The result of "resourceful ability" may relate to the design of the scenario — basic life support. It aims at letting students gradually achieve the objective of BSN program curriculum of our college that "students could keep clam and act immediately during the emergency situation as a graduate" (Kiang Wu Nursing College of Macau, 2015). Based on the mentioned simulation model, it is very important to design the scenario that can support course goals (Jeffries, 2005). The scenario of these HFS activities was that there was a patient who suddenly became unconscious, where students should manage the situation and save the life of the SimMan through the skill of basic life support. The goals of the course are: (1) to demonstrate the ability of basic life support; and (2) to manage the emergency situation simply. And the result indicated that students thought that the HFS activities could train their resourceful ability. On the other hand, the "positive feelings" and "a suitable atmosphere for learning" were two other categories of the theme "appreciation". No matter the students joined the HFS activities before or after the real clinical placement, they appreciated that the HFS activities were able to let them learn with positive feelings, especially the development of resourceful ability and health assessment ability. Both are the objectives of the HFS activities. But for the feedback about learning of communication skill, more participants in group B than group A mentioned about it. It was the same as the result of Pike and O'Donnell's (2010) that there was a lack of communication skill during the simulated interaction. The students of group B went to the clinic first, they had the experience of talking with real patients. Hence, they could talk more with the SimMan and target to learn communication skill. Students of both group A and group B also mentioned knowledge application. The result matched with the result of the study of Yuan et al. (2012b) that knowledge score would be increased after using HFS. It was evident that HFS contributed to the knowledge of students. As mentioned, there were advantages for arranging HFS activities before or after the clinical placement. It was not the same as the finding of the study of Darcy Mahoney et al. (2013).

Nursing educators should consider that nursing students would misunderstand the objective of HFS activities, the findings of this study showed that "contempt" and "rote learning" were the two categories of the theme "misunderstanding". During the HFS activities, a few students (no matter they were from group A or group B) thought SimMan's life was not an issue and might got the rote learning method. The mentioned findings could not be found in other studies. Nursing educators should pay attention to this because the result of McCaughey and Traynor (2010) showed that 95% of the students would plan to apply what they learnt in simulation to clinical practice. The reason might be that the students thought that the SimMan was not the real human being, so they could not put themselves in the scenario. And some students might use the surface learning style that promoted the rote learning. The design of scenario of HFS activities should be taken seriously to avoid the mentioned situations.

Limitation

Due to the arrangement of clinical study course of the BSN program, only year one students were included in this study. Thus the perception expressed was from those who did not have much clinical experience. A comparison of the perception of different classes and long-term follow-up should be done in the future.

Since this is only an exploratory study, an open-ended questionnaire was used. This prohibited a further and deeper exploration of the information provided, while the data of facial expressions and body language were also missing. A face-to-face interview should be considered in future studies.

Conclusion

This was the first time to utilize HFS activities as a part of the clinical study course in one of the nursing colleges in Macau. It was based on the advantages of the use of high-fidelity simulation, and was the preparation for the transition to the modified curriculum that decreases the amount of hours of the real clinical placement for the future. Hence, the HFS activities were carried out for year one students during clinical study course period. This study explored the perception of nursing students of HFS activities instead of clinical placement. The findings showed that these HFS activities instead of part of real clinical placement were appreciated by nursing students. And it mainly contributed to the resourceful ability in the students' view which was not found before. During the HFS activities, nursing educators should consider the misunderstanding of HFS activities of students that a few nursing students despised simulator's life and got rote learning method. The mentioned findings were also not found in other studies. The design of

scenario of HFS activities should be taken seriously to avoid the mentioned situations.

References

Arthur, C., Levett-Jones, T., Kable, A., 2013. Quality indicators for design and implementation of simulation experiences: a Delphi study. Nurse Educ. Today 33 (11), 1357–1361. http://dx.doi.org/10.1016/j.nedt.1012.07.012.

California Board of Registered Nursing, 2015. Title 16, California code of regulations. Article 3. Prelicensure nursing programs. Retrieved from http://www.rn.ca.gov/regulations/title16.shtml#1426.

Colorado Department of Regulatory Agencies Board of Nursing, 2015,. Department of regulatory agencies: division of professions and occupations. Chapter 2 — rules and regulations for approval of nursing education programs. Retrieved from http://cdn.colorado.gov/cs/Satellite/DORA-Reg/CBON/DORA/1251631690394.

Crytzer, M.L., 2011. The effect of high-fidelity home health simulations on nursing students' clinical performance (doctoral dissertation). Available from ProQuest Dissertations and Theses Global (UMI No. 3486967).

Darcy Mahoney, A.E., Hancock, L.E., Iorianni-Cimbak, A., Curley, M.A.Q., 2013. Using high-fidelity simulation to bridge clinical and classroom learning in undergraduate pediatric nursing. Nurse Educ. Today 33 (6), 648–654. http://dx.doi.org/10.1016/j.nedt.2012.01.005.

Fabro, K., Schaffer, M., Scharton, J., 2014. The development, implementation, and evaluation of an end-of-life simulation experience for baccalaureate nursing students. Nurs. Educ. Perspect. 35 (1), 19–25. http://dx.doi.org/10.5480/11–593.1.

Gates, M.G., Parr, M.B., Hughen, J.E., 2012. Enhancing nursing knowledge using high-fidelity simulation. J. Nurs. Educ. 51 (1), 9–15. http://dx.doi.org/10.3928/01484834-20111116-01.

Jeffries, P.R., 2005. A framework for designing, implementing and evaluating simulations used as teaching strategies in nursing. Nurs. Educ. Perspect. 26 (2), 96–103.

Kiang Wu Nursing College of Macau, 2015. The Curriculum of Baccalaureate Nursing (BSN) Program. Kiang Wu Nursing College of Macau, Macau.

Levett-Jones, T., Lapkin, S., Hoffman, K., Arthur, C., Roche, J., 2011. Examining the impact of high and medium fidelity simulation experiences on nursing students' knowledge acquisition. Nurse Educ. Pract. 11 (6), 380–383. http://dx.doi.org/10.1016/j.nepr.2011.03.014.

McCaughey, C.S., Traynor, M.K., 2010. The role of simulation in nurse education. Nurse Educ. Today 30 (8), 827–832. http://dx.doi.org/10.1016/j.nedt.2010.03.005.

Ogilvie, S., Cragg, B., Foulds, B., 2011. Perceptions of nursing students on the process and outcomes of a simulation experience. Nurse Educ. 36 (2), 56–58. http://dx.doi.org/10.1097/NNE.0b013e31820b4fd5.

Oxford University Press, 2015. Oxford English dictionary: resourceful. Retrieved from http://www.oed.com/view/Entry/163769?redirectedFrom=resourceful+#eid.

Pike, T., O'Donnell, V., 2010. The impact of clinical simulation on learner self-efficacy in pre-registration nursing education. Nurse Educ. Today 30 (5), 405–410. http://dx.doi.org/10.1016/j.nedt.2009.09013.

Polit, D.F., Beck, C.T., 2012. Nursing Research: Generating and Assessing Evidence for Nursing Practice. nineth ed. Lippincott Williams & Wilkins, Philadelphia, PA.

Reid-Searl, K., Eaton, A., Vieth, L., Happell, B., 2011. The educator inside the patient: students' insights into the use of high fidelity silicone patient simulation. J. Clin. Nurs. 20 (9), 2752–2760. http://dx.doi.org/10.1111/j.1365-2702.2011.03795.x.

Samawi, Z., Miller, T., Haras, M.S., 2014. Using high-fidelity simulation and concept mapping to cultivate self-confidence in nursing students. Nurs. Educ. Perspect. 35 (6), 408–409. http://dx.doi.org/10.5480/1042.1.

Shin, S., Park, J.H., Kim, J.H., 2015. Effectiveness of patient simulation in nursing education: meta-analysis. Nurse Educ. Today 35 (1), 176–182. http://dx.doi.org/10.1016/j.nedt.2014.09.009.

Silvia, S., 2013. Faculty perceptions of simulation on student learning for safe clinical nursing practice (Doctoral dissertation) Available from ProQuest Dissertations and Theses Global (UMI No. 3600876).

Streubert, H.J., 2011. Qualitative Research in Nursing: Advancing the Humanistic Imperative. fifth ed. Lippincott Williams & Wilkins, Philadelphia, PA.

Wotton, K., Davis, J., Button, D., Kelton, M., 2010. Third-year undergraduate nursing students' perceptions of high-fidelity simulation. J. Nurs. Educ. 49 (11), 632–639. http://dx.doi.org/10.3928/01484834-20100831-01.

Yuan, H.B., Williams, B.A., Fang, J.B., 2012a. The contribution of high-fidelity simulation to nursing students' confidence and competence: a systematic review. Int. Nurs. Rev. 59 (1), 26–33. http://dx.doi.org/10.1111/j.1466-7657.2011.00964.x.

Yuan, H.B., Williams, B.A., Fang, J.B., Ye, Q.H., 2012b. A systematic review of selected evidence on improving knowledge and skills through high-fidelity simulation. Nurse Educ. Today 32 (3), 294–298. http://dx.doi.org/10.1016/j.nedt.2011.07.010.