



**HAL**  
open science

# Facteurs psychosociaux et économiques en milieu professionnel à l'heure de la médecine 4P : évolution longitudinale et prédictive du comportement des travailleurs

Rémi Colin

► **To cite this version:**

Rémi Colin. Facteurs psychosociaux et économiques en milieu professionnel à l'heure de la médecine 4P : évolution longitudinale et prédictive du comportement des travailleurs. Santé. Université Clermont Auvergne, 2024. Français. NNT : 2024UCFA0028 . tel-04624444

**HAL Id: tel-04624444**

**<https://theses.hal.science/tel-04624444>**

Submitted on 25 Jun 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



UNIVERSITE CLERMONT AUVERGNE

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE,  
SANTÉ, AGRONOMIE, ENVIRONNEMENT

THESE D'UNIVERSITE

en vue de l'obtention du grade de DOCTEUR D'UNIVERSITE  
spécialité INGENIERIE DE LA SANTE

par

Rémi COLIN

Présentée et soutenue publiquement le 05 avril 2024

Facteurs psychosociaux et économiques en milieu professionnel  
à l'heure de la médecine 4P : évolution longitudinale et  
prédictive du comportement des travailleurs

Membres du jury :

Arnaud GUILLIN	LMBP, CNRS, Université Clermont Auvergne	Président
Sophie BERJOT	C2S, Université de Reims Champagne-Ardenne	Rapporteuse
Yolande ESQUIROL	CERPOP, INSERM, Université Toulouse III	Rapporteur
Laëtitia ROLLIN	LIMICS, INSERM, Université de Rouen Normandie	Examinatrice
Frédéric DUTHEIL	LAPSCO, CNRS, Université Clermont Auvergne	Directeur de thèse
Bruno PEREIRA	DRCI, CHU Clermont-Ferrand	Co-directeur de thèse
Thomas CORNET	Wittyfit, Cegid	Membre invité





UNIVERSITE CLERMONT AUVERGNE

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE,  
SANTÉ, AGRONOMIE, ENVIRONNEMENT

THÈSE D'UNIVERSITÉ

en vue de l'obtention du grade de DOCTEUR D'UNIVERSITÉ  
spécialité INGÉNIERIE DE LA SANTÉ

par

Rémi COLIN

Présentée et soutenue publiquement le 05 avril 2024

Facteurs psychosociaux et économiques en milieu professionnel  
à l'heure de la médecine 4P : évolution longitudinale et  
prédictive du comportement des travailleurs

Membres du jury :

Arnaud GUILLIN	LMBP, CNRS, Université Clermont Auvergne	Président
Sophie BERJOT	C2S, Université de Reims Champagne-Ardenne	Rapporteuse
Yolande ESQUIROL	CERPOP, INSERM, Université Toulouse III	Rapporteur
Laëtitia ROLLIN	LIMICS, INSERM, Université de Rouen Normandie	Examinatrice
Frédéric DUTHEIL	LAPSCO, CNRS, Université Clermont Auvergne	Directeur de thèse
Bruno PEREIRA	DRCI, CHU Clermont-Ferrand	Co-directeur de thèse
Thomas CORNET	Wittyfit, Cegid	Membre invité



## **Remerciements**

Je tiens tout naturellement à remercier mes deux directeurs de thèse, le Pr Frédéric Dutheil et le Dr Bruno Pereira, pour leur accompagnement au cours de ses trois années de thèse et leur soutien dans tous les projets que nous avons menés.

Merci à l'ensemble des membres de jury d'être présent ce jour pour juger de mon parcours de ces trois dernières années.

Je tiens à remercier tous les membres de l'équipe Wittyfit, passés et présents, pour avoir rendu ce projet de thèse possible grâce à leur envie et leur motivation sans faille. Samuel, Thomas, Benoît, Nour, Willy, Annabelle, Laura, Céline, Raphaël, Margaud, Emilie, Clémence, Margot, Nathalie, Justine, Thibault, Olivier, Jérémy, Vincent, Manuelle, Pierre-André, Sofiane, et enfin Maxime, Alexandre, Chloé, merci à vous tous.

Merci à l'Institut de Médecine du Travail et à ses membres pour l'accueil chaleureux qui m'a été réservé à mon arrivée dans le service et pour m'avoir supporté durant ces trois années. Sarah, Brigitte, Claire, Gil, Luc, Senem, merci pour tous les bons moments passés à vos côtés.

Merci à Guillaume Heller, Stéphanie Léger et Luc Fontana à qui je dois beaucoup pour m'avoir aidé, motivé et accompagné tout au long de ma thèse.

A ma famille et mes amis, sans qui je ne serai rien, merci pour tout.

Enfin, à vous, Maëlys et Luther, pour qui j'espère avoir été d'une aide aussi précieuse que celle que vous m'avez apportée, merci pour tous les moments de partage, de déconne et d'entraide. Hâte d'un jour enfin pouvoir me délecter de toutes les bières que vous me devez !

## Table des matières

Introduction et revue de la littérature .....	12
1. Santé et qualité de vie des travailleurs .....	13
2. Risques et facteurs psychosociaux en milieu professionnel.....	15
2.1. Origines, enjeux pour les entreprises et exemples de risque .....	15
2.2. Les principaux facteurs de risques psychosociaux .....	17
2.3. Les conséquences sanitaires et économiques des risques psychosociaux .....	19
3. La médecine 4P au service de la santé au travail .....	21
4. Wittyfit : un outil de mesure de la qualité de vie au travail .....	22
5. L'analyse des données de santé au travail.....	23
5.1. Données du monde réel et données de santé au travail : définitions et concepts... 24	
5.2. Les enjeux méthodologiques de l'analyse de données de santé au travail .....	25
Résultats et travaux de recherche .....	36
Article n°1 : Methodological issues in analyzing real-world longitudinal occupational health data: a useful guide to approaching the topic .....	39
Article n°2 : The protective role of job control/autonomy on mental strain of managers: a cross-sectional study among Wittyfit's users.....	61
Article n°3 : The managerial role and psychosocial factors of job satisfaction: a cross-sectional study among Wittyfit's users .....	74
Article n°4 : Stress and job satisfaction over time, the influence of the managerial position: a bivariate longitudinal modelling of Wittyfit data.....	83
Article n°5 : Psychosocial well-being index and sick leave in the workplace: a structural equation modelling of Wittyfit data .....	100
Article n°6 : Psychosocial well-being as a protective factor against short- and long-term work absenteeism: from censored-data survival models to multistate analysis .....	125
Discussion générale.....	151

1. L'étude des données de santé au travail au service de la lutte contre les risques psychosociaux .....	152
2. Les défis de l'analyse de données de santé au travail : enjeux et réponses méthodologiques .....	155
Conclusions et perspectives .....	159
Références .....	163
Annexes .....	174
Annexe A : Captures d'écran de la plateforme Wittyfit.....	175
Annexe B : Présentation orale au Small Meeting d'hiver du LAPSCO (2021).....	179
Annexe C : Présentation orale au Small Meeting d'hiver du LAPSCO (2023).....	181
Annexe D : Présentation orale au Small Meeting d'hiver du LAPSCO (2024).....	183
Annexe E : Présentation orale aux Journées de l'Ecole Doctorale SVSAE (2023).....	185

## Liste des tableaux et des figures

### Figures de l'introduction et de la revue de littérature

Figure 1. Facteurs et risques psychosociaux, conséquences sanitaires et économiques .....	20
Figure 2. Enjeux méthodologiques de l'analyse de données de santé au travail .....	26
Figure 3. Exemple de structure de données multiniveaux : population de travailleurs d'une entreprise à plusieurs pôles .....	28
Figure 4. Identification de la nature des données manquantes avant imputation de données..	31

### Tables et figures de l'article 1

Table 1. Imputation methods by type of missing data .....	46
Figure 1. Data diversity of real-world longitudinal occupational health databases .....	42
Figure 2. Example of cluster-correlated data .....	44
Figure 3. Missing data structures .....	46
Figure 4. Evolutionary trajectories of job satisfaction among Wittyfit users .....	50

### Tables et figures de l'article 2

Table 1. Characteristics of workers according to their classification using Karasek's Job Demand-Control-Support model (JDCS).....	67
Figure 1. Flow chart of Wittyfit's users.....	66
Figure 2. Job demand, job control and job support by job position.....	67
Figure 3. Karasek's Job Demand-Control-Support model.....	68
Figure 4. Prevalence of job strain and isostrain by job position .....	68
Figure 5. Sociodemographic factors of the Karasek model dimensions .....	69
Figure 6. Sociodemographic factors of job strain and isostrain.....	69

### **Figures de l'article 3**

Figure 1. Flow chart of Wittyfit's users .....	78
Figure 2. Job satisfaction by job position.....	78
Figure 3. Influence of socio-demographic factors of job satisfaction.....	79
Figure 4. Psychosocial feelings by job position.....	79
Figure 5. Influence of psychosocial factors of job satisfaction by job position.....	80

### **Tables et figures de l'article 4**

Table 1. Stress and job satisfaction by job position over time .....	90
Figure 1. Flowchart of the study .....	90
Figure 2. Stress and job satisfaction by job position across all years (cross-sectional univariate and bivariate analyses) .....	91
Figure 3. Evolution of stress and job satisfaction over time by job position (longitudinal bivariate analysis).....	92
Figure 4. Group-based multi-trajectories of workers' stress and job satisfaction.....	93
Figure 5. Cross-lagged effects of stress and job satisfaction by job position .....	94

### **Tables et figures de l'article 5**

Table 1. Fit indices of models .....	111
Figure 1. Flowchart of the study .....	108
Figure 2. Illustration of the structural equation model of the psychosocial well-being index.....	110
Figure 3. Evolution of goodness-of-fit across models .....	111
Figure 4. Influence of psychosocial factors on psychosocial well-being index.....	113
Figure 5. Influence of psychosocial well-being index on duration of sick leave.....	113

## **Tables et figures de l'article 6**

Table 1. Estimated covariate effect of psychosocial feelings and psychosocial well-being on transition intensities between states .....	140
Figure 1. Flowchart of the study .....	135
Figure 2. Evolution of the probability of surviving short-term absenteeism (simple-event survival model).....	135
Figure 3. Evolution of the probability of surviving long-term absenteeism (simple-event survival model).....	136
Figure 4. Evolution of the probability of surviving short-term absenteeism (multiple-event survival model).....	137
Figure 5. Evolution of the probability of surviving long-term absenteeism (multiple-event survival model).....	138
Figure 6. Estimated time-to-time transition probabilities between states .....	139

## **Figures de la discussion générale**

Figure 5. Contribution de la thèse à l'approfondissement des connaissances sur les relations entre les facteurs psychosociaux, la qualité de vie au travail, le bien-être et l'absentéisme..	154
Figure 6. Solutions et méthodes pour répondre aux enjeux méthodologiques de l'analyse de données de santé au travail.....	156

## Liste des abréviations

<b>95CI/95% CI</b>	“95% confidence interval” (Intervalle de confiance à 95%)
<b>AIC</b>	“Akaike’s information criteria” (Critère d’information d’Akaike)
<b>ANOVA</b>	“Analysis of variance” (Analyse de variance)
<b>BIC</b>	“Bayesian information criteria” (Critère d’information bayésien)
<b>CFA</b>	“Confirmatory factor analysis” (Analyse factorielle confirmatoire)
<b>CLPM</b>	“Cross-lagged panel model” (Modèle de médiation par panel à décalage croisé)
<b>CNIL</b>	Commission nationale de l’informatique et des libertés
<b>DE</b>	“Design effect” (Effet plan de sondage)
<b>ES</b>	“Effect size” (Taille d’effet)
<b>ESS</b>	“Effective sample size” (Taille effective de l’échantillon)
<b>GBTM</b>	“Group-based trajectory model” (Modèle de trajectoire basé sur le groupe)
<b>GCM</b>	“Growth curve model” (Modèle de courbe de croissance)
<b>GLM</b>	“Generalized linear model” (Modèle linéaire généralisé)
<b>GLMM</b>	“Generalized linear mixed model” (Modèle linéaire mixte généralisé)
<b>GMM</b>	“Growth mixture model” (Modèle de mélange de croissance)
<b>GRoLTS</b>	“Guidelines for reporting on latent trajectory studies” (Lignes directrices pour les rapports sur les études de trajectoire latente)
<b>ICC</b>	“Intra-cluster correlation coefficient” (Coefficient de corrélation intra classe)
<b>JDCS</b>	“Job demand-control-support” (Demande psychologique/latitude décisionnelle/soutien social au travail)
<b>LCA</b>	“Latent class analysis” (Analyse des classes latentes)
<b>LM</b>	“Linear model” (Modèle linéaire)

<b>LMM</b>	“Linear mixed model” (Modèle linéaire mixte)
<b>MANOVA</b>	“Multivariate analysis of variance” (Analyse de variance multivariée)
<b>MAR</b>	“Missing at random” (Manquant au hasard)
<b>MCAR</b>	“Missing completely at random” (Manquant complètement au hasard)
<b>MNAR</b>	“Missing not at random” (Ne manquant pas au hasard)
<b>OMS</b>	Organisation mondiale de la santé
<b>OR</b>	Odds ratio
<b>PWI</b>	“Psychosocial well-being index” (Indice de bien-être psychosocial)
<b>Q1</b>	“First quartile” (Premier quartile)
<b>QVT</b>	Qualité de vie au travail
<b>REF</b>	“Reference” (Référence)
<b>RMSEA</b>	“Root mean square error of approximation” (Racine de l'erreur quadratique moyenne)
<b>RPS</b>	Risques psychosociaux
<b>SD</b>	“Standard deviation” (Ecart-type)
<b>SEM</b>	“Structural equation model” (Modèle d'équation structurelle)
<b>STROBE</b>	“Strengthening the reporting of observational studies in epidemiology” (Renforcer les rapports d'études observationnelles en épidémiologie)
<b>T1</b>	“First tercile” (Premier tercile)
<b>VAS</b>	“Visual analog scale” (Echelle visuelle analogique)



# Introduction et revue de la littérature

# Introduction et revue de la littérature

## 1. Santé et qualité de vie des travailleurs

« La santé est un état de complet bien-être physique, mental et social, et ne consiste pas seulement en une absence de maladie ou d'infirmité ». Cette définition, inscrite dans le Préambule de la Constitution par l'Organisation mondiale de la santé (OMS) pour la toute première fois en 1946 [1], et toujours présente à ce jour, met en exergue l'aspect multidimensionnel de la santé. Selon l'OMS, un individu en parfait état de santé physique mais dont l'état de bien-être mental et/ou social est altéré n'est pas en bonne santé. Cette définition de la santé a exercé une influence majeure sur la manière dont la santé est comprise et abordée dans le monde entier. Elle met en avant le besoin d'approcher la santé dans son ensemble, en considérant simultanément le bien-être physique, mental et social des patients. En outre, cette définition explique que la santé d'un individu ne se limite pas à l'absence d'un mal-être à un moment donné, mais se définit un état dynamique, susceptible d'évoluer dans le temps, au grès de l'environnement dans lequel l'individu évolue. Cela souligne l'importance de suivre l'évolution de la santé des patients au fil du temps afin de mieux comprendre leur état de santé à chaque temps et de leur fournir des soins adaptés et efficaces. Parmi les trois axes qui définissent la santé humaine, le bien-être social d'un individu se caractérise par un état de satisfaction, de prospérité, et de qualité de vie au sein d'une société. En fait, le bien-être social, ou plus largement le bien-être psychosocial [2], se rapproche de la notion de qualité de vie [3]. Le concept de qualité de vie, popularisé au cours du siècle dernier, élargit la définition de la santé pour y inclure des aspects de la vie qui, en général, ne sont pas intuitivement associés à la santé, bien qu'ils soient reconnus comme pouvant l'altérer [4]. Comme pour la santé, la notion de qualité de vie fait écho au bien-être physique, mental et social des individus. En outre, elle intègre des indicateurs objectifs et subjectifs, des valeurs individuelles et un large panel de domaines de la vie telles des dimensions psychologiques, culturelles et environnementales qui

contribuent à la perception globale du bien-être. En milieu professionnel, la qualité de vie au travail (QVT) désigne un concept qui englobe les conditions dans lesquelles les employés exercent leur travail et leur impact sur leur bien-être physique, mental et social, autrement dit leur qualité de vie [5]. Les deux objectifs majeurs de la QVT sont (1) d'améliorer la qualité de l'expérience du travailleur et (2) de garantir la productivité et le bien-être des organisations [6]. Veiller à ce que les travailleurs jouissent d'un haut niveau de QVT entraîne de nombreux avantages, notamment une plus grande implication et participation au travail, ainsi qu'une satisfaction professionnelle accrue [7]. Parallèlement, cela contribue à atténuer les effets néfastes sur leur bien-être et celui de l'entreprise, tels les risques d'absentéisme et de turnover [7]. En résumé, l'objectif principal d'une démarche de QVT consiste à instaurer un milieu de travail où les collaborateurs se sentent appréciés, respectés, épanouis et motivés, tout en préservant un équilibre harmonieux entre leurs obligations professionnelles et leur vie personnelle. La QVT englobe de nombreux facteurs tels que le stress, la satisfaction au travail, l'autonomie, le sens du travail, le soutien social, la justice et l'équité, l'anxiété, le bonheur, les perspectives professionnelles, la charge de travail, l'équilibre entre vie professionnelle et vie personnelle, la reconnaissance, l'environnement de travail, la culture et le climat de l'organisation, ou encore les relations sociales [8–13]. Les institutions se sont saisies de cette question, à l'image du rapport Gollac qui recense différents facteurs d'origine psychosociale à risque pour la santé des travailleurs [14]. Des approches telles que l'échelle de bien-être psychosocial de Ryff [15], le Copenhagen Psychosocial Questionnaire [16], et l'échelle de stress perçu [17] existent déjà pour évaluer la qualité de vie des travailleurs. Portant, il demeure pertinent de développer de nouvelles méthodes d'évaluation se concentrant davantage sur les caractéristiques propres au milieu professionnel, le tout dans des approches holistiques, tenant compte des relations qui existent entre les différents facteurs. Cela devrait permettre de mieux appréhender les liens entre facteurs de risque et QVT, mais aussi les conséquences de l'amélioration ou de la dégradation de la santé des travailleurs sur le bien-être des organisations.

La plupart des recherches portant sur la QVT adoptant une approche transversale, basculer vers des études longitudinales fournirait un niveau de preuve accru.

## **2. Risques et facteurs psychosociaux en milieu professionnel**

### **2.1. Origines, enjeux pour les entreprises et exemples de risque**

Les risques psychosociaux (RPS) désignent toute situation, condition ou élément inhérent au milieu professionnel pouvant influencer la santé mentale, le bien-être psychologique ou le comportement d'un individu, donc sa QVT. Contrairement aux risques chimiques, biologiques, ou physiques, qui sont généralement liés à des contextes spécifiques de l'environnement de travail, les RPS peuvent émerger dans n'importe quel lieu de travail et dans toute sorte de situation [18]. Par exemple, dans une entreprise, le stockage de produits chimiques peut représenter un risque pour les travailleurs qui opèrent à proximité des zones de stockage, mais ne concerne pas nécessairement les autres employés de l'entreprise. En revanche, des RPS tels que la surcharge de travail ou les interruptions fréquentes peuvent affecter tous les travailleurs quel que soit leur activité, leur statut socio-professionnel ou leur environnement de travail. Il n'existe pas de consensus clair concernant la classification des RPS. Des risques comme le stress [19] ou l'insatisfaction au travail [20], en raison de leur potentiel effet néfaste sur la qualité de vie des travailleurs, peuvent être considérés comme faisant partie des RPS.

Le stress professionnel, ou stress lié au travail, désigne la pression physique, émotionnelle ou mentale que les individus ressentent dans leur milieu professionnel. Il se manifeste lorsque les individus sont soumis à des demandes et des contraintes au travail qui excèdent leurs compétences et leurs ressources, les mettant ainsi à l'épreuve quant à leur capacité à y faire face [21]. La gestion du stress représente un défi crucial pour les entreprises, influençant directement la performance des employés et l'efficacité organisationnelle [22]. En plus de ses effets délétères sur le bien-être physique et mental des individus [23], notamment sur le risque cardiovasculaire [24,22,25], le stress professionnel peut induire une baisse de productivité [26], et une augmentation du risque d'absentéisme [27] et de turnover [28]. La littérature fait état de

deux types de stress : l'eustress, qui est positif, et la détresse, qui est négative, ayant chacun des origines et des implications variées. Le « stress positif », ou eustress, peut aider les individus à faire face aux sources de stress et à s'adapter à leur environnement, là où la détresse, quant à elle, peut profondément impacter la santé des individus [29]. Si l'eustress semble être un facteur de la croissance, la détresse tend à entraver le progrès [30]. Divers outils existent pour évaluer le stress professionnel, tels des questionnaires [31] ou des échelles visuelles analogiques [32]. Ces évaluations, toutefois, reposent sur des mesures auto-déclarées et donc subjectives du stress, les rendant sujettes à des biais dans la manière dont les individus perçoivent et ressentent le stress.

Parmi les dimensions les plus étudiées en lien avec le stress, et reconnu comme le résultat positif d'une politique de gestion des RPS en milieu professionnel, la satisfaction au travail, ou à l'opposée l'insatisfaction au travail, « peut être définie comme l'évaluation de l'appréciation ou de la non-appréciation du travail par les individus » [33]. La satisfaction au travail résulte de l'appréciation personnelle et subjective qu'un individu fait de l'équilibre entre ses attentes professionnelles et ce que son emploi lui offre comme opportunités d'épanouissement [34]. C'est un marqueur primordial de la santé au travail, de plus en plus employé pour évaluer divers aspects professionnels, ce qui en fait un indicateur crucial à l'échelle mondiale dans de nombreux milieux professionnels [20]. La littérature fait état d'un effet bénéfique de la satisfaction au travail sur le bien-être et la santé des travailleurs [35] ou leur productivité [36].

En revanche, l'insatisfaction au travail peut entraîner un accroissement du niveau de stress chez les travailleurs [37], une augmentation de l'absentéisme [38,39] ou du risque de turnover [40].

A ce jour, il n'existe que peu d'outils d'évaluation de la satisfaction au travail des employés, hormis le Minnesota (Job) Satisfaction Questionnaire. Conçu pour obtenir des données précises sur les éléments du travail que les individus trouvent gratifiants, il fournit ainsi une perspective plus approfondie que les évaluations générales de la satisfaction professionnelle [41]. Étant donné que la satisfaction au travail est une évaluation subjective, l'utilisation d'une échelle

visuelle analogique pourrait fournir une mesure pertinente de ce ressenti, bien que potentiellement biaisée.

Les RPS sont associés à des facteurs psychologiques et sociaux, aussi appelés facteurs de RPS ou facteurs psychosociaux, et peuvent contribuer au développement de problèmes de santé chez les individus. La gestion des RPS en milieu professionnel revêt une importance cruciale pour prévenir des troubles de santé mentale chez les travailleurs et encourager un environnement de travail sain [42]. Un lieu de travail qui prend en compte les aspects psychosociaux peut contribuer à accroître la satisfaction des employés, à stimuler leur productivité et à les maintenir engagés [23]. La charge de la lutte contre les RPS revient à l'employeur qui doit mettre en place des politiques, des initiatives et des pratiques destinées à réduire l'exposition des salariés à ces risques, à promouvoir la santé mentale et à apporter un soutien approprié aux employés en cas de besoin. Cependant, il demeure essentiel pour les employeurs d'établir si certains groupes de travailleurs sont potentiellement plus vulnérables sur le plan psychosocial que d'autres, en particulier en examinant si le profil sociodémographique des travailleurs peut influencer leur exposition aux RPS et l'impact des facteurs psychosociaux. Cette démarche leur permettrait d'adresser spécifiquement les besoins de chaque groupe et de mettre en place des mesures préventives adaptées en conséquence.

## **2.2. Les principaux facteurs de risques psychosociaux**

Plusieurs approches existent pour évaluer les facteurs psychosociaux présents dans l'environnement de travail et analyser leur impact sur la santé des travailleurs. Deux approches, les modèles de Karasek et de Siegrist, figurent parmi les plus couramment utilisées. Ces modèles font appel à différents facteurs psychosociaux pour évaluer les RPS sur le lieu de travail. Le modèle de Karasek permet d'identifier les situations de travail à risque pour les travailleurs en se basant sur les trois dimensions psychosociales que sont la demande psychologique, la latitude décisionnelle et le soutien social au travail. Il explique ainsi qu'un travailleur en situation de forte demande psychologique et de faible latitude décisionnelle se

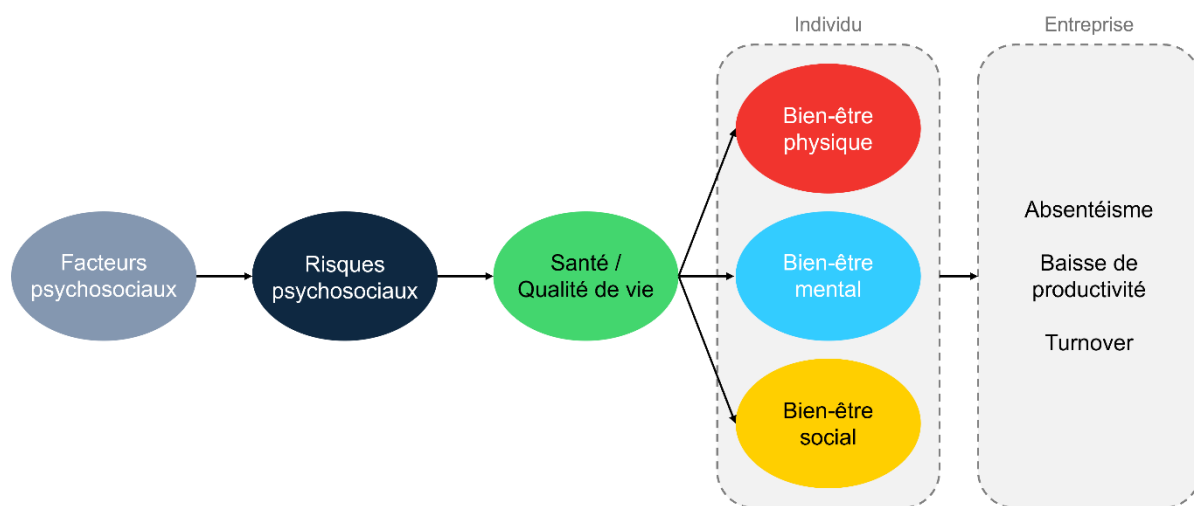
trouve dans une situation de tension au travail, aussi appelée job strain [43,44]. Si en plus d'être en situation de job strain, le travailleur perçoit un faible soutien social, il se trouve dans une situation d'isostrain [45]. Comme le modèle de Karasek, le modèle « déséquilibre effort / récompense » de Siegrist permet de détecter et d'identifier les travailleurs en situation de job strain, mais se base sur d'autres facteurs du travail [46–48]. Le modèle est né du constat qu'un déséquilibre entre les efforts demandés à un travailleur dans l'exercice de son activité et les récompenses qu'il perçoit en retour peut engendrer une altération de son état émotionnel et physiologique. Ces deux modèles, bien que différents dans leur approche, ont contribué à offrir une meilleure compréhension des facteurs psychosociaux au travail et de leur impact sur la santé des travailleurs. Ils ont été utilisés dans de nombreuses études de recherche et ont exercé une influence majeure dans les pratiques de gestion des ressources humaines visant à améliorer les conditions de travail et le bien-être des travailleurs. Ces modèles ont également permis de mettre en évidence les populations de travailleurs les plus à risques de stress selon les caractéristiques personnelles (âge, sexe, ...), le secteur d'activité, ou encore la catégorie socio-professionnelle (cadre, employé, ...) [49]. Aussi efficaces qu'ils soient pour évaluer les facteurs psychosociaux sur le lieu de travail, les modèles de Karasek et de Siegrist ne couvrent cependant qu'une infime partie de l'ensemble des facteurs de risques. Même si la littérature sur le sujet est vaste, il reste ambitieux de prétendre pouvoir énumérer tous les facteurs de RPS. Parmi les classes de facteurs psychosociaux souvent décrites dans la littérature, on trouve les facteurs liés aux comportements offensifs sur le lieu de travail, aux exigences du travail, à l'interface « vie privée – vie travail », à la justice et au respect, à l'organisation et au contenu du travail, aux relations sociales et au leadership, à la santé et au bien-être, et aux valeurs professionnelles [50–52]. L'évolution rapide du monde du travail et les changements engendrés dans nos modes et conditions de travail nous encouragent à penser que de nouveaux risques restent encore à identifier et que cette classification peut encore évoluer dans les décennies à venir [53]. Par ailleurs, à l'image des modèles de Karasek et de Siegrist, il est tout à fait plausible que les

travailleurs soient exposés à des RPS en raison de l'interaction, tant simultanée que différée, de multiples facteurs de risque. Il reste donc plus que pertinent d'analyser comment ces différents facteurs interagissent et impactent les travailleurs, à court, moyen et long terme.

### **2.3. Les conséquences sanitaires et économiques des risques psychosociaux**

Les conséquences associées aux facteurs psychosociaux sur le bien-être des travailleurs et des organisations font l'objet d'une étude approfondie depuis de nombreuses années dans le domaine de la recherche scientifique. On sait par exemple que l'exposition prolongée des travailleurs à une situation de travail exigeante sur le plan émotionnel peut se traduire dans des comportements stressants ou violents [54]. Par effet domino, ces comportements peuvent avoir des répercussions tant sur la santé mentale que sur la santé physique des travailleurs [55], voir leur être fatals [56]. Sur le plan psychique, les RPS sont associés à un risque accru de troubles mentaux tels dépression et anxiété [57,58]. Les comportements violents sur le lieu de travail peuvent représenter des évènements traumatisants pour les travailleurs et les mener à développer des troubles du stress post-traumatique [59]. Les travailleurs soumis à des conditions psychosociales exigeantes peuvent subir un épuisement professionnel plus important les menant au burnout [60]. Sur le plan physique, les RPS sont associés à un risque accru de troubles musculosquelettiques [61,62]. Les RPS sont également associés à d'autres comportements néfastes pour leur santé des travailleurs comme le tabagisme ou l'absence d'activité physique [63]. Le stress est connu pour être facteur de risque de maladie cardiovasculaire [64,65] et d'obésité [63,66], voire de mortalité [67,68]. Enfin, les travailleurs confrontés à des RPS sont susceptibles de connaître une baisse de productivité, une perte d'engagement au travail, une augmentation de comportements absentéistes ou un risque accru de turnover [69–71] (Figure 1).





**Figure 1.** Facteurs et risques psychosociaux, conséquences sanitaires et économiques

L'absentéisme sur le lieu de travail se définit par « l'absence d'une personne à un endroit et à un moment donnés, alors qu'il y a une attente sociale à ce qu'elle s'y trouve » [72]. En d'autres termes, l'absentéisme désigne l'absence imprévue d'un travailleur à son poste de travail. L'absentéisme regroupe, de manière non-exhaustive, les arrêts pour cause de maladie, les arrêts à la suite d'un accident de travail ou d'une maladie professionnelle, ou toute absence non justifiée par le salarié. Par définition, les absences imprévues s'opposent aux absences prévues comme les congés payés, les réductions du temps de travail, les jours de formation ou les jours de repos. Le phénomène de l'absentéisme est un enjeu majeur pour les entreprises tant les risques pour la santé des salariés et pour la perte de productivité globale sont conséquents [73,74]. Les coûts engendrés par l'absentéisme n'incombent pas seulement aux entreprises mais aussi aux salariés [75]. Pour lutter contre l'absentéisme et promouvoir le bien-être des employés, les entreprises disposent de nombreux leviers. Cela comprend la mise en place de programmes de prévention, de politiques de congés maladie ou de programmes de bien-être des employés. La prévention, par exemple, est particulièrement efficace pour lutter contre les RPS [23,76]. Les employeurs disposent de différents moyens pour évaluer l'efficacité de stratégies de prévention qu'ils mettent en place. L'évaluation des RPS, la captation des ressentis et des remarques et suggestions des employés, le suivi des indicateurs clés, l'évaluation du processus

et des résultats de leurs interventions, et la conduite de recherches sur des méthodologies efficaces sont autant d'outils à disposition des employeurs. Ces mesures aident les employeurs à évaluer le succès de leurs efforts en matière de gestion des RPS sur le lieu de travail. De plus, l'implication conjointe des salariés et des employeurs garantit une évaluation plus efficace des RPS sur le lieu de travail [23,77].

### **3. La médecine 4P au service de la santé au travail**

La médecine 4P, pour prédictive, préventive, personnalisée et participative, est une forme de médecine théorisée au début des années 2000 qui vise à appréhender la santé des individus par une approche holistique [78–81]. Ses objectifs sont la quantification du bien-être, la démystification de la maladie, et la diminution des coûts des soins en santé [79,82]. Cette forme de médecine s'appuie sur l'approche systémique de la maladie, sur les technologies émergentes et sur les outils d'analyse statistique pour garantir le bien-être des individus. La médecine 4P repose sur trois piliers : (1) la capacité croissante de la médecine systémique à comprendre la complexité des maladies, (2) les avancées en termes de collecte, de stockage, d'analyse et de partage de données, et (3) l'attention croissante portée par les individus à leur santé, faisant d'eux non plus seulement les simples bénéficiaires mais les acteurs de l'amélioration de leur propre état de santé [81]. Les implications de médecine 4P sont multiples et se traduisent notamment dans le développement de nouveaux systèmes et de nouvelles technologies, si bien qu'on peut considérer qu'elle participe activement à la digitalisation de la médecine [79–81]. De toutes les disciplines couvertes par la médecine, la médecine du travail est celle qui a su le mieux s'emparer des concepts de la médecine 4P pour faire évoluer son approche et améliorer l'efficacité et la qualité des soins prodigués aux travailleurs [83]. Une approche possible contre les RPS en milieu professionnel et fidèle aux principes de la médecine 4P consiste en (1) l'identification des situations à risques, (2) la mise en débat entre tous les salariés d'une évolution dans le fonctionnement organisationnel de l'entreprise, et dans (3) l'accompagnement de cette démarche. Encourager les salariés à participer à la résolution des problèmes auxquels

ils font face peut être un atout pour traiter les RPS. Plus précisément, tenir compte du vécu et des expériences des travailleurs constitue un moyen efficace pour accroître leur engagement, ce qui facilite la mise en œuvre d'interventions sur leur lieu de travail [42].

#### **4. Wittyfit : un outil de mesure de la qualité de vie au travail**

Wittyfit est une plateforme collaborative de ressources humaines éditée par l'entreprise Cegid en partenariat public-privé avec le CHU de Clermont-Ferrand [84]. A destination des entreprises, le logiciel permet de comprendre de manière holistique les problèmes rencontrés par les travailleurs. Wittyfit propose diverses fonctionnalités pour évaluer, surveiller et promouvoir le bien-être des travailleurs, tout en facilitant la communication et la collaboration au sein de l'entreprise. L'ambition de Wittyfit est de garantir et d'améliorer la santé des travailleurs et la productivité des entreprises, ainsi que de mesurer et prévenir les RPS sur le lieu de travail. La plateforme permet à ses utilisateurs, sur une base volontaire et anonyme, quand et autant de fois qu'ils le souhaitent, d'évaluer leur ressenti dans diverses thématiques liées à la santé et au bien-être au travail au moyen de questionnaires et de sondages. La plateforme utilise l'échelle visuelle analogique comme principal outil de mesure, ce qui permet de capter la perception subjective des individus [85]. L'adoption des échelles visuelles analogiques, bien que beaucoup restent à valider, découle principalement de leur complétion rapide en comparaison avec les questionnaires traditionnels, rendant ainsi le processus d'évaluation plus efficace et rapide. A l'issue de son évaluation, chaque utilisateur de la plateforme bénéficie d'un retour personnalisé de son bien-être actuel et de l'évolution de ces indicateurs au cours du temps, et a la possibilité de contribuer à l'amélioration de son bien-être en suggérant, émettant, ou mettant en avant des idées d'actions pouvant être mises en place par son entreprise. Dans le même temps, les employeurs et managers reçoivent un bilan anonymisé de l'état de santé de leurs équipes et peuvent mettre en place des actions, suggérées par les employés ou non, ciblant une ou plusieurs thématiques afin d'améliorer la QVT de leurs salariés. A l'issue de la mise en place d'une action, les travailleurs ciblés par l'action sont invités

à l'évaluer, ce qui permet aux employeurs de mesurer son efficacité. Cette démarche continue permet d'adresser un à un les problèmes rencontrés par les travailleurs et de répondre à leurs besoins afin d'améliorer leur bien-être de manière progressive et cohérente. Des captures d'écran issues des différents modules de la plateforme sont présentées en Annexe A.

Par sa démarche d'amélioration continue et validée scientifiquement, la plateforme Wittyfit apparaît comme une application directe de la médecine 4P. En encourageant la participation active et volontaire des travailleurs à l'évaluation de leur propre état de santé et en leur fournissant un retour personnalisé, elle s'inscrit dans les composantes participative et personnalisée de la médecine 4P. En mettant en avant aux employeurs et aux managers leurs équipes en difficulté, les points forts et les axes d'amélioration, Wittyfit se positionne comme un outil de prévention efficace contre les RPS. Comme elle permet de suivre l'évolution du ressenti des travailleurs dans le temps, l'ambition de la plateforme est de prédire et d'anticiper l'évolution future de la QVT des travailleurs afin de les maintenir dans un état de bien-être psychosocial. Par sa volonté de mesurer le bien-être des travailleurs, d'identifier les facteurs à risque pour leur santé et de les prévenir de ces risques, ce qui par conséquent participe à diminuer les coûts de soins en santé, les objectifs de Wittyfit sont en effet pleinement alignés avec ceux de la médecine 4P. Analyser les données de Wittyfit pourrait consolider l'importance de la médecine 4P en fournissant une vue approfondie sur les liens entre les facteurs psychosociaux, les RPS et la santé des travailleurs, ainsi qu'en facilitant l'élaboration de stratégies de prévention plus efficaces et plus adaptées.

## **5. L'analyse des données de santé au travail**

La base de données de Wittyfit est une base de données cliniques conséquente, dont l'analyse peut nourrir la connaissance autour de la QVT. Par exemple, elle pourrait aider à identifier les RPS auxquels les travailleurs sont confrontés et leurs facteurs. Toutefois, analyser les données d'une base comme la base Wittyfit n'est pas sans défi et nécessite de bien appréhender des concepts spécifiques à l'analyse de données de santé au travail.

### **5.1. Données du monde réel et données de santé au travail : définitions et concepts**

L'émergence des nouvelles technologies numériques a révolutionné de nombreux secteurs, y compris celui de la santé au travail. Ces avancées ont permis une transformation majeure dans la prestation des soins de santé, la collection et la gestion des données médicales, la surveillance des patients et la promotion du bien-être. La démocratisation de ces technologies, combinée à la capacité de collecter des quantités massives de données individuelles, a favorisé l'émergence de nouvelles approches de santé plus préventives et personnalisées, contribuant à une amélioration significative des soins fournis aux individus. Par exemple, les smartphones et applications connectées permettent un suivi constant des utilisateurs et peuvent leur suggérer des programmes et des interventions personnalisées [86,87]. L'accessibilité généralisée d'Internet a permis la création et la mise en œuvre d'initiatives de promotion de la santé sur le lieu de travail adaptées à chaque individu [88]. De telles données, collectées en dehors de cadres non-conventionnels, sont appelées données du monde réel [89]. Les données du monde réel jouent un rôle significatif dans la constitution des bases de données qui alimentent ce qu'on désigne comme les Big Data en matière de soins de santé [90]. Parce que ces bases de données sont conçues principalement pour collecter de la donnée sur le long-terme dans le but de suivre les changements qui peuvent s'opérer dans une population donnée, les données du monde réel présentent en général une nature longitudinale [91]. Les Big Data peuvent être considérées comme l'infrastructure fondamentale de la conduite des études de preuves du monde réel [92]. Souvent opposées aux essais cliniques randomisés menés dans un environnement clinique contrôlé et réglementé [93], les études de preuves du monde réel sont aussi vues comme leur complément. En adressant nombre de limitations, par exemple la taille de la population d'étude qui est souvent restreinte [92], les études de preuves du monde réel jouent un rôle déterminant dans l'aide à la prise de décision.

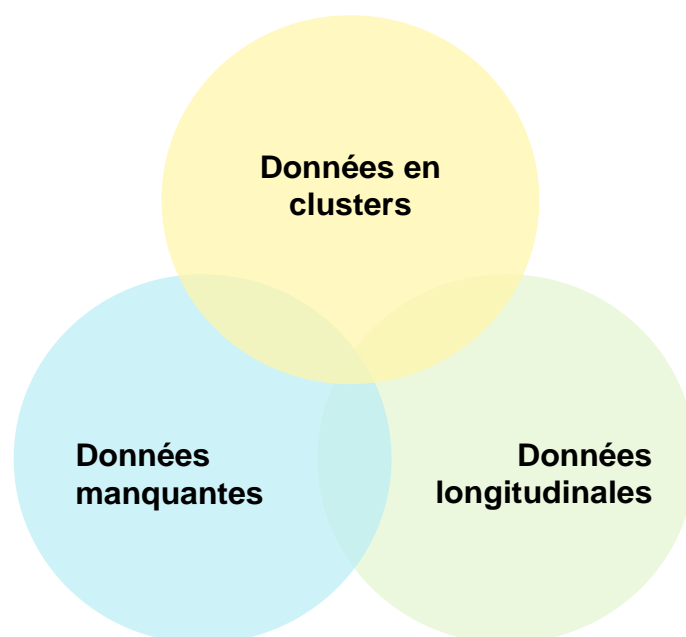
On qualifie de donnée de santé au travail toute donnée à caractère personnel relative à la santé d'un individu, révélant des informations sur son état de santé. Elles englobent un large éventail

de facteurs autour des trois aspects de la santé (physique, mentale, sociale). Ces informations peuvent être extraites de diverses origines, incluant les dossiers médicaux des travailleurs, les enquêtes menées sur le lieu de travail, les dispositifs portables, ainsi que les systèmes de surveillance environnementale. En d'autres termes, les données de santé au travail constituent une sous-famille des données du monde réel, et partagent donc des caractéristiques communes comme leur aspect longitudinale. L'analyse de ces données revêt une importance capitale et trouve de nombreuses applications. Cela peut aider tant, de manière non-exhaustive, à évaluer le bien-être des travailleurs [94], à identifier et évaluer les risques qu'ils encourent [95], à étudier des événements spécifiques au milieu professionnel tels que les accidents du travail et les maladies professionnelles et en identifier les causes [96], à évaluer les effets à long-terme des interventions conduites sur le lieu de travail [97], qu'à effectuer un suivi des patients atteints de maladie [98].

## **5.2. Les enjeux méthodologiques de l'analyse de données de santé au travail**

Bien que les données de santé au travail soient riches et variées en termes d'informations, leur nature les rend complexes à analyser et exige une compréhension approfondie des concepts spécifiques inhérents à ce type de données. L'analyse de données de santé au travail se heurte à trois principaux défis : (1) la gestion des données en clusters (ou données corrélées en grappes), (2) le traitement des données manquantes, et (3) l'analyse de données longitudinales elle-même qui est soumise à de nombreuses problématiques, par exemple la non-stabilité ou la non-stationnarité des données, ou encore l'attrition. La principale difficulté à laquelle font face les chercheurs lorsqu'ils souhaitent analyser des données de santé au travail réside dans le chevauchement de ces trois problématiques bien souvent liées entre elles (Figure 2). Par exemple, la plupart des modèles pour données longitudinales ne peuvent être appliqués que sur des jeux de données complets, c'est-à-dire ne contenant aucune donnée manquante [99]. De même, en présence de données corrélées au cluster, une méthode d'imputation globale nécessite d'être menée dans chaque cluster de manière indépendante pour être efficace [100]. L'analyse

des données de santé au travail nécessite donc de traiter chacune des trois problématiques simultanément afin de garantir la fiabilité des résultats obtenus.



**Figure 2.** Enjeux méthodologiques de l'analyse de données de santé au travail

### ***5.2.1. Modélisation des données longitudinales***

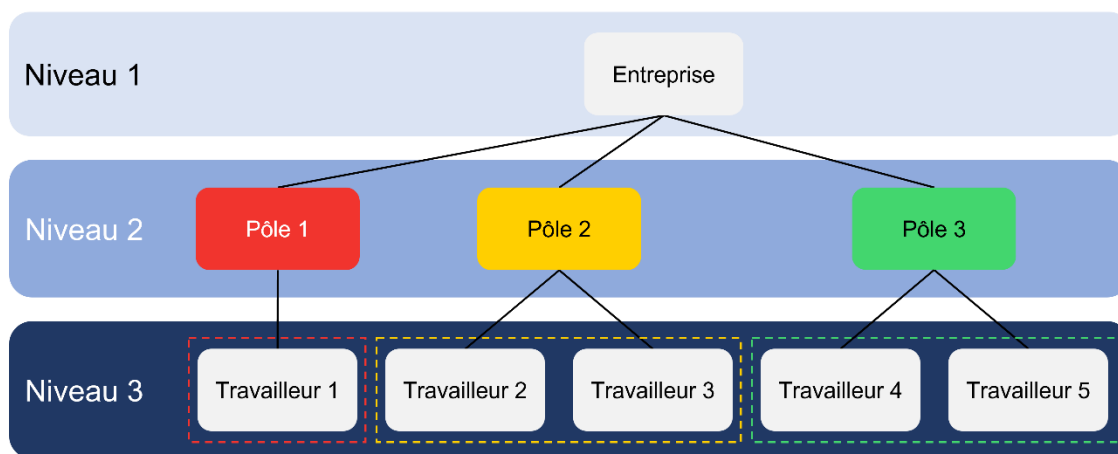
Contrairement à une étude transversale qui se situe à un moment déterminé dans le temps, une étude longitudinale permet d'étudier les changements d'un panel de variables et d'individus dans le temps. Ce type d'étude statistique est employé dans une variété de domaines, allant de la médecine à l'économie, en passant par l'épidémiologie ou les sciences sociales. Par exemple, en médecine, des études longitudinales sont fréquemment menées pour observer l'influence de facteurs dans le développement d'une maladie ou pour mesurer l'efficacité de traitements à long terme. Une étude longitudinale repose sur l'analyse de données répétées, c'est-à-dire des données recueillies à plusieurs moments dans le temps pour chaque individu. Si elles présentent de nombreux avantages, les études longitudinales font face à de nombreux défis qu'il convient aux chercheurs d'appréhender afin de garantir la fiabilité de leurs résultats. La conduite d'études longitudinales exige un suivi des participants sur une durée prolongée, une tâche qui s'avère bien souvent délicate. La perte de suivi ou l'abandon des participants constituent une perte

d'information et une absence de donnée qui peut affecter les résultats de l'étude [101]. La sélection des moments opportuns pour la collecte de données ainsi que le choix du nombre optimal de mesures peuvent également s'avérer challengeant pour les chercheurs [102]. Les chercheurs doivent également avoir recours à des modèles plus spécifiques afin de mesurer de manière simultanée mais séparée les effets de cohorte d'un côté et les effets du temps de l'autre [103]. Sélectionner le modèle approprié peut être une tâche complexe, en particulier pour les chercheurs qui ne sont pas familiers avec ces techniques [102].

### ***5.2.2. Données corrélées en grappes et effets cluster***

En analyse de données, un cluster désigne « un ensemble d'observations sur différentes unités [...] partageant certaines caractéristiques communes » [104]. Plus simplement, un cluster désigne un sous-groupe naturel d'individus dont les données intrinsèques présentent des similarités. La présence de clusters dans une population induit une hiérarchisation de la structure des données à deux niveaux ou plus et où chaque individu, en plus d'être affilié à un cluster, se situe à un niveau spécifique de la structure [105]. Si les données des individus d'un même cluster sont fortement corrélées, les données des individus appartenant à des clusters différents sont indépendantes et donc décorrélés. Par exemple, les travailleurs d'une entreprise répartis au sein de différents pôles peuvent être classés selon une structure hiérarchique (Figure 3). Dans cette population, on peut imaginer que les travailleurs d'un même pôle présentent un niveau de reconnaissance de leur supérieur similaire. En revanche, on peut aussi envisager que les travailleurs de deux pôles différents aient un niveau de reconnaissance différent car ils n'ont pas forcément le même supérieur hiérarchique. Dans une telle configuration, analyser le niveau de reconnaissance perçu par les travailleurs exige de prendre en compte leur pôle d'appartenance comme pouvant exercer une influence sur les observations. On dit alors que le pôle induit un effet de cluster (ou effet cluster).





**Figure 3.** Exemple de structure de données multiniveaux : population de travailleurs d’une entreprise à plusieurs pôles

En outre, dans une étude longitudinale, un cluster peut également définir « l’ensemble des données répétées d’un individu » [104]. En effet, la réponse d’un individu à un temps donné dépend de ses réponses antérieures, ce qui entraîne une intercorrélacion entre toutes ses observations, ce qui caractérise bien un cluster de données [106]. Ainsi, en présence de données en clusters, il est nécessaire de prendre en compte la structure des données et les effets des clusters induits lors des analyses. Une mauvaise spécification de la structure des données peut en effet conduire à des estimations biaisés faites par les modèles [107].

### 5.2.3. *Nature des données manquantes et méthodes d’imputation*

Il est rare pour un jeu de données d’être complet, autrement dit de ne présenter aucune donnée manquante. Les données manquantes constituent un problème majeur en analyse statistique car rendent l’ajustement des modèles plus difficile, voire impossible [108]. Le phénomène est très fréquent dans les études longitudinales car suivre un panel d’individus sur la durée est compliqué. Les études longitudinales peuvent souffrir de manques d’observation aléatoires et isolés. Cela se produit par exemple lorsque qu’un individu ne se présente pas à un rendez-vous ou en cas de défaillance d’un appareil de mesure lors d’une journée d’étude. Cependant, le plus grand problème que rencontre les études longitudinales est l’attrition, qui désigne le phénomène

où des individus quittent prématurément l'étude, entraînant une perte d'information conséquente [109].

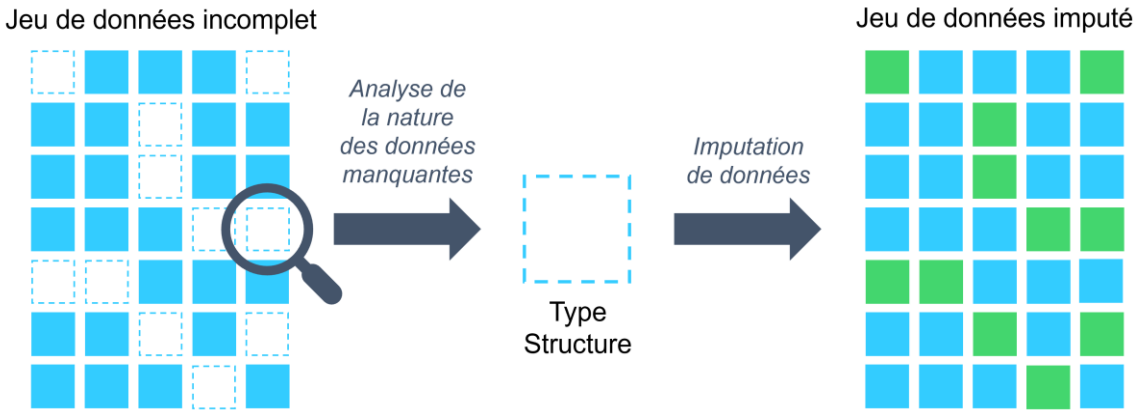
Lorsqu'ils font face à des données manquantes, les chercheurs ne peuvent pas simplement ignorer le problème et ne prendre en compte que les données dont ils disposent. En effet, une analyse des cas disponibles, voire des cas complets (i.e. uniquement les individus dont toutes les observations sont mesurées), conduit à des biais et à une perte d'efficacité des modèles [110,111]. Pour pallier ce problème, les chercheurs ont la possibilité de recourir à une imputation de données. Cette méthode consiste à assigner à chaque observation manquante une estimation de la vraie valeur, bien qu'inconnue. Plusieurs typologies de méthodes s'offrent aux chercheurs : les méthodes d'imputation simple dites *ad-hoc* [112], regroupant des méthodes comme l'imputation par la moyenne ou l'imputation aléatoire, l'interpolation polynomiale (linéaire par exemple) [113], ou encore l'imputation multiple [114,115]. Bien que le choix de la méthode soit laissé à la discrétion des chercheurs, la nature des données manquantes peut les conduire à privilégier une méthode par rapport à une autre.

La nature des données manquantes est caractérisée par deux éléments : leur type et leur structure. Le type désigne le mécanisme sous-jacent à l'absence de données comme fonction des variables observées et non-observées. La littérature fait état de trois types distincts de données manquantes : les données manquantes complètement au hasard (ou MCAR pour « missing completely at random »), les données manquantes au hasard (ou MAR pour « missing at random »), et les données ne manquant pas au hasard (ou MNAR pour « missing not at random ») [116,117]. Si le mécanisme responsable de l'absence de données est indépendant des variables observées comme des variables non-observées, alors les données manquantes induites sont MCAR. Un exemple souvent utilisé pour décrire ce mécanisme est le suivant : si dans le cadre d'un questionnaire sur leur état de santé, chaque participant décide de ne répondre à une question spécifique que s'il obtient six à un lancer de dé alors la donnée manquante induite par son échec est complètement au hasard [118]. Si cette hypothèse est très prisee car n'induisant

aucun biais dans les résultats des analyses, elle est peu plausible et donc peu considérée par les chercheurs. Si le mécanisme responsable de l'absence de données est dépendant des variables observées uniquement, alors les données manquantes induites sont MAR. Par exemple, dans le cadre du suivi des travailleurs en arrêt, si un salarié bénéficie d'une retraite anticipée pour raison médicale, ce salarié ne sera plus suivi par la médecine du travail et les observations liées aux visites en médecine du travail seront donc manquantes. Enfin, si le mécanisme responsable de l'absence de données est dépendant des variables observées et non-observées, alors les données manquantes induites sont MNAR. Un exemple courant concerne les questions filtrées dans les questionnaires, où certaines questions ne sont posées qu'à un groupe défini de répondant, ce qui induit une absence d'observation pour tous les non-répondants. En pratique, il est difficile d'identifier avec certitude le type des données manquantes, et les chercheurs doivent alors s'efforcer de justifier le choix de leur hypothèse [119]. Sous l'hypothèse de données MNAR, les chercheurs n'ont d'autre choix que de mener une analyse de sensibilité [120,121]. L'analyse de sensibilité est une approche essentielle pour évaluer la robustesse des résultats d'une étude. Cette approche consiste à observer comment les résultats sont influencés par des modifications dans les conditions et outils utilisés dans l'étude [122]. Cela peut se faire, par exemple, en comparant les résultats obtenus après imputation des données selon la méthode choisie ou en altérant volontairement le jeu de données initial.

Autre élément définissant la nature des données manquantes, la structure décrit la répartition des observations mesurées et des valeurs manquantes dans un jeu de données [117]. Une structure de données manquantes est dite monotone si les données concernées suivent un schéma ordonné par rapport à une ou plusieurs variables. Dans une telle configuration, il est possible de classer les variables selon leur nombre de données manquantes. Un tel schéma se retrouve par exemple dans les études longitudinales concernées par l'attrition [109]. En effet, dans une telle situation, le schéma des données manquantes est monotone, car une fois qu'un individu quitte l'étude, aucune observation ultérieure n'est possible sur lui. En opposition, une

structure est dite non-monotone (ou arbitraire) si elle ne suit aucun schéma spécifique. Comprendre et être capable d'identifier les différents schémas peut être utile pour repérer les liens entre les données manquantes et d'autres variables, ainsi que pour choisir les méthodes appropriées de gestion des données manquantes.



**Figure 4.** Identification de la nature des données manquantes avant imputation de données

# Enjeux et objectifs

## **Enjeux et objectifs**

Ce travail de thèse vise à approfondir les connaissances autour des RPS et de ses facteurs dans le but d'améliorer la QVT et le bien-être des travailleurs. Notre volonté de construire ce projet de thèse découle de la conviction qu'une analyse approfondie des données de santé au travail, telles que celles de la base Wittyfit, pourrait offrir une compréhension accrue des liens entre les risques et les facteurs psychosociaux, la qualité de vie au travail et le bien-être organisationnel, en mettant l'accent notamment sur l'impact sur l'absentéisme. Ainsi, nous avons analysé les données des utilisateurs de Wittyfit depuis sa création et l'inclusion des premiers participants, aux moyens de modèles statistiques complexes. En raison de la variété des thématiques liées au travail abordées et de l'augmentation constante du nombre d'utilisateurs de la plateforme, le choix de la base Wittyfit comme base de données privilégiées pour la conduire nos études était évident. Situé au croisement de l'épidémiologie, de la médico-économie et des mathématiques appliquées et statistiques, cette thèse englobe des enjeux multiples. Du point de vue épidémiologique, les objectifs de ce travail consistent, dans un premier temps, à mesurer l'exposition des travailleurs aux RPS et, dans un second temps, à évaluer l'influence de leurs facteurs, donc de facteurs de RPS, sur la qualité de vie et le bien-être des travailleurs. Nous envisageons également dans cette thèse des perspectives médico-économiques, souhaitant évaluer l'impact des facteurs psychosociaux ainsi identifiés sur l'absentéisme en milieu professionnel. Nous basculerons d'une approche transversale vers une approche longitudinale pour étendre nos résultats et garantir un niveau de preuve plus élevé de nos résultats. Du point de vue statistique, les objectifs sont de développer et d'appliquer des modèles spécifiques à l'analyse de données de santé au travail en tenant compte des défis méthodologiques rencontrés dans ce type d'analyse. Cela inclut la prise en compte des effets de clusters, la gestion des données manquantes, et tous les autres aspects liés à l'analyse de données longitudinales. Finalement, nous totalisons six études, dont quatre déjà publiées dans des revues scientifiques

indexées PubMed, une soumise, et une en cours de travail, accompagnées de quatre présentations orales (Annexes B à E).

Cette thèse se découpe en quatre chapitres comme suit : (i) un chapitre consacré à la revue de la littérature scientifique autour des données de santé au travail (article 1), (ii) un chapitre centré autour de la question des risques et facteurs psychosociaux et de l'importance du statut managérial sur l'exposition à ces risques (articles 2 à 4), et (iii) un chapitre dédié à la création d'un indice de bien-être psychosocial (articles 5), et un chapitre consacré à l'étude du lien entre bien-être psychosocial et absentéisme (article 6). Dans un premier article (article 1), nous avons conduit une revue de la littérature autour des enjeux méthodologiques liés à l'analyse longitudinale de données réelles de santé au travail, dans lequel nous avons proposé une démarche à suivre pour répondre à ces enjeux et améliorer la qualité des études longitudinales. Dans les trois articles suivants (articles 2 à 4), nous avons mesurer l'exposition des travailleurs à des RPS selon leur statut managérial (manager ou employé), et évaluer l'influence de leurs facteurs de risques. Étant donné le faible nombre d'articles de littérature sur les liens entre statut managérial et facteurs de RPS, nous avons jugé pertinent de choisir cet axe d'étude. Un deuxième article (article 2) s'est ainsi concentré sur les différences de charge de travail, d'autonomie et de soutien social de ces deux populations, ainsi que sur l'exposition au job strain et d'isostrain. Un troisième article (article 3), basé sur une approche similaire, a permis de mettre en évidence les différences de satisfaction au travail perçues entre managers et employés et de risque d'insatisfaction au travail. Nous avons également étudié l'influence de six facteurs de la satisfaction au travail, l'ambiance, la reconnaissance, l'équilibre entre vie professionnelle et vie personnelle, le sens, l'organisation du travail, et les valeurs professionnelles selon le statut managérial des travailleurs. Dans un quatrième article (article 4), nous avons étudié l'évolution du stress et de la satisfaction au travail des managers et des employés au cours du temps. Pour cela, nous avons appliqué trois modèles pour données longitudinales : les modèles linéaires mixtes, les modèles de trajectoires, et les modèles de médiation par panel à décalage croisé.

Dans un cinquième article (article 5), nous avons construit un indicateur de bien-être psychosocial au moyen d'un modèle d'équation structurelle basé sur dix facteurs psychosociaux pour étudier les liens entre facteurs psychosociaux, bien-être et risque d'absentéisme pour cause de maladie non-professionnelle. Enfin, un dernier article (article 6) a permis d'étudier l'effet du bien-être psychosocial, mesuré grâce à l'indicateur construit dans notre précédente étude, sur le risque d'absentéisme de court ou long terme dans une approche par modèle de survie et par modèle multi-états.



# Résultats et travaux de recherche

## **Résultats et travaux de recherche**

### **Article n°1**

Methodological issues in analyzing real-world longitudinal occupational health data: a useful guide to approaching the topic

Colin-Chevalier R, Dutheil F, Cambier S, Dewavrin S, Cornet T, Baker JS, Pereira B

### **Article n°2**

The protective role of job control/autonomy on mental strain of managers: a cross-sectional study among Wittyfit's users

Colin-Chevalier R, Pereira B, Benson AC, Dewavrin S, Cornet T, Dutheil F

### **Article n°3**

The managerial role and psychosocial factors of job satisfaction: a cross-sectional study among Wittyfit's users

Colin-Chevalier R, Pereira B, Dewavrin S, Cornet T, Zak M, Benson AC, Dutheil F

### **Article n°4**

Stress and job satisfaction over time, the influence of the managerial position: a bivariate longitudinal modelling of Wittyfit data

Colin-Chevalier R, Dutheil F, Benson AC, Dewavrin S, Cornet T, Lambert C, Pereira B

### **Article n°5**

Psychosocial well-being index and sick leave in the workplace: a structural equation modelling of Wittyfit data

Colin-Chevalier R, Pereira B, Dewavrin S, Cornet T, Baker JS, Dutheil F

### **Article n°6**

Psychosocial well-being as a protective factor against short- and long-term work absenteeism: from censored-data survival models to multistate analysis

Colin-Chevalier R, Dutheil F, Cambier S, Dewavrin S, Cornet T, Baker JS, Pereira B

# Article n°1

Methodological issues in analyzing  
real-world longitudinal occupational  
health data: a useful guide  
to approaching the topic

## **Article n°1**

Methodological issues in analyzing real-world longitudinal occupational health data: a useful guide to approaching the topic

Colin-Chevalier R, Dutheil F, Cambier S, Dewavrin S, Cornet T, Baker JS, Pereira B

### **Résumé**

Ce premier article méthodologique fait un état de l'art de la littérature autour de l'analyse de données du monde réel de santé au travail, et présente les enjeux méthodologiques liés à leur analyse. Les trois plus grands défis auxquels font face les chercheurs y sont exposés, à savoir les questions autour du traitement de données en clusters, de la gestion des données manquantes, et de l'analyse de données longitudinales en général, avec une attention particulière tournée autour des principaux modèles utilisés dans ce type d'études. Dans une première partie, nous avons étudié chacun des trois enjeux en présentant les problèmes engendrés et les solutions, méthodes et modèles développés dans la littérature capable d'y répondre. Dans une deuxième partie, nous avons synthétisé toutes les informations collectées pour développer et présenter une démarche capable d'adresser les trois enjeux simultanément. Dans une troisième partie, nous avons illustré nos propos au travers d'une étude de cas en utilisant les données de la plateforme Wittyfit. Une recherche bibliographique exhaustive nous a permis de mettre en évidence l'absence d'une démarche similaire dans la littérature, appuyant la pertinence de notre solution. Notre espoir est que cette nouvelle approche puisse aider les chercheurs et améliorer la qualité et la fiabilité de leurs futures études.



Review

# Methodological Issues in Analyzing Real-World Longitudinal Occupational Health Data: A Useful Guide to Approaching the Topic

Rémi Colin-Chevalier <sup>1,2,3,\*</sup> , Frédéric Dutheil <sup>1,2,3</sup> , Sébastien Cambier <sup>4</sup>, Samuel Dewavrin <sup>3</sup>, Thomas Cornet <sup>3</sup>, Julien Steven Baker <sup>5</sup> and Bruno Pereira <sup>4</sup>

<sup>1</sup> CNRS, LaPSCo, Physiological and Psychosocial Stress, Université Clermont Auvergne, F-63000 Clermont-Ferrand, France; frederic.dutheil@uca.fr

<sup>2</sup> Preventive and Occupational Medicine, CHU Clermont-Ferrand, F-63000 Clermont-Ferrand, France

<sup>3</sup> Wittyfit, F-75000 Paris, France; samuel.dewavrin@wittyfit.com (S.D.); thomas.cornet@wittyfit.com (T.C.)

<sup>4</sup> Biostatistics Unit, The Clinical Research and Innovation Direction, University Hospital of Clermont-Ferrand, CHU Clermont-Ferrand, F-63000 Clermont-Ferrand, France; scambier@chu-clermontferrand.fr (S.C.); bpereira@chu-clermontferrand.fr (B.P.)

<sup>5</sup> Centre for Health and Exercise Science Research, Hong Kong Baptist University, Kowloon Tong, Hong Kong 999077, China; jsbaker@hkbu.edu.hk

\* Correspondence: r.colin6374@gmail.com



**Citation:** Colin-Chevalier, R.; Dutheil, F.; Cambier, S.; Dewavrin, S.; Cornet, T.; Baker, J.S.; Pereira, B. Methodological Issues in Analyzing Real-World Longitudinal Occupational Health Data: A Useful Guide to Approaching the Topic. *Int. J. Environ. Res. Public Health* **2022**, *19*, 7023. <https://doi.org/10.3390/ijerph19127023>

Academic Editors: Vasile Palade, Alireza Daneshkhah, Amin Hosseinian-Far and Samer A. Kharroubi

Received: 17 May 2022

Accepted: 6 June 2022

Published: 8 June 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Ever greater technological advances and democratization of digital tools such as computers and smartphones offer researchers new possibilities to collect large amounts of health data in order to conduct clinical research. Such data, called real-world data, appears to be a perfect complement to traditional randomized clinical trials and has become more important in health decisions. Due to its longitudinal nature, real-world data is subject to specific and well-known methodological issues, namely issues with the analysis of cluster-correlated data, missing data and longitudinal data itself. These concepts have been widely discussed in the literature and many methods and solutions have been proposed to cope with these issues. As examples, mixed and trajectory models have been developed to explore longitudinal data sets, imputation methods can resolve missing data issues, and multilevel models facilitate the treatment of cluster-correlated data. Nevertheless, the analysis of real-world longitudinal occupational health data remains difficult, especially when the methodological challenges overlap. The purpose of this article is to present various solutions developed in the literature to deal with cluster-correlated data, missing data and longitudinal data, sometimes overlapped, in an occupational health context. The novelty and usefulness of our approach is supported by a step-by-step search strategy and an example from the Wittyfit database, which is an epidemiological database of occupational health data. Therefore, we hope that this article will facilitate the work of researchers in the field and improve the accuracy of future studies.

**Keywords:** methodological issues; modeling; occupational health; longitudinal data; missing data; cluster-correlated data; real-world data

## 1. Introduction

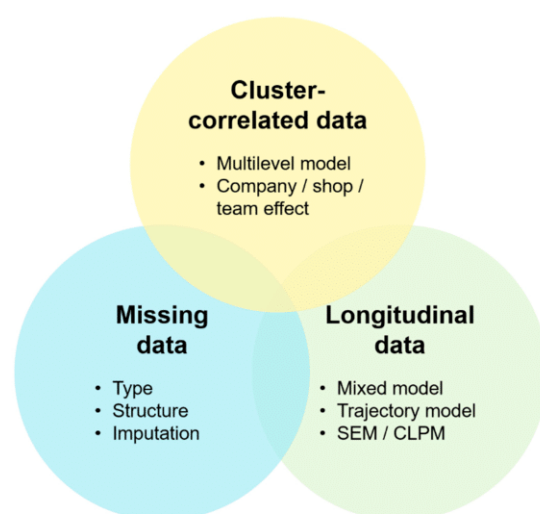
In previous decades, randomized clinical trials have been the main experimental methodology used to collect clinical data. Collected within a strict framework, the analysis of randomized clinical trials data is used to answer specific questions related to a specific population (namely the population selected for the study). This remains true provided that analyses are performed rigorously with consideration for all required hypotheses and methods (e.g., definition of a representative population sample, randomization, controlled tests, application of appropriate models to the data). However, the advances in technology, the democratization of computer tools such as computers and smartphones as well as the increase in data storage capacities offer researchers new possibilities for collecting

large amounts of health-related data for analysis. Such data, collected independently of traditional trials, is also known as real-world data [1–3]. More broadly, the data can be obtained from various sources such as patient health records, product and disease registers or even digital health data collection platforms and applications. Many examples of real-world databases can be listed and classified according to the sources from which they come. As examples of this, we can cite records from patient registries such as the European Cystic Fibrosis Society Registry database [4], healthcare databases such as Wittyfit and Wittyfit Research [5], pharmacy and health insurance databases such as the Food and Drug Administration’s Sentinel Initiative [6], social media such as PatientsLikeMe [7] or patient-powered research networks such as PCORnet [8]. Big data management approaches such as the management of chronic kidney disease [9], chronic obstructive pulmonary disease [10], or lymphoma subtypes [11] have naturally emerged from such databases.

As a perfect complement to randomized clinical trials data and presenting a longitudinal structure [12], the results of the analysis of real-world data, also called real-world evidence, are very popular; for example, within pharmacology research, these data sets are considered essential because they are crucial in pharmacological decision-making [13,14]. However, by construction, real-world databases are not immune to missing, noisy (outliers), duplicate or inconsistent data phenomena [15].

A longitudinal study is a study investigating repeated data, i.e., where the same data has been measured/collected on an individual several times over time [16]. Unlike cross-sectional studies, cohort effects and time-related effects can be measured separately. The main goal of a longitudinal study is to describe changes over time and measure the individual influence of variables to explain changes observed [17]. Longitudinal data may also facilitate the change over time when investigating particular individuals [18]. For example, it is possible to measure risk factors in the development of disease for particular individuals in the population [19]. However, many challenges and methodological issues make it difficult to analyze longitudinal data. These analytical problems include the correlated structure of intra-individual data, the considerable size of data sets, irregular time-spaced measurements, non-linear patterns (such as rapid growth or stationary responses), latent constructs, and the mix of time-varying and static covariates.

There are also more difficult problems to consider, such as cluster-correlated data, missing data and longitudinal modeling itself [20,21]. To sum up, studying such data types requires understanding the specific concepts that are outlined in Figure 1.



**Figure 1.** Data diversity of real-world longitudinal occupational health databases (see Appendix A for the novelty and usefulness of our approach and main formulas surrounding cluster-correlated data, missing data, and longitudinal data). SEM: structural equation model, CLPM: “cross-lagged” panel model.



Although a multitude of approaches have been proposed in the literature, it is still difficult to account for all these concerns in a synchronized way. The purpose of this paper is to review the state of the art of different methods and models designed to address issues related to real-world longitudinal occupational health data in a holistic way to improve the quality of future surveys on the topic. In the first part, we aim to introduce each methodological issue and to present advanced methods for addressing them. In the second part, we aim to present a case study in which all the issues overlap and present a method to meet the objectives of the study using the solutions previously featured.

## 2. Methodological Issues and Methods for the Analysis of Longitudinal Data

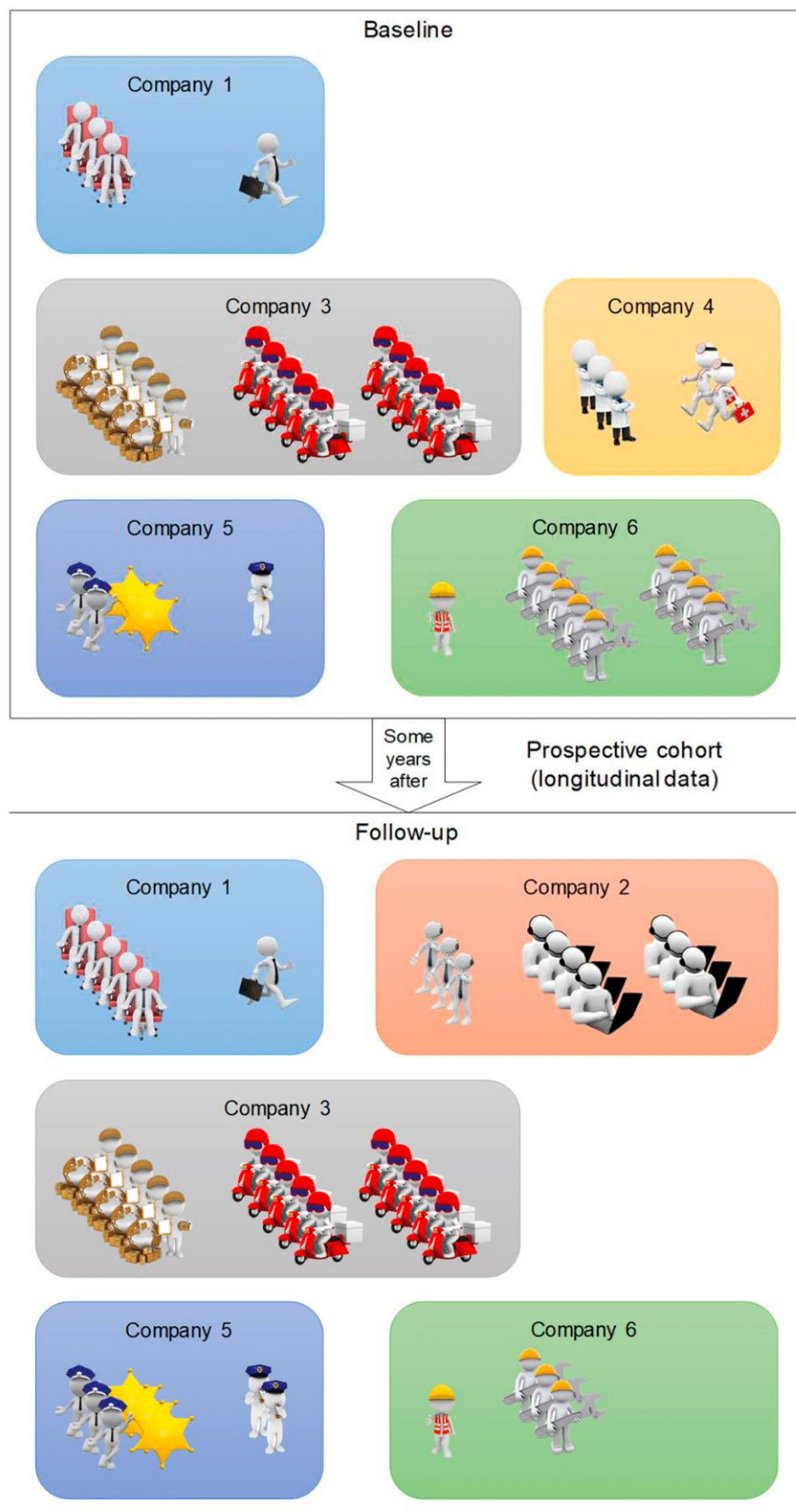
### 2.1. Cluster-Correlated Data

Cluster-correlated data occur when pre-existing groups of individuals exist within the population, allowing for a natural classification of subjects into groups, otherwise known as clusters. Thus, the observations of the same cluster are correlated with each other while the observations between the different clusters are uncorrelated. This structure requires considering two possible sources of variability, namely intra- and inter-cluster variability. These similarities in the data can, however, lead to a loss of statistical efficiency and integrity of the models used, necessitating an increase in the size of the sample [22]. It is therefore necessary to find an appropriate balance between the number of clusters and the number of patients.

Correlated data occur in a number of different settings. They can be clustered, spatial (grouping by location), multilevel or longitudinal data or even several at the same time [23]. A hierarchical or multilevel structure can be thought of as a series of several levels nested within each other. Thus, the lowest level represents level 1, and the levels follow each other upwards to the level where all the data is contained. The units formed at each level can then be likened to clusters. A hierarchical structure is therefore a succession of more and more specific clusters and, from observation of the lower levels, is a structure that is “neither accidental nor ignorable” [24]. Multilevel structures widen the field of possibilities in terms of questions that can be answered (e.g., presence of a group effect, disparities within the different clusters). On the other hand, the structures question the use of statistical methods that do not consider the nature of the structures, which could also provide erroneous results [25,26]. Longitudinal data is a special form of clustered data. By nature, an individual’s data is plausibly more positively correlated with each other than with other individuals. Like correlated data, this intra-subject correlation must be taken into account in the analyses, otherwise this can provide false positive results and erroneous confidence intervals [21,27]. Figure 2 presents an example of a population divided into clusters followed in a longitudinal study.

Thus, cluster-correlated data, regardless of its type, requires special analytical treatment [28]. Cluster-correlated data analysis attempts to take into account the variability associated with each level of the structure and must be interpreted separately from the overall variability [29]. Indeed, misspecification of the cluster effect or careless interpretation of the model parameters can lead to erroneous results [30,31]. Hence, in the literature, multilevel models (MLMs) have been designed to address these various issues and analyze correlated-clustered data [24,29,32,33]. These models are similar to mixed models. In fact, the “levels” described by the multilevel models can be assimilated to be the “random effects” of the mixed model. This implies that cluster-correlated data can be analyzed properly using the models defined above by considering the levels of the data as random effects. Similarly, for longitudinal data, this implies that the individual effect must be considered as a random effect also.





**Figure 2.** Example of cluster-correlated data: here, a prospective cohort of workers from multiple companies is followed over time.

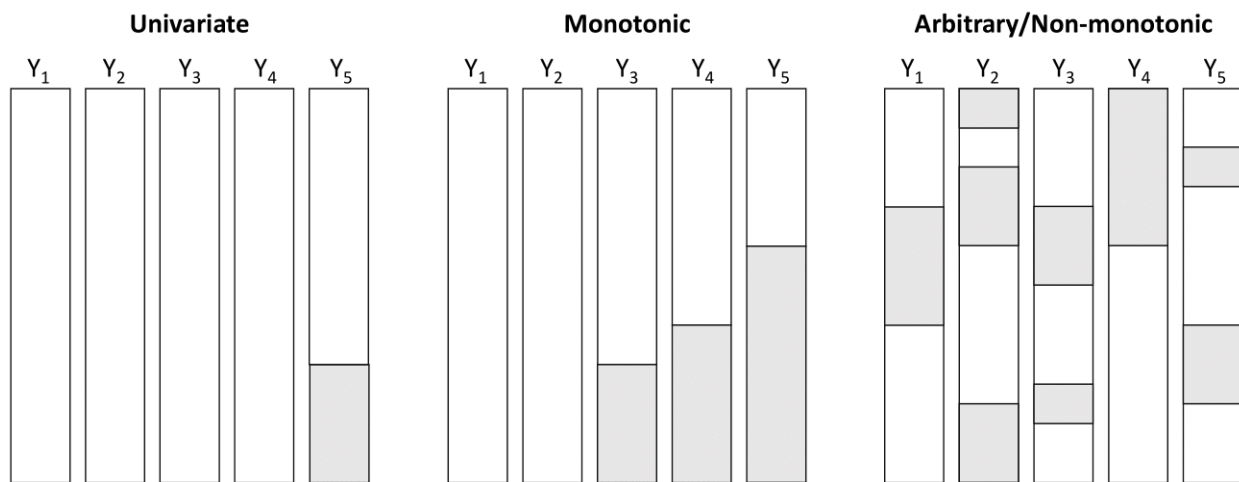
Clustered data are not immune to the phenomenon of missing data. In the absence of data, conventional imputation methods can be applied. However, the particular structure of cluster-correlated data, which leads to multiple sources of variability (intra- and inter-cluster), suggests proceeding differently. As a result, it appears preferable to apply the imputation method chosen cluster by cluster [34]. To sum up, in the presence of cluster-correlated data, it is preferable to estimate the missing values within clusters. When the data of an individual is missing, estimates can be made using individuals from the same cluster. For a hierarchical structure, the lowest level cluster in which the individual is located will be preferred. If an entire cluster leaves the study, the missing data can be estimated from the data of individuals at the level above. Finally, at the highest level of the hierarchy, it can be estimated using the entire data set.

## 2.2. Missing Data

The phenomenon of missing data is a common problem in data analysis, especially in longitudinal surveys. It is rare for a data set to be complete, particularly for long studies, either because of occasional omissions or because of subject withdrawal. As a consequence, data missingness causes three main issues [35]: it can introduce a significant amount of bias, make it more difficult to process and analyze the data and induce a loss of statistical power [36,37]. When faced with missing data, the most common used methods include the complete case analysis (or listwise deletion) and the “last observation carried forward” methods. However, these methods are too hazardous and can introduce a significant bias in the estimation of the parameter under investigation [38–40]. As previously observed, these ad hoc methods are no longer desirable or necessary. Likewise, in the presence of clustered data, it is difficult to estimate what effect the loss of an individual or even an entire cluster has on the outcome measures. Accordingly, it is better not to ignore missing values but to try to estimate them using imputation methods. The purpose of these methods is to estimate missing values from previously observed values, which can be considered as measured data and therefore used for fitting analytical models. While these methods may seem attractive, their use is not without problems [41]. Even though there are some methods that are more effective than others, it is not really possible to say which method is especially good [42]. These methods require understanding and apprehension of certain concepts that we aim to outline below.

To be more effective and valid, statistical analysis must be performed with suitable mechanisms and assumptions for missing data [37]. In other words, it is crucial to understand the process behind it to ensure the validity of the statistical inferences and the absence of bias [43]. Therefore, a careful preparation should be observed before using a method that is very similar to an identification process. In particular, it is essential to identify the two main parameters: the structure and the type of missing data.

The structure plays an important role in the choice of imputation methods that can be applied [44]. There are three possible structures, which can overlap. The structure is said to be univariate if one and only one variable of an individual has missing values. It is said to be monotonic when several variables of an individual have missing values and these variables can be classified according to their percentage. For longitudinal surveys, this structure can be found when an individual leaves the study, for example in case of attrition or even abandonment. Attrition, or random drop-out [45], is a major and common problem in longitudinal studies, representing an important challenge when modeling [46] as it may produce estimates of the effects that are often underestimated [47]. If it is not possible to obtain the data of an individual at a given time (e.g., forgotten appointment, forgotten connection), we often consider intermittent monotonic structures. The structures can qualify as arbitrary (or non-monotonic) if they are neither univariate nor monotonic. Figure 3 illustrates the different structures of missing data.



**Figure 3.** Missing data structures. White sections indicate the presence of data, shaded sections their absence.

After identifying the structure of missing data, it is necessary to deal with the identification of its type. We consider here again three different types: data which is missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR) [41]. A missing value is considered to be MCAR if its missingness does not depend on other observations, whether they are observed or not. Under this assumption, the results of the analyses appear to be generally unbiased, but it is rarely verified because it is restrictive. A less restrictive assumption supposes that if it depends on observed observations only, its type will be MAR. Otherwise, if it is neither MCAR nor MAR, it will be MNAR. Despite this formulation, it is not always possible in practice to identify the causes of data missingness and therefore to determine its type [48].

Finally, once the type and structure of the missing data are identified, an imputation method can be used. Table 1 provides a non-exhaustive overview of the different methods that can be used to calculate estimates of missing values [41,49].

**Table 1.** Imputation methods by type of missing data.

Missing Completely at Random	Missing at Random	Missing Not at Random
<p><b>Ad hoc methods</b> Complete case analysis, available-case analysis, weighting methods</p> <p><b>Single imputation</b></p> <p><b>Implicit modeling</b> Hot/cold deck imputation, substitution, composite methods</p>	<p>Expectation maximization algorithm</p> <p>Multiple imputation</p>	<p>“Sensitivity analysis”</p>
<p><b>Explicit modeling</b> Mean/regression/stochastic regression imputation</p>		

For the particular case of MNAR type data, it is necessary to perform a “sensitivity analysis”. Indeed, although there is a multitude of models to analyze this type of data, they all rely on assumptions (e.g., the missing data mechanism) which, for this type of missing data, are unverifiable. When the results are particularly sensitive to the assumptions made, it becomes difficult to choose the most suitable model. That is, instead of fitting a single model, it is wiser to consider alternative models and to assess the sensitivity of the results to the assumptions made about the missing data mechanism [31,41,48].

### 2.3. Longitudinal Data and Modeling

Repeated data models have been the subject of many years of research and development. Today, the literature on the subject is substantial and many of these models have been used in longitudinal surveys. Among the models used most for the analysis for longitudinal data [20,50], the most often used are variance models, whether univariate (ANOVA) or multivariate (MANOVA), random effects models such as mixed models or



generalized linear models (GLMs) using generalized estimating equations. Nevertheless, other models such as trajectory models or alternatives to linear and mixed models such as structural equation models (SEMs) or “cross-lagged” panel models (CLPMs) have appeared in the literature and constitute promising alternatives for the analysis of longitudinal data. These models are not all similar and are intended to answer different questions. While mixed models and their alternatives are used to estimate the impact of factors on a response variable, trajectory models are used to model individual trajectories within a population and follow their evolution over time. Given a large amount of data, modeling should be performed on a training sample and model validation on a test sample. We will now describe the above-mentioned models in more detail.

### 2.3.1. Analysis of Variance for Repeated Measures

The analysis of variance for repeated measures represents a classical method to analyze longitudinal data. Whether for ANOVA or MANOVA, they allow comparing groups' means on a dependent variable across time. However, it does not allow learning about individual trajectories. In addition to other parameters, time is also treated as an explicative variable. This means, among other things, that these models assume that individuals cannot have a proper slope over time, which is rarely true. Moreover, these models are difficult to apply in the absence of data and are applied using the complete cases or the “last observation carried forward” methods. As a consequence, it is therefore best to avoid these models and to turn to more suitable methods for repeated data analysis [27]. For example, mixed models offer more advantages [36], especially by adding a subject-specific component to the model and allowing the conduct of the analysis despite possible lack of data.

### 2.3.2. Mixed Models

Mixed models can be seen as natural extensions of regression models [51,52]. Unlike the latter, they involve random effects specific to each individual. The mean and the variance of the response are respectively modeled as a linear combination of fixed-effect and random-effect components, where the impact of the factors is weighted and associated with the coefficients of the model [53,54]. Unlike analysis of variance models, both marginal and mixed models are unconditional. This feature makes it possible to model the response variable as a function of both the covariates and time separately representing both within- and between-subject effects [20]. They are also particularly effective for making individual predictions even despite eventual missing data because the randomly missing estimates are unbiased [41]. More broadly, generalized linear mixed models combine the specificities of generalized linear models (GLMs) and mixed models, allowing for generalizing the type of the explained variable. This change in construction is ideal for the analysis of longitudinal data structures since it makes it possible to account for the intra-subject correlation through random effects [55,56].

This makes mixed models a safe and effective method for the analysis of longitudinal data. Other more specific models, as we will see below, have been designed subsequently and added to the literature, but mixed models remain the most popular. Once a trajectory is identified, it is possible to describe it over time but also in terms of static covariates using a mixed model or equivalent.

### 2.3.3. Generalized Estimating Equations

In general, the estimation of GLM parameters is based on the maximum likelihood method [57]. Yet, the generalized estimating equations method does represent a popular alternative to likelihood-based generalized linear mixed models, allowing the model to be extended to the analysis of correlated data. Afterwards, they simultaneously model the link between explicative variables, the response and the within-subject dependence. This method is particularly appreciated on the one hand because it gives efficient and unbiased estimates of the parameters of a generalized linear model and on the other hand because it accounts for intra-individual correlation [58,59]. Generalized estimating equations can deal

with missing data under the assumptions that they are MCAR. When these data are MAR, the estimates might be biased [60].

#### 2.3.4. SEMs and CLPMs, Complementary Approaches to the Mixed Model

Despite its effectiveness, the mixed model is only a first step towards more complex statistical methods allowing the move from a global analysis of the population to a more personalized/individualized analysis of occupational data. Therefore, models such as SEMs or CLPMs can be used as alternatives to mixed models for the analysis of longitudinal data.

SEMs can be viewed as a much more comprehensive regression method, including dependent and independent variables. Where SEMs stand out is in the ability to additionally take into account hypothetical latent constructs, and to examine relationships between observed variables and these concepts [61–65]. As such, they can be seen as a natural combination between factor analysis and regression or path analysis. Although they are particularly powerful, SEMs stay very sensitive to the problem of missing values. In addition to the assumptions of normality and independence that they require, SEMs need a very large sample size to fit. On average, at least 500 individuals and up to >2500 individuals if one of the assumptions is not verified are needed to expect good estimates [65].

CLPMs are used to describe and estimate reciprocal relationships or directional influences between longitudinal variables [66,67]. Cross-analysis is particularly used to describe causal relationships between variables. Often contested because of their low statistical power and their limitations (e.g., the need for a large sample of longitudinal data, the stationarity and synchronicity assumptions or the causal relationship assertion), simpler methods such as multiple regression are often preferred [68]. Because of this, alternative models to SEMs such as the random intercept “cross-lagged” panel models have been developed to overcome their shortcomings [69]. In addition to the temporal stability assumption of the classical method, these consider the trait-type stability invariant of individuals over time.

#### 2.3.5. Trajectory Models

By plotting longitudinal data, we obtain curves otherwise called individual trajectories or developmental trajectories. A developmental trajectory describes and provides information on the evolution of an individual over time. A trajectory is used to describe a latent process that cannot be observed directly but can be explained over time using measured (therefore observed) variables from which its trajectory is inferred [70,71]. These methods that take into account unobserved heterogeneity are called person-centered [72]. Nowadays, several common methods [73,74] are used: generalizations of mixed effects models such as growth curve models [75], growth mixture models [76,77], group-based trajectory models [78–81], latent class analysis [82,83] and latent transition analysis [72].

Most of these methods identify several different trajectories within the population, also called latent classes. For each trajectory, the estimated share of individuals in the population belonging to it is given, which reflects its shape, and each individual has a certain probability of belonging to each trajectory. Individuals are then assigned to the group corresponding to the trajectory for which the probability of belonging is the highest. Once a trajectory is identified, it is possible to explain its shape over time but also in terms of static covariates using a mixed model or equivalent.

Although an individual trajectory can be estimated despite prospective missing values, it is difficult to conclude on the validity of this trajectory, especially when data is NMAR [84]. It is essential to have a sufficient amount of data for an individual, otherwise it is impossible to determine its trajectory [85]. These models are also considerably sensitive to assumptions, and misspecification can lead to biased estimates for trajectories and an overestimation of the number of classes by the model [73].

### 3. Case Report

#### 3.1. Introduction

We aim here to illustrate our suggestions with a complete and illustrated example using the Wittyfit database [5]. Here, we focus on a small part of the approaches described above, namely the contribution of trajectory models to mixed models in a real-world longitudinal occupational health data framework.

Let us imagine that we want to analyze the annual evolution in job satisfaction of workers at different companies between 2018 and 2021, and to observe if the gender and the job position of a worker can affect job satisfaction. A first look at the objective allows us to confirm that we are in the presence of clustered (multiple companies), correlated (individual effect) and longitudinal (follow-up over time) data. If a worker did not express his or her sentiment in a year, he or she is assigned a missing value. Thus, we are faced with the main methodological issues mentioned in this paper.

#### 3.2. Methods

##### 3.2.1. Participants and Exclusion Criteria

Wittyfit software is a web-based platform designed to assess workers' health through a holistic approach of the individual. Volunteers are invited to express their feelings on different health-related outcomes using visual analog scales. With more than 40,000 active users in about nearly 80 companies, whose first registrations began in January 2018, the Wittyfit database offers researchers a substantial behavioral and longitudinal database to study the evolution of workers' occupational health through various indicators such as job satisfaction and stress. Workers not present at the baseline (2018) or with too few data [85] were excluded.

##### 3.2.2. Outcomes

Job satisfaction of workers was assessed using a related visual analog scale, scaled from 0 to 100. Workers could rate their personal feeling of job satisfaction as many times as they wanted. A worker's overall annual job satisfaction score was computed as the average of the notes that he or she filled in over the year.

Socio-demographic characteristics of workers (gender and job position) were filled in by corporates clients of Wittyfit. The job position of a worker is defined according to whether he is an employee or a manager.

##### 3.2.3. Statistics

Statistical analyses were performed using R (version 4.1.1) in the RStudio (version 1.4.1717) platform. The imputation of the data was done using "longitudinal-Data" (with the 'locf' and 'linear.interpol' methods) and "mice" packages. Group-based trajectory modeling was realized with the "latrend" package. We combine the commands 'lcMethodLcmmGBTM' and 'lcMethods' to define the model and the 'latrendBatch' command to fit it, with nonstructured matrix of variance-covariance. Model assumptions (residual independence and normality, and variance homogeneity) were verified a posteriori. Unless specified, we considered a  $p$ -value  $< 0.05$  as statistically significant for analyses.

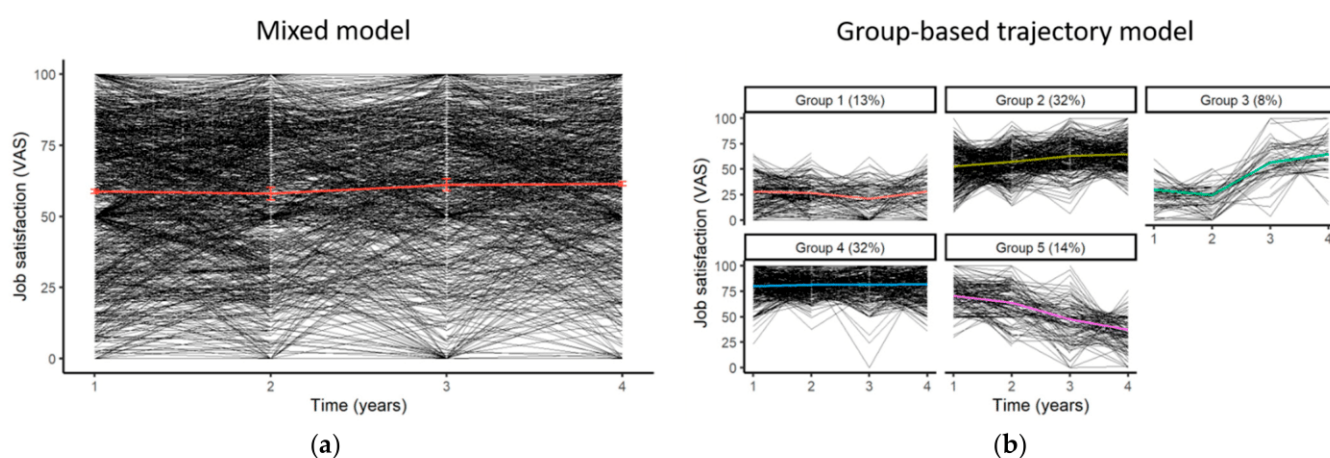
#### 3.3. Data Application

First, we need to analyze the data structure to determine if it has a multilevel structure in addition to a longitudinal structure itself. In our example, we count multiple companies to which workers belong. Therefore, this observation indicates to consider the companies as clusters. We now need to focus on missing data by analyzing its type and structure. The missing data presents two different patterns: (1) a non-monotonic (e.g., due to a worker who did not express in a year) and (2) a non-monotonic pattern (e.g., due to a worker who did stop expressing over the years). For the second one, as data is longitudinal, a drop-out phenomenon may occur. Since it is reasonable to assume that an individual's drop-out may depend on unobserved data (e.g., for an individual who quit during the studied period



because of a lack of job satisfaction), we need to suppose that the drop-out is informative, and therefore that the type of missing values is MNAR. This assumption should therefore lead us to perform a sensitivity analysis by applying several imputation methods, for example with the “last observation carried forward” and the linear regression interpolation methods, but also using multiple imputation, taking into account the company effect.

To study the evolution of workers’ job satisfaction, we can first apply a linear mixed model assessing the influence of time on job satisfaction (Figure 4a). Starting with a single trajectory given by the mixed model, we can observe whether different evolutions exist in the population. To do this, we can apply a trajectory model such as the GBTM on the original dataset and confirm the results using the imputed datasets. The model allows us to identify five latent classes within the population, with a minimal average posterior probability of assignment of 70.5% and minimal odds of correct classification of 5.5, thus exceeding the classical thresholds of 70% and 5 expected for this type of model [79] (Figure 4b).



**Figure 4.** Evolutionary trajectories of job satisfaction among Wittyfit users: (a) mean trajectory using a mixed model (standard errors of coefficient are symbolized by error bars), (b) individual trajectories using a group-based trajectory model, where each group represent a possible evolution of the workers’ job satisfaction in the population (e.g., individuals belonging to “Group 3” are characterized by a slight decrease between times 1 and 2, then by a sharp increase beyond time 2).

Finally, to explore the relationship between a worker’s sociodemographic characteristics and job satisfaction, we can apply a generalized mixed model with the trajectory to which the worker belongs as the outcome and the company effect as a random effect. Thus, according to the model, both gender ( $\beta = 1.60$ , 95% CI 1.15 to 2.24,  $p = 0.005$ ) and job position affect the job satisfaction of the worker ( $\beta = 5.49$ , 95% CI 2.72 to 13.13,  $p < 0.001$ ). This result remains true regardless of the dataset, non-imputed or imputed.

### 3.4. Conclusions

From a self-imposed context, we presented here the many challenges we had to face in analyzing real-world longitudinal occupational health data. This study had two objectives, namely the identification of the existence of different evolutionary trajectories of job satisfaction within a population of workers and the role of the job position on the level of job satisfaction. Each methodological issue has been raised and addressed using the appropriate methods in agreement with the purpose of our article, allowing us to meet the aims of the study.

## 4. Conclusions

As a direct result of the rise of massive databases [5–8] and advances in computational and digital tools and their democratization, real-world data analysis has allowed researchers to confirm the results of previous randomized clinical trials and provide new knowledge, known as real-world evidence [1,3]. The progressive and incessant accumulation

of data thus offers the possibility of collecting health-related, longitudinal, individual data, all at a low cost, thus allowing the development of data-analysis-centered medical approaches.

Although the analysis of longitudinal data has become widespread over the last 30 years, particularly through mixed models, and has made it possible to study population evolution, this article demonstrates that methodological issues remain and have begun to find an echo in more recent approaches, allowing for a move towards a more predictive, preventive, personalized and participatory medicine. Despite the advances offered by these new approaches, some of these issues remain to be addressed before proceeding with the analyses [20,21], in addition to the usual issues linked to any data analysis, limiting their use and overall understanding [74].

In this article, we aimed to present the state of the art of the methods developed for the analysis of real-world longitudinal occupational health data while exposing the three main issues encountered during the analyses of such type of data, namely the concepts of cluster-correlated data, missing data and longitudinal data itself. In addition, we have provided an example of data analysis presenting these issues and discussed the steps to be taken in the analysis process. A search strategy using numerous general article databases was conducted, showing that there is currently no approach to deal with these three methodological issues.

We believe this article can serve as a practical guide for future studies on the topic to improve experimental quality, especially for non-statistician researchers who wish to keep abreast of the new approaches available and the issues involved. Future studies will focus on the comparison between the different models and approaches presented here. Simulation work has already been done [86–89], but not in this specific framework.

**Author Contributions:** Conceptualization, R.C.-C., B.P. and F.D.; methodology, R.C.-C. and B.P.; software, R.C.-C.; validation, B.P. and F.D.; formal analysis, R.C.-C.; investigation, F.D., S.D. and T.C.; resources, S.D. and T.C.; data curation, R.C.-C.; writing—original draft preparation, R.C.-C.; writing—review and editing, B.P., F.D., J.S.B. and S.C.; visualization, R.C.-C.; supervision, B.P. and F.D.; project administration, B.P.; funding acquisition, S.D. and T.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** This study was approved by the National Commission for Data Protection and Liberties (CNIL), and the South-East VI ethics committee (clinicaltrials.gov identification number NCT02596737).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study. This information is set out in Wittyfit’s terms and conditions of use.

**Data Availability Statement:** Data from Wittyfit cannot be transmitted without the prior consent of the company’s corporate clients, except to the University Hospital of Clermont-Ferrand, France, which may use the data for research purposes.

**Acknowledgments:** We express our sincere gratitude to all voluntary workers using Wittyfit, who participated in this study.

**Conflicts of Interest:** R.C., S.D. and T.C. are part of Wittyfit. Other authors have declared that no competing interests exist (F.D. is responsible for the scientific accuracy of Wittyfit, but is not paid by Wittyfit; as previously published, Wittyfit is a public private partnership with the CHU Clermont-Ferrand).

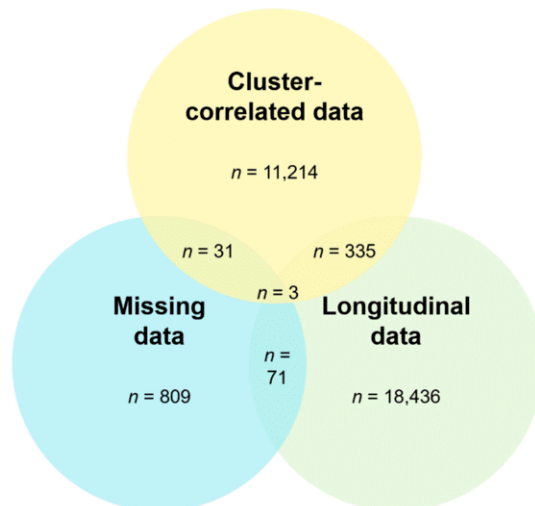
## Appendix A. Novelty and Usefulness of Our Approach and Main Formulas Surrounding Cluster-Correlated Data, Missing Data and Longitudinal Data

Many challenges and methodological issues make it difficult to analyze longitudinal data. These analytical problems include the correlated structure of intra-individual data, the considerable size of data sets, irregular time-spaced measurements, non-linear patterns (such as rapid growth or stationary responses), latent constructs, mix of time-varying and static covariates. There are also more difficult problems to consider, such as longitudinal



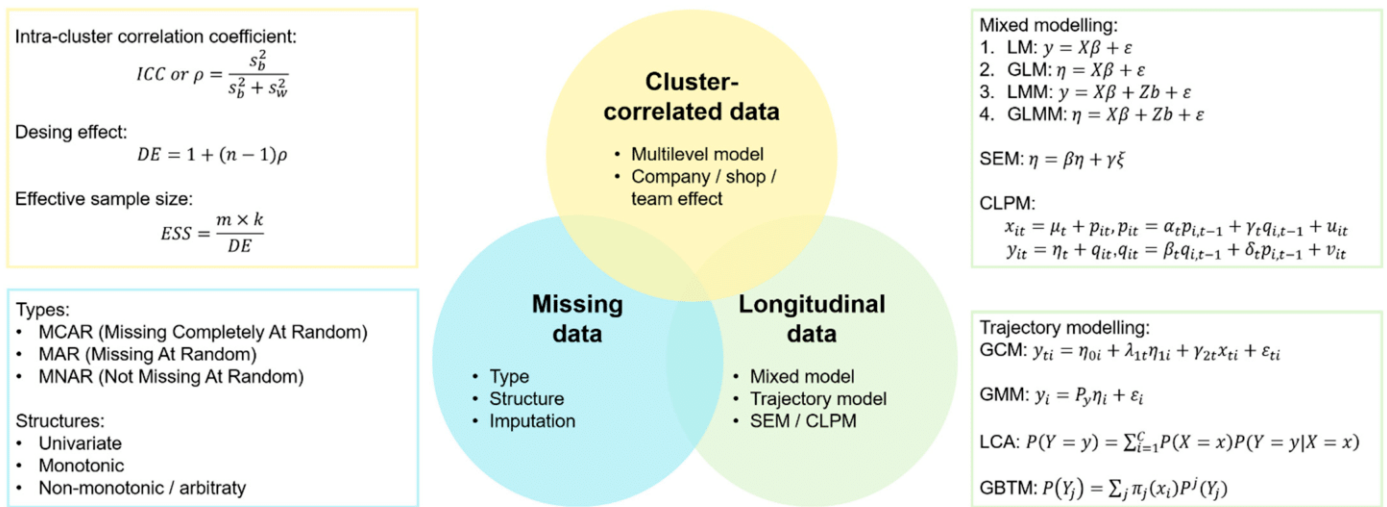
modeling itself, missing data and cluster-correlated data. To sum up, studying such data types requires understanding the specific concepts that are outlined in Figure 1.

The novelty and usefulness of our approach is validated by a step-by-step search strategy. The search presented below has been computed on 8 February 2022 using the PubMed database (details of other research can be found in Appendix B). First, we computed the number of articles dealing with occupational data using keywords (“occupation\*” OR “profession\*” OR “job-related” OR “work-related”) followed by one of the three concepts: correlated-clustered data using keyword (“cluster\*”), missing data using keywords (“missing data” OR “missing value\*”) and longitudinal data using keyword (“longitudinal”). We retrieved  $n = 11,214$  articles using the keywords (“occupation\*” OR “profession\*” OR “job-related” OR “work-related”) AND (“cluster\*”),  $n = 809$  articles using the keywords (“occupation\*” OR “profession\*” OR “job-related” OR “work-related”) AND (“missing data” OR “missing value\*”) and  $n = 19,436$  articles using the keywords (“occupation\*” OR “profession\*” OR “job-related” OR “work-related”) AND (“longitudinal”). Then, we combined the two concepts. We retrieved  $n = 31$  articles using the keywords (“occupation\*” OR “profession\*” OR “job-related” OR “work-related”) AND (“cluster\*”) AND (“missing data” OR “missing value\*”),  $n = 335$  articles using the keywords (“occupation\*” OR “profession\*” OR “job-related” OR “work-related”) AND (“cluster\*”) AND (“longitudinal”) and  $n = 71$  articles using the keywords (“occupation\*” OR “profession\*” OR “job-related” OR “work-related”) AND (“missing data” OR “missing value\*”) AND (“longitudinal”). Lastly, we found only  $n = 3$  articles using the keywords (“occupation\*” OR “profession\*” OR “job-related” OR “work-related”) AND (“cluster\*”) AND (“missing data” OR “missing value\*”) AND (“longitudinal”). Therefore, there is currently no approach to deal with the methodological issues in analyzing real-world longitudinal occupational health data. The results of our research can be summarized as shown in Figure A1.



**Figure A1.** Results of the step-by-step search strategy using PubMed.

To date, the understanding of occupational health data is still under research. As mentioned above, there are three main issues in analyses, namely cluster-correlated data, missing data and longitudinal data itself. Details of the main formulas surrounding cluster-correlated data, missing data and longitudinal data can be found below (Figure A2).



**Figure A2.** Main formulas for the methods presented to address methodological issues for analyzing real-world longitudinal occupational health data. Table A1 provides a summary of the main models and formulas presented here.

**Main Formulas Surrounding Cluster-Correlated Data, Missing Data, and Longitudinal Data.**

*Appendix A.1. Cluster-Correlated Data*

The intra-cluster correlation coefficient (ICC) is a measure of similarity of cluster-correlated data. The ICC can be defined as follows:

$$ICC \text{ or } \rho = \frac{s_b^2}{s_b^2 + s_w^2} \tag{A1}$$

where  $s_b^2$  represents the between-cluster variability and  $s_w^2$  the within-cluster variability. As the ICC increases, the sample size required to detect a significant effect becomes larger. The “design effect” (DE) estimator is then used to estimate the increase in sample size needed to account for the homogeneity of the clustered data:

$$DE = 1 + (n - 1)\rho \tag{A2}$$

where  $n$  is the average size of a cluster. Finally, the “effective sample size” (ESS) can be calculated as bellow:

$$ESS = \frac{m \times k}{DE} \tag{A3}$$

where  $m$  corresponds to the number of subjects in a cluster and  $k$  the total number of clusters.

*Appendix A.2. Longitudinal Data*

*Appendix A.2.1. Linear Modeling*

The linear model describes the relationship between one or more independent variables  $X_i$  ( $i = 1, \dots, n$ ), also called the predictors and a continuous dependent variable  $y$ , also called the response. The equation of a linear model is:

$$y = X\beta + \varepsilon \tag{A4}$$

where  $X$  represents the matrix of the predictors  $X_i$ ,  $\beta$  the vector of fixed effects and  $\varepsilon$  the vector of errors terms following a normal distribution with mean zero and variance  $\sigma^2 I_n$ .

The generalized linear model is a generalization of the linear model, including a linking function specifying the relationship between the predictors and the response [90]. The equation of a generalized linear model then is:

$$\eta = X\beta + \varepsilon \quad (\text{A5})$$

where  $\eta$  represent the link function. In the logistic regression case, the link function will be the *logit* function described as follows:

$$\eta = \text{logit}(p) = \log\left(\frac{p}{1-p}\right) \quad (\text{A6})$$

with  $p = P(y = 1)$ .

The linear mixed model and generalized mixed model appear as extensions of the linear model and generalized mixed model, respectively. Mixed models contain both fixed effects (as describe in the two first models) and random effects [20,53]. The equation of a generalized mixed linear model is:

$$\eta = X\beta + Zb + \varepsilon \quad (\text{A7})$$

where  $b$  is the vector of random effects and  $Z$  the design matrix relating  $b$  to  $y$ .

Although likelihood-based approaches are still the most commonly used to estimate model parameters, there are alternatives such as Bayesian approaches [57], generalized estimating equations [58,59] or Monte Carlo-based methods [91].

#### Appendix A.2.2. Structural Equation Modeling

The structural equation modeling is a set of techniques allowing the estimation of the relationships between dependent and independent variables. Factually, the structural equation models (SEMs) are, as the generalized model, a generalization of the linear model. As such, the equation of a structural equation model seems very close to that of the linear model. It can, be described as follows [61,63]:

$$\eta = \beta\eta + \gamma\bar{\zeta} \quad (\text{A8})$$

where  $\eta$  is the vector of dependent variables,  $\beta$  the matrix of the regression coefficients between dependent variables,  $\gamma$  the matrix of regression coefficients between dependent and independent variables and  $\bar{\zeta}$  the vector of independent variables.

Maximum likelihood, generalized least squares, elliptical distribution theory [92] or the asymptotically distribution free method [93] represent possible estimation methods for this model. The choice of the parameter estimation depends on the choice of the weight matrix [65].

#### Appendix A.2.3. Cross-Lagged Panel Modeling

Cross-lagged panel models (CLPMs) can be useful to measure the evolution of several variables and evaluate their mutual influences over time. Considering two variables  $x$  and  $y$  measured at multiple times (i.e.,  $x = (x_1, \dots, x_n)$  and  $y = (y_1, \dots, y_n)$ ), the measurement equations between the two variables for an individual  $i$  are:

$$x_{it} = \mu_t + p_{it}y_{it} = \eta_t + q_{it} \quad (\text{A9})$$

where  $\mu_t$  and  $\eta_t$  are the means of the two variables at time  $t$ , respectively and where the temporal deviations are defined as:

$$p_{it} = \alpha_t p_{i,t-1} + \gamma_t q_{i,t-1} + u_{it} q_{it} = \beta_t q_{i,t-1} + \delta_t p_{i,t-1} + v_{it} \quad (\text{A10})$$

where  $\alpha_t$  and  $\beta_t$  represent the auto-regressive effects, i.e., the effect of a variable on itself, and  $\gamma_t$  and  $\delta_t$  the cross-lagged effects, i.e., the effect on a variable on the other, the whole over time.

As for the SEMs, the estimation of the parameters can be computed using the maximum likelihood method [94]. Nevertheless, Bayesian estimation can also represent an alternative to the first method [95].

### Appendix A.3. Trajectory Modeling

#### Appendix A.3.1. Growth Curve Modeling

Trajectory models consist of statistical methods used to analyze the change over time among a population of individuals. The growth curve model (GCM) allows in particular analyzing both inter- (also called the between-person) and intra-individual (the within-person) changes thanks to the mean and the covariance structure, respectively. As well, considering a response vector  $y$  measured  $t$  times, the model can be written as [75]:

$$y_{ti} = \eta_{0i} + \lambda_{1t}\eta_{1i} + \gamma_{2t}x_{ti} + \varepsilon_{ti} \quad (\text{A11})$$

where  $\lambda_{1t}$  is the measurement of the variable at the first time, the individual intercept  $\eta_{0i}$  with mean  $v_0$  and random departure  $\zeta_{0i}$  is defined by:

$$\eta_{0i} = v_0 + \gamma_0\zeta_i + \xi_{0i} \quad (\text{A12})$$

and the individual slope with mean  $v_1$  and random departure  $\zeta_{1i}$  by:

$$\eta_{1i} = v_1 + \gamma_1\zeta_i + \xi_{1i} \quad (\text{A13})$$

Finally,  $\gamma_0$  and  $\gamma_1$  represent the effects of the time-invariant covariate  $\zeta_i$  and  $\gamma_{2t}$  of the time-varying covariate  $x_{ti}$  and  $\varepsilon_{ti}$  the time-specific deviation, following a normal distribution with mean zero and variance  $\sigma_\varepsilon^2$ .

#### Appendix A.3.2. Growth Mixture Modeling

The growth mixture model (GMM) is quite similar to the GCM, with the exception that it combines the techniques of the growth models with the latent concepts defined in the latent class analysis. Considering an observed response  $y$  and a set of latent continuous variables  $\eta$ , the model can be written as [76]:

$$y_i = P_y\eta_i + \varepsilon_i \quad (\text{A14})$$

where  $P_y$  describes the matrix of the parameters and  $\varepsilon$  the vector of residuals with mean zero and covariance  $D$ . Continuous latent variables  $\eta$  are for them related to the observed covariate  $x$  and the latent categorical variables  $c$  by the relation:

$$\eta_i = Lc_i + Q_\eta x_i + \xi_i \quad (\text{A15})$$

where  $L$  describes the matrix of the intercept parameters of latent classes  $c$ ,  $Q_\eta$  describes the matrix of the parameters and  $\xi$  the vector of residuals with mean zero and covariance  $D$ .

#### Appendix A.3.3. Latent Class Analysis

Latent class analysis assumes the existence of several underlying latent classes in a population. Although the actual class membership of an individual is unknown, it can be inferred through the covariates. The basic equation of the LCA is [82]:

$$P(Y = y) = \sum_{i=1}^C P(X = x)P(Y = y|X = x) \quad (\text{A16})$$



where  $P(X = x)$  is the probability of an individual belonging to latent class  $c$  among the  $C$  existing latent classes.

Appendix A.3.4. Group-Based Trajectory Modeling

The estimation of the GBTM parameters is quite similar to the estimation of the LCA parameters. Indeed, the equation of the GBTM is:

$$P(Y_j) = \sum_j \pi_j(x_i)P^j(Y_j) \tag{A17}$$

where  $P^j(Y_j)$  is the probability of  $Y_j$  (i.e., the trajectory of the individual  $i$ ) if the subject belongs to the group  $j$  and  $\pi_j(x_i)$  is the membership's probability of the group  $j$ , determined from the covariate  $x_i$ .

Appendix A.4. Summary of Models Formula

Summary of the main formulas surrounding models for cluster-correlated data, missing data and longitudinal data can be found below (Table A1).

Table A1. Summary of the different models.

Model	Search Strategy	Mathematical Formulation	Missing Data	Advantages	Drawbacks
MULTILEVEL MODELING					
MLM	3585	$y_{ij} = \beta_0 + \beta_1 x_{ij} + (u_{0j} + u_{1j} x_{ij} + e_{0ij})$	<i>a</i>	- Calculation of the intra-cluster variability apart from the overall variability	- Validity of conclusions highly dependent on the cluster effect specification and definition and interpretation or the model parameters
FIRST APPROACHES					
ANOVA for repeated measures	1547	$SS_{error} = SS_{total} - SS_{between} - SS_{subject}$ (SS: Sum of Squares)	<i>a</i>	- Comparison of group's means across time	- No information on possible individual trajectories - Time effect as fixed effect
Mixed model	1983	$y = X\beta + \varepsilon$ $\eta = X\beta + \varepsilon$ $\eta = X\beta + Zb + \varepsilon$	<i>b</i> *	- Combination of fixed and random effects - Randomly missing estimates unbiased - Possible consideration of subject-specific/cluster effect	- Only provides information on the average trajectory followed by the population
TRAJECTORIES					
GCM	126	$y_{ti} = \eta_{0i} + \lambda_{1i}\eta_{1i} + \gamma_{2i}x_{ti} + \varepsilon_{ti}$		- Person-centered method - Identification of different trajectories within population	- Questioning of the validity of individual trajectories caused by data missingness - Need for sufficient amount of data for each individual
GMM	50	$y_i = P_y \eta_i + \varepsilon_i$		- Assigning individuals to a trajectory	
LCA	154	$P(Y = y) = \sum_{i=1}^C P(X = x)P(Y = y X = x)$	<i>b</i>	- Possibility to study each trajectory shape with static covariates	
GBTM	78	$P(Y_j) = \sum_j \pi_j(x_i)P^j(Y_j)$		- Possible consideration of cluster effect	
COMPLEMENTARY APPROACHES					
SEM	380	$\eta = \beta\eta + \gamma\xi$	<i>a</i>	- Description of latent process	- Need for a large sample size to fit

Table A1. Cont.

Model	Search Strategy	Mathematical Formulation	Missing Data	Advantages	Drawbacks
CLPM	105	$x_{it} = \mu_t + p_{it}$ $y_{it} = \eta_t + q_{it}$	<i>b</i>	- Study of causal relationships	- Need for a large sample size to fit - Strong assumptions (stationarity and synchronicity)

Legend: Missing data: *a*: complete-cases analysis, *b*: possible estimates computation despite missing data, \*: unbiased estimates. Number of articles (search strategy in PubMed): The keywords (“occupation\*” OR “profession\*” OR “job-related” OR “work-related”) were linked with the following keywords: MLM: (multilevel OR multi-level). Repeated measures ANOVA: (ANOVA). Mixed model: (mixed). GCM: (GCM OR growth curve). GMM: (GMM OR growth mixture). LCA: (LCA OR LCGA OR latent class). GBTM: (GBTM or group-based trajectory\*). SEM: (SEM or structural equation). CLPM: (CLPM or cross-lagged).

## Appendix B. Details for the Search Strategy Used within Each Database

Results of different research conducted on the Cochrane Library/Embase/PsycInfo/PubMed databases:

### Cochrane Library

#1 occupational OR professional OR job-related OR work-related

#2 cluster

#3 missing data OR missing values

#4 longitudinal

#5 #1 AND #2 AND #3 AND #4

Filter Language = none in CENTRAL

Filter Dates = none

Results = 0

### Embase

‘occupational’ OR ‘professional’ OR ‘job-related’ OR ‘work-related’ AND (‘cluster\*’) AND (‘missing data’ OR ‘missing value\*’) AND ‘longitudinal’

Filter Language = none

Filter Dates = none

Results = 2

### PsycInfo

**Any Field:** occupational OR **Any Field:** professional OR **Any Field:** job-related OR **Any Field:** work-related AND **Any Field:** cluster AND **Any Field:** missing data OR **Any Field:** missing values AND **Any Field:** longitudinal

Filter Language = none

Filter Dates = none

Results = 2

### PubMed

(“occupation\*” OR “profession\*” OR “job-related” OR “work-related”) AND “cluster\*” AND (“missing data” OR “missing value\*”) AND “longitudinal”

Filter Language = none

Filter Dates = none

Results = 3

## References

1. Basch, E.; Schrag, D. The Evolving Uses of “Real-World” Data. *JAMA* **2019**, *321*, 1359–1360. [[CrossRef](#)] [[PubMed](#)]
2. Makady, A.; de Boer, A.; Hillege, H.; Klungel, O.; Goettsch, W. What Is Real-World Data? A Review of Definitions Based on Literature and Stakeholder Interviews. *Value Health* **2017**, *20*, 858–865. [[CrossRef](#)] [[PubMed](#)]
3. Corrigan-Curay, J.; Sacks, L.; Woodcock, J. Real-world evidence and real-world data for evaluating drug safety and effectiveness. *JAMA* **2018**, *320*, 867–868. [[CrossRef](#)]
4. McCormick, J.; Mehta, G.; Olesen, H.V.; Viviani, L.; Macek, M.; Mehta, A. Comparative demographics of the European cystic fibrosis population: A cross-sectional database analysis. *Lancet* **2010**, *375*, 1007–1013. [[CrossRef](#)]

5. Dutheil, F.; Duclos, M.; Naughton, G.; Dewavrin, S.; Cornet, T.; Huguet, P.; Chatard, J.-C.; Pereira, B. Wittyfit-live your work differently: Study protocol for a workplace-delivered health promotion. *JMIR Res. Protoc.* **2017**, *6*, e6267. [[CrossRef](#)] [[PubMed](#)]
6. Platt, R.; Brown, J.S.; Robb, M.; McClellan, M.; Ball, R.; Nguyen, M.D.; Sherman, R.E. The FDA Sentinel Initiative—An Evolving National Resource. *N. Engl. J. Med.* **2018**, *379*, 2091–2093. [[CrossRef](#)]
7. Smith, C.A.; Wicks, P.J. PatientsLikeMe: Consumer Health Vocabulary as a Folksonomy. *AMIA Annu. Symp. Proc.* **2008**, *2008*, 682–686.
8. Randhawa, G.S. Building electronic data infrastructure for comparative effectiveness research: Accomplishments, lessons learned and future steps. *J. Comp. Eff. Res.* **2014**, *3*, 567–572. [[CrossRef](#)]
9. James, G.; Nyman, E.; Fitz-Randolph, M.; Niklasson, A.; Hedman, K.; Hedberg, J.; Wittbrodt, E.T.; Medin, J.; Moreno Quinn, C.; Allum, A.M.; et al. Characteristics, symptom severity, and experiences of patients reporting chronic kidney disease in the patientslikeme online health community: Retrospective and qualitative study. *J. Med. Internet Res.* **2020**, *22*, e18548. [[CrossRef](#)]
10. Benjdir, M.; Audureau, É.; Beresniak, A.; Coll, P.; Epaud, R.; Fiedler, K.; Jacquemin, B.; Niddam, L.; Pandis, S.N.; Pohlmann, G.; et al. Assessing the impact of exposome on the course of chronic obstructive pulmonary disease and cystic fibrosis: The REMEDIA European Project Approach. *Environ. Epidemiol.* **2021**, *5*, e165. [[CrossRef](#)]
11. McCaffrey, S.; Black, R.A.; Nagao, M.; Sepassi, M.; Sharma, G.; Thornton, S.; Kim, Y.H.; Braverman, J. Measurement of quality of life in patients with mycosis fungoides/sézary syndrome cutaneous t-cell lymphoma: Development of an electronic instrument. *J. Med. Internet Res.* **2019**, *21*, e11302. [[CrossRef](#)] [[PubMed](#)]
12. Maissenhaelter, B.E.; Woolmore, A.L.; Schlag, P.M. Real-world evidence research based on big data. *Onkologie* **2018**, *24* (Suppl. S2), 91–98. [[CrossRef](#)]
13. Garrison, L.P.; Neumann, P.J.; Erickson, P.; Marshall, D.; Mullins, C.D. Using Real-World Data for Coverage and Payment Decisions: The ISPOR Real-World Data Task Force Report. *Value Health* **2007**, *10*, 326–335. [[CrossRef](#)] [[PubMed](#)]
14. Barrett, J.S.; Heaton, P.M. Real-World Data: An Unrealized Opportunity in Global Health? *Clin. Pharmacol. Ther.* **2019**, *106*, 57–59. [[CrossRef](#)] [[PubMed](#)]
15. Han, J.; Pei, J.; Kamber, M. *Data Mining: Concepts and Techniques*, 3rd ed.; Elsevier: Amsterdam, The Netherlands, 2011; 740p.
16. Diggle, P.; Heagerty, P.; Liang, K.-Y.; Zeger, S. *Analysis of Longitudinal Data*, 2nd ed.; OUP: Oxford, UK, 2002; 396p.
17. Fitzmaurice, G.M.; Laird, N.M.; Ware, J.H. *Applied Longitudinal Analysis*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2012; 742p.
18. Caruana, E.J.; Roman, M.; Hernández-Sánchez, J.; Solli, P. Longitudinal studies. *J. Thorac. Dis.* **2015**, *7*, E537–E540. [[PubMed](#)]
19. Van Belle, G.; Fisher, L.D.; Heagerty, P.J.; Lumley, T. *Biostatistics: A Methodology for the Health Sciences*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2004; 895p.
20. Edwards, L.J. Modern statistical techniques for the analysis of longitudinal data in biomedical research. *Pediatr. Pulmonol.* **2000**, *30*, 330–344. [[CrossRef](#)]
21. Weiss, R.E. *Modeling Longitudinal Data*; Springer Science & Business Media: New York, NY, USA, 2005; 445p.
22. Killip, S.; Mahfoud, Z.; Pearce, K. What Is an Intraclass Correlation Coefficient? Crucial Concepts for Primary Care Researchers. *Ann. Fam. Med.* **2004**, *2*, 204–208. [[CrossRef](#)] [[PubMed](#)]
23. Song, P.X.-K. *Correlated Data Analysis: Modeling, Analytics, and Applications*; Springer Science & Business Media: New York, NY, USA, 2007; 356p.
24. Goldstein, H. *Multilevel Statistical Models*, 4th ed.; John Wiley & Sons: Hoboken, NJ, USA, 2011; 376p.
25. Bliese, P.D.; Hanges, P.J. Being Both Too Liberal and Too Conservative: The Perils of Treating Grouped Data as though They Were Independent. *Organ. Res. Methods* **2004**, *7*, 400–417. [[CrossRef](#)]
26. Hayes, A.F. A Primer on Multilevel Modeling. *Hum. Commun. Res.* **2006**, *32*, 385–410. [[CrossRef](#)]
27. Gibbons, R.D.; Hedeker, D.; DuToit, S. Advances in analysis of longitudinal data. *Annu. Rev. Clin. Psychol.* **2010**, *6*, 79–107. [[CrossRef](#)]
28. Murray, D.M. *Design and Analysis of Group-Randomized Trials*; Oxford University Press: Oxford, UK, 1998; 481p.
29. Snijders, T.A.B.; Bosker, R.J. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*, 2nd ed.; SAGE: Thousand Oaks, CA, USA, 2011; 370p.
30. Begg, M.D.; Parides, M.K. Separation of individual-level and cluster-level covariate effects in regression analysis of correlated data. *Stat. Med.* **2003**, *22*, 2591–2602. [[CrossRef](#)] [[PubMed](#)]
31. Bruckers, L.; Molenberghs, G.; Pulinx, B.; Hellenthal, F.; Schurink, G. Cluster analysis for repeated data with dropout: Sensitivity analysis using a distal event. *J. Biopharm. Stat.* **2018**, *28*, 983–1004. [[CrossRef](#)] [[PubMed](#)]
32. Hox, J.J.; Moerbeek, M.; van de Schoot, R. *Multilevel Analysis: Techniques and Applications*, 3rd ed.; Routledge: Abingdon, UK, 2017; 365p.
33. Raudenbush, S.W.; Bryk, A.S. *Hierarchical Linear Models: Applications and Data Analysis Methods*, 2nd ed.; SAGE: Thousand Oaks, CA, USA, 2002; 520p.
34. Graham, J.W. Missing Data Analysis: Making It Work in the Real World. *Annu. Rev. Psychol.* **2008**, *60*, 549–576. [[CrossRef](#)] [[PubMed](#)]
35. Little, T.D.; Lang, K.M.; Wu, W.; Rhemtulla, M. Missing Data. In *Developmental Psychopathology*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2016; pp. 1–37.
36. Hedeker, D.; Gibbons, R.D. Application of random-effects pattern-mixture models for missing data in longitudinal studies. *Psychol. Methods* **1997**, *2*, 64–78. [[CrossRef](#)]
37. Kang, H. The prevention and handling of the missing data. *Korean J. Anesthesiol.* **2013**, *64*, 402–406. [[CrossRef](#)]



38. Donner, A. The Relative Effectiveness of Procedures Commonly Used in Multiple Regression Analysis for Dealing with Missing Values. *Am. Stat.* **1982**, *36*, 378–381.
39. Newgard, C.D.; Lewis, R.J. Missing Data: How to Best Account for What Is Not Known. *JAMA* **2015**, *314*, 940–941. [[CrossRef](#)]
40. Li, P.; Stuart, E.A.; Allison, D.B. Multiple Imputation: A Flexible Tool for Handling Missing Data. *JAMA* **2015**, *314*, 1966–1967. [[CrossRef](#)]
41. Little, R.J.A.; Rubin, D.B. *Statistical Analysis with Missing Data*, 2nd ed.; John Wiley & Sons: New York, NY, USA, 2002.
42. Allison, P.D. *Missing Data; Quantitative Applications in the Social Sciences*; SAGE Publications: Thousand Oaks, CA, USA, 2001; Volume 136.
43. Rubin, D.B. Inference and missing data. *Biometrika* **1976**, *63*, 581–592. [[CrossRef](#)]
44. Kenward, M.G.; Carpenter, J. Multiple imputation: Current perspectives. *Stat. Methods Med. Res.* **2007**, *16*, 199–218. [[CrossRef](#)]
45. Diggle, P.; Kenward, M.G. Informative Drop-Out in Longitudinal Data Analysis. *J. R. Stat. Soc. Ser. C (Appl. Stat.)* **1994**, *43*, 49–73. [[CrossRef](#)]
46. Little, R.J.A. Modeling the Drop-Out Mechanism in Repeated-Measures Studies. *J. Am. Stat. Assoc.* **1995**, *90*, 1112–1121. [[CrossRef](#)]
47. Twisk, J.; de Vente, W. Attrition in longitudinal studies: How to deal with missing data. *J. Clin. Epidemiol.* **2002**, *55*, 329–337. [[CrossRef](#)]
48. Fitzmaurice, G. Missing data: Implications for analysis. *Nutrition* **2008**, *24*, 200–202. [[CrossRef](#)] [[PubMed](#)]
49. Rosenthal, S. Data Imputation. In *The International Encyclopedia of Communication Research Methods*; American Cancer Society: Atlanta, GA, USA, 2017; pp. 1–12.
50. Liu, C.; Cripe, T.P.; Kim, M.-O. Statistical Issues in Longitudinal Data Analysis for Treatment Efficacy Studies in the Biomedical Sciences. *Mol. Ther.* **2010**, *18*, 1724–1730. [[CrossRef](#)] [[PubMed](#)]
51. Verbeke, G. Linear Mixed Models for Longitudinal Data. In *Linear Mixed Models in Practice: A SAS-Oriented Approach*; Lecture Notes in Statistics; Verbeke, G., Molenberghs, G., Eds.; Springer: New York, NY, USA, 1997; pp. 63–153.
52. Verbeke, G.; Molenberghs, G. *Linear Mixed Models for Longitudinal Data*; Springer: New York, NY, USA, 2000.
53. Laird, N.M.; Ware, J.H. Random-effects models for longitudinal data. *Biometrics* **1982**, *38*, 963–974. [[CrossRef](#)] [[PubMed](#)]
54. Fahrmeir, L.; Tutz, G. *Multivariate Statistical Modelling Based on Generalized Linear Models*, 2nd ed.; Springer Science & Business Media: New York, NY, USA, 1994; 537p.
55. Cnaan, A.; Laird, N.M.; Slasor, P. Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. *Stat. Med.* **1997**, *16*, 2349–2380. [[CrossRef](#)]
56. McCulloch, C.E.; Neuhaus, J.M. Generalized Linear Mixed Models. In *Encyclopedia of Biostatistics*; American Cancer Society: Atlanta, GA, USA, 2005.
57. Ju, K.; Lin, L.; Chu, H.; Cheng, L.-L.; Xu, C. Laplace approximation, penalized quasi-likelihood, and adaptive Gauss–Hermite quadrature for generalized linear mixed models: Towards meta-analysis of binary outcome with sparse data. *BMC Med. Res. Methodol.* **2020**, *20*, 152. [[CrossRef](#)]
58. Liang, K.-Y.; Zeger, S.L. Longitudinal data analysis using generalized linear models. *Biometrika* **1986**, *73*, 13–22. [[CrossRef](#)]
59. Ballinger, G.A. Using generalized estimating equations for longitudinal data analysis. *Organ. Res. Methods* **2004**, *7*, 127–150. [[CrossRef](#)]
60. Zorn, C.J.W. Generalized estimating equation models for correlated data: A review with applications. *Am. J. Political Sci.* **2001**, *45*, 470–490. [[CrossRef](#)]
61. Bentler, P.M.; Weeks, D.G. Linear structural equations with latent variables. *Psychometrika* **1980**, *45*, 289–308. [[CrossRef](#)]
62. Hoyle, R.H. *Structural Equation Modeling: Concepts, Issues, and Applications*; SAGE: Thousand Oaks, CA, USA, 1995; 313p.
63. Ullman, J.B. Structural equation modeling: Reviewing the basics and moving forward. *J. Pers. Assess* **2006**, *87*, 35–50. [[CrossRef](#)] [[PubMed](#)]
64. Savalei, V.; Bentler, P.M. Structural Equation Modeling. In *The Corsini Encyclopedia of Psychology*; American Cancer Society: Atlanta, GA, USA, 2010; pp. 1–3.
65. Ullman, J.B.; Bentler, P.M. Structural Equation Modeling. In *Handbook of Psychology*, 2nd ed; American Cancer Society: Atlanta, GA, USA, 2012.
66. Kenny, D.A. Cross-lagged panel correlation: A test for spuriousness. *Psychol. Bull.* **1975**, *82*, 887–903. [[CrossRef](#)]
67. Selig, J.P.; Little, T.D. Autoregressive and cross-lagged panel analysis for longitudinal data. In *Handbook of Developmental Research Methods*; The Guilford Press: New York, NY, USA, 2012; pp. 265–278.
68. Kenny, D.A.; Harackiewicz, J.M. Cross-lagged panel correlation: Practice and promise. *J. Appl. Psychol.* **1979**, *64*, 372–379. [[CrossRef](#)]
69. Hamaker, E.L.; Kuiper, R.M.; Grasman, R.P.P.P. A critique of the cross-lagged panel model. *Psychol. Methods* **2015**, *20*, 102–116. [[CrossRef](#)]
70. Curran, P.J.; Willoughby, M.T. Implications of latent trajectory models for the study of developmental psychopathology. *Dev. Psychopathol.* **2003**, *15*, 581–612. [[CrossRef](#)]
71. Schumacker, R.; Lomax, R. *A Beginner's Guide to Structural Equation Modeling*, 4th ed.; Routledge: Mahwah, NJ, USA, 2016; Volume 288.
72. Muthén, B.; Muthén, L.K. Integrating Person-Centered and Variable-Centered Analyses: Growth Mixture Modeling with Latent Trajectory Classes. *Alcohol. Clin. Exp. Res.* **2000**, *24*, 882–891. [[CrossRef](#)]



73. Herle, M.; Micali, N.; Abdulkadir, M.; Loos, R.; Bryant-Waugh, R.; Hübel, C.; Bulik, C.M.; De Stavola, B.L. Identifying typical trajectories in longitudinal data: Modelling strategies and interpretations. *Eur. J. Epidemiol.* **2020**, *35*, 205–222. [[CrossRef](#)]
74. Nguena Nguéfack, H.L.; Pagé, M.G.; Katz, J.; Choinière, M.; Vanasse, A.; Dorais, M.; Samb, O.M.; Lacasse, A. Trajectory Modelling Techniques Useful to Epidemiological Research: A Comparative Narrative Review of Approaches. *Clin. Epidemiol.* **2020**, *12*, 1205–1222. [[CrossRef](#)]
75. Hox, J.; Stoel, R.D. Multilevel and SEM approaches to growth curve modeling. In *Encyclopedia of Statistics in Behavioral Science*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2005.
76. Muthén, B.; Shedden, K. Finite Mixture Modeling with Mixture Outcomes Using the EM Algorithm. *Biometrics* **1999**, *55*, 463–469. [[CrossRef](#)] [[PubMed](#)]
77. Muthén, B. Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class–latent growth modeling. In *New Methods for the Analysis of Change*; American Psychological Association: Washington, DC, USA, 2001; pp. 291–322.
78. Nagin, D.S. Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychol. Methods* **1999**, *4*, 139–157. [[CrossRef](#)]
79. Nagin, D.S. *Group-Based Modeling of Development*; Harvard University Press: Cambridge, MA, USA, 2005; 226p.
80. Nagin, D.S.; Odgers, C.L. Group-based trajectory modeling in clinical research. *Annu. Rev. Clin. Psychol.* **2010**, *6*, 109–138. [[CrossRef](#)] [[PubMed](#)]
81. Nagin, D.S.; Jones, B.L.; Passos, V.L.; Tremblay, R.E. Group-based multi-trajectory modeling. *Stat. Methods Med. Res.* **2018**, *27*, 2015–2023. [[CrossRef](#)] [[PubMed](#)]
82. Lanza, S.T.; Rhoades, B.L. Latent Class Analysis: An Alternative Perspective on Subgroup Analysis in Prevention and Treatment. *Prev. Sci.* **2013**, *14*, 157–168. [[CrossRef](#)]
83. Lanza, S.T.; Cooper, B.R. Latent Class Analysis for Developmental Research. *Child Dev. Perspect.* **2016**, *10*, 59–64. [[CrossRef](#)]
84. Dupéré, V.; Lacourse, E.; Vitaro, F.; Tremblay, R.E. Méthodes d’analyse du changement fondées sur les trajectoires de développement individuel. Modèles de régression mixtes paramétriques et non paramétriques. *Bull. Méthodol. Sociol. Bull. Sociol. Methodol.* **2007**, *95*, 26–57. [[CrossRef](#)]
85. Rogosa, D.; Brandt, D.; Zimowski, M. A growth curve approach to the measurement of change. *Psychol. Bull.* **1982**, *92*, 726–748. [[CrossRef](#)]
86. Martin, D.P.; von Oertzen, T. Growth mixture models outperform simpler clustering algorithms when detecting longitudinal heterogeneity, even with small sample sizes. *Struct. Equ. Model. A Multidiscip. J.* **2015**, *22*, 264–275. [[CrossRef](#)]
87. McNeish, D.; Harring, J.R. The effect of model misspecification on growth mixture model class enumeration. *J. Classif.* **2017**, *34*, 223–248. [[CrossRef](#)]
88. McNeish, D.; Matta, T. Differentiating between mixed-effects and latent-curve approaches to growth modeling. *Behav. Res.* **2018**, *50*, 1398–1414. [[CrossRef](#)] [[PubMed](#)]
89. Den Teuling, N.G.P.; Pauws, S.C.; van den Heuvel, E.R. A comparison of methods for clustering longitudinal data with slowly changing trends. *Commun. Stat.-Simul. Comput.* **2021**, *20*, 1–28. [[CrossRef](#)]
90. Nelder, J.A.; Wedderburn, R.W.M. Generalized Linear Models. *J. R. Stat. Soc. Ser. A (Gen.)* **1972**, *135*, 370–384. [[CrossRef](#)]
91. Booth, J.G.; Hobert, J.P. Maximizing generalized linear mixed model likelihoods with an automated Monte Carlo EM algorithm. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **1999**, *61*, 265–285. [[CrossRef](#)]
92. Shapiro, A.; Browne, M.W. Analysis of covariance structures under elliptical distributions. *J. Am. Stat. Assoc.* **1987**, *82*, 1092–1097. [[CrossRef](#)]
93. Browne, M.W. Asymptotically distribution-free methods for the analysis of covariance structures. *Br. J. Math. Stat. Psychol.* **1984**, *37*, 62–83. [[CrossRef](#)] [[PubMed](#)]
94. Allison, P.D.; Williams, R.; Moral-Benito, E. Maximum likelihood for cross-lagged panel models with fixed effects. *Socius* **2017**, *3*, 1–17. [[CrossRef](#)]
95. Zyphur, M.J.; Hamaker, E.L.; Tay, L.; Voelkle, M.; Preacher, K.J.; Zhang, Z.; Allison, P.D.; Pierides, D.C.; Koval, P.; Diener, E.F. From data to causes III: Bayesian priors for general cross-lagged panel models (GCLM). *Front. Psychol.* **2021**, *12*, 612251. [[CrossRef](#)]

## Article n°2

The protective role of job  
control/autonomy on mental strain  
of managers: a cross-sectional study  
among Wittyfit's users

## **Article n°2**

The protective role of job control/autonomy on mental strain of managers: a cross-sectional study among Wittyfit's users

Colin-Chevalier R, Pereira B, Benson AC, Dewavrin S, Cornet T, Dutheil F

### **Résumé**

Cet article constitue le premier d'une série de trois articles autour d'un facteur peu pris en compte lors de l'évaluation des RPS sur le lieu de travail : le statut managérial. Dans cet article, nous souhaitons comparer les situations de travail des managers et des employés en utilisant le modèle de Karasek. En amont de nos analyses, nous avons identifié deux enjeux méthodologiques auxquels nous allons faire face. D'une part, l'inclusion de travailleurs issus de plusieurs entreprises nous a conduit à appliquer des modèles mixtes en prenant en compte l'« effet entreprise » (l'entreprise à laquelle appartient un travailleur) comme un effet de cluster. D'autre part, confronté à un problème de données manquantes identifiées comme ayant une nature peu complexe, nous avons appliqué une méthode d'imputation par la moyenne. Etant donnée la présence de clusters de données, cette imputation s'est effectuée au sein de chaque cluster. Une fois nos enjeux résolus, nous étions en mesure de procéder à nos analyses. Dans un premier temps, nous avons étudié les différences dans la perception que se font les deux populations en termes de charge de travail, d'autonomie et de soutien social. En moyenne, si les deux populations de travailleurs ressentaient des niveaux de charge de travail équivalents, les managers percevaient de plus hauts niveaux d'autonomie et de soutien social que les employés. Dans un second temps, nous avons mesuré et comparé l'exposition au job strain et à l'isostrain des managers et des employés. Nous avons pu mettre en évidence que les managers étaient moins exposés que les employés, notamment du fait d'une meilleure autonomie et d'un meilleur soutien social au travail. D'autres facteurs, notamment l'âge, l'ancienneté et le fait d'être une femme, ont été associées à une charge de travail plus élevée ainsi qu'à un niveau

d'autonomie et de soutien social moindre. Ces trois facteurs ont alors été identifiés comme des facteurs de risque de contraintes au travail et de contraintes d'isolement. En implémentant des stratégies visant à améliorer ces deux dimensions du travail, notamment pour les employés qui sont plus exposés au job strain et à l'isostrain, les entreprises pourraient être capable d'améliorer la QVT des travailleurs et de les protéger de ces risques susceptibles d'avoir des effets délétères, voir mortels.



Article

# The Protective Role of Job Control/Autonomy on Mental Strain of Managers: A Cross-Sectional Study among Wittyfit's Users

Rémi Colin-Chevalier <sup>1,\*</sup> , Bruno Pereira <sup>2</sup> , Amanda Clare Benson <sup>3</sup> , Samuel Dewavrin <sup>4</sup>, Thomas Cornet <sup>4</sup> and Frédéric Dutheil <sup>1</sup>

<sup>1</sup> CNRS, LaPSCo, Physiological and Psychosocial Stress, University Hospital of Clermont-Ferrand, Preventive and Occupational Medicine, Université Clermont Auvergne, Wittyfit, 63000 Clermont-Ferrand, France; fred\_dutheil@yahoo.fr

<sup>2</sup> Biostatistics Unit, Clinical Research and Innovation Direction, University Hospital of Clermont-Ferrand, 63000 Clermont-Ferrand, France; bpereira@chu-clermontferrand.fr

<sup>3</sup> Department of Health and Biostatistics, Swinburne University of Technology, Hawthorn, VIC 3122, Australia; abenson@swin.edu.au

<sup>4</sup> Wittyfit, 75000 Paris, France; samuel.dewavrin@wittyfit.com (S.D.); thomas.cornet@wittyfit.com (T.C.)

\* Correspondence: r.colin6374@gmail.com

**Abstract:** Background: Karasek's Job Demand-Control-Support model is the gold standard to assess the perception of work; however, this model has been poorly studied among managers. We aimed to explore the perception of work (job demand, control, and support) in managers, and to quantify their risk of job strain (high job demand and low job control) and isostrain (job strain with low job support). Methods: We conducted a cross-sectional study on workers from various French companies using the Wittyfit software. Job demand, control, and support were evaluated by self-reported questionnaires, as well as sociodemographic data. Results: We included 9257 workers: 8488 employees (median age of 45 years, median seniority of 10 years, 39.4% women) and 769 managers (463 were more than 45 years old, 343 with more than 10 years of service, 33.3% women). Managers had higher mean  $\pm$  SD levels than employees in job control ( $79.2 \pm 14.9$  vs.  $75.4 \pm 16.9$ ) and job support ( $25.2 \pm 5.1$  vs.  $24.0 \pm 6.1$ ) ( $p < 0.001$ ). Compared to employees, managers had a 37% decreased risk of job strain (OR = 0.63, 95% CI 0.52 to 0.77) and a 47% decreased risk of isostrain (OR = 0.53, 95% CI 0.40 to 0.69) ( $p < 0.001$ ). Workers over age 45 (OR = 1.26, 95% CI 1.14 to 1.40,  $p < 0.001$ ) and women (OR = 1.12, 95% CI 1.01 to 1.25,  $p = 0.03$ ) were at greater risk of job strain. Furthermore, workers over age 45 (OR = 1.51, 95% CI 1.32 to 1.73,  $p < 0.001$ ), workers with over 10 years of service (OR = 1.35, 95% CI 1.16 to 1.56,  $p < 0.001$ ), and women (OR = 1.15, 95% CI 1.00 to 1.31,  $p = 0.04$ ) were at greater risk of isostrain. Conclusions: Managers seem to have higher autonomy and greater social support and therefore are less at risk of job strain or isostrain than employees. Other factors such as age, seniority, and sex may influence this relationship. Trial Registration: Clinicaltrials.gov: NCT02596737.

**Keywords:** job demand; job control; job support; job strain; isostrain; manager; Karasek's model; work



**Citation:** Colin-Chevalier, R.; Pereira, B.; Benson, A.C.; Dewavrin, S.; Cornet, T.; Dutheil, F. The Protective Role of Job Control/Autonomy on Mental Strain of Managers: A Cross-Sectional Study among Wittyfit's Users. *Int. J. Environ. Res. Public Health* **2022**, *19*, 2153. <https://doi.org/10.3390/ijerph19042153>

Academic Editor: Paul B. Tchounwou

Received: 20 December 2021

Accepted: 10 February 2022

Published: 14 February 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Karasek's Job Demand-Control-Support (JDCS) model [1–5] is a theory that helps to explain the relationship between job characteristics and the psychological well-being of workers. The model explains how the interaction of job demand (or psychological demand, e.g., work overload or role conflict), job control (or decision latitude, e.g., control over work pace or even workplace autonomy), and job support (or social support, e.g., emotional, physical, physiological, or practical assistance) can cause mental strain in individuals [1,2,6,7]. Job strain results from the interaction between high job demand and low job control, and isostrain is characterized by job strain with low social support. Job strain and isostrain have emerged as major public health problems with hazardous effects on workers [8–20].



Several factors such as age, seniority, or even sex are known to have an impact on job strain among workers [21–24]. To the best of our knowledge, there are few studies on the links between one’s position within a company (manager or employee) and job strain or isostrain risk [25,26], or studies which were carried out in one industry only, such as healthcare [15,27]. Despite managers’ increased risk of stress [28] due to the multitude of tasks and responsibilities they face [29], their autonomy may overcome their job demand. As managers occupy a crucial role in a company, it therefore seems necessary to evaluate the impact of managerial status on the risk of mental strain, but also to identify other potential factors which may affect this relationship.

Thus, the main objective of our study was to evaluate the influence of job position (manager or employee) on the perception of work (job demand, job control, and job support) and mental strain (job strain and isostrain), particularly the putative protective role of job control among managers. The secondary objectives were to investigate other sociodemographic factors that may influence participant responses, using Karasek’s model, and to quantify the impact of all these factors.

## 2. Materials and Methods

### 2.1. Recruitment

Wittyfit software is a web platform designed to promote and improve well-being in the workplace, implemented in various French industries that are clients of the Wittyfit company (Paris, France) [30]. Any worker wishing to participate in the Wittyfit study could register for free and access the eponymous application, which can be downloaded to a desktop computer or smartphone. Volunteer workers were asked to answer various questionnaires. They could anonymously provide personal information about their physical, mental, and work-related feelings in order to participate in the improvement of their own well-being. Managers and employees received personal and individualized feedback that could be used to implement a preventive strategy. The Wittyfit software was designed in partnership with the University Hospital of Clermont-Ferrand. All companies and workers using Wittyfit on a voluntary basis were included in the study. We limited the inclusion of data up to February 2020, prior to the COVID-19 pandemic.

### 2.2. The Job Demand-Control-Support Model

The three measurements of Karasek’s model, namely, perceived job demand, job control, and job support, were assessed through the Wittyfit platform using visual analogical scales (VASs). VASs are widely used to assess perceived stress in workers [31–33]. Wittyfit’s VASs are graded from 0 to 100, with 0 indicating a low level and 100 a high level. Therefore, job demand and job control were, respectively, assessed using the Wittyfit’s VASs “workload” and “autonomy” from the “stress” questionnaire. Job support was assessed by the “ambiance” VAS from the “job satisfaction” questionnaire, designed to assess social support at work [30]. Then, the values were rescaled to match the true measurement scales. From a score of 0 to 100, the levels of job demand, job control, and job support were rescaled to match the true measurement scales of the Karasek questionnaire. Thus, the VAS from 0 to 100 for job demand was converted into a score from 9 to 36, job control into a score from 24 to 96, and job support into a score from 8 to 32. From this conversion, we were able to apply thresholds to our variables to transform them into qualitative variables to determine whether a worker was in a job strain or isostrain situation. Workers with more than 21 out of 36 in job demand and less than 70 out of 96 in job control were considered to be in a job strain situation. Job strain workers with less than 23 out of 32 in job support were considered to be in an isostrain situation. These values were chosen in accordance with the literature [6]. Throughout the duration of the study, when logging back into the platform, the workers could, if they wished, re-express how they felt in the different categories. The last scores entered by the user were used for the study.

### 2.3. Sociodemographic Characteristics of Workers

Sociodemographic characteristics considered were age, seniority, sex, position (manager or employee), and company they work for. A manager was defined as a worker with primary responsibility, having a leadership position over a group of workers, as opposed to an employee. This information was provided by Wittyfit's client companies and stored in the Wittyfit database. Thus, for each worker, age, seniority (in age groups), sex (male or female), and position (manager or employee) were known. Age was transformed from a multiclass variable into a binary variable indicating whether the worker was over or under 45 years old. In the same way, seniority was transformed into a binary variable indicating whether the worker had over or under 10 years of service.

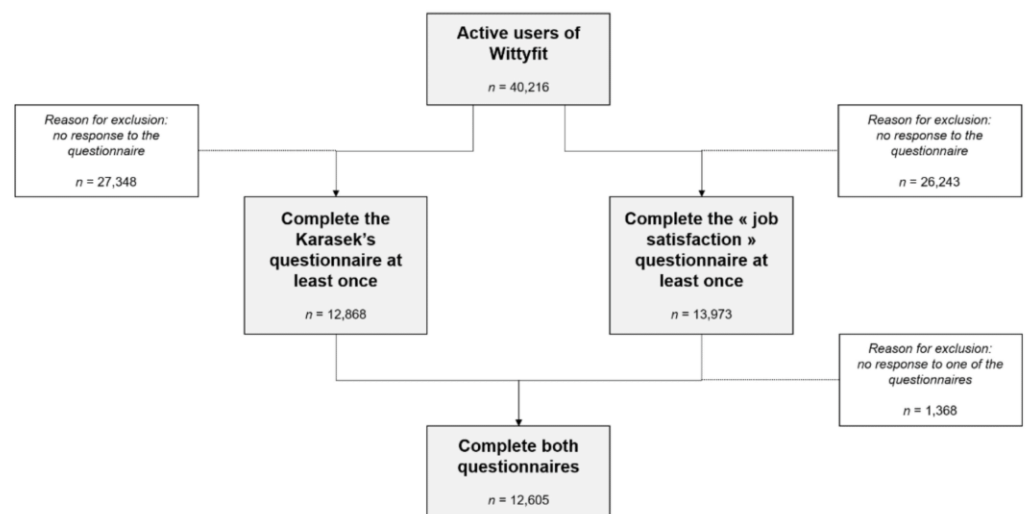
### 2.4. Statistical Analyses

Quantitative data (job demand, control, and support), expressed by the mean  $\pm$  standard deviation (SD), were normalized using a "refit" method [34] and were compared between independent groups (by age, seniority, sex, and job position) with ANOVA. Qualitative data (job strain and isostrain) were described by a number of participants, and associated frequencies were compared between groups with the  $\chi^2$  test. Then, multivariate analysis was conducted using linear or logistic mixed regressions depending on the outcome's nature (quantitative or qualitative, respectively) to determine variables associated with different Karasek model outcomes. The "company" effect was included as a random effect. Model assumptions (residual independence and normality, and variance homogeneity) were verified a posteriori. Multicollinearity of covariates was checked prior to analysis. Unless specified, all estimates were interpreted in terms of effect sizes (ESs) or standardized odds ratios (ORs) and 95% confidence intervals (95% CIs). ESs were interpreted according to Funder's rules [35]. Statistical analyses were performed with R (version 4.0.4) [36] in the RStudio (version 1.3.1056) platform. A  $p$ -value  $< 0.05$  was considered statistically significant for all analyses.

## 3. Results

### 3.1. Participants

Data of 15,562 workers were collected between January 2018 and February 2020. We obtained complete data from 5656 men and 3601 women, i.e., 9257 workers, including 769 managers (8.3% of workers) (Figure 1).



**Figure 1.** Flow chart of Wittyfit's users.

Table 1 shows the sociodemographic characteristics of workers by score obtained using Karasek's model.

**Table 1.** Characteristics of workers according to their classification using Karasek’s Job Demand-Control-Support model (JDCS).

Variable	JDCS Model Groups				p-Value
	Active	High Strain	Low Strain	Passive	
JDCS (mean ± SD)					
n (%)	5537 (59.8%)	2235 (24.1%)	927 (10.0%)	558 (6.0%)	
Job demand	29.3 ± 4.2	27.3 ± 4.5	15.9 ± 3.4	15.8 ± 3.6	<0.001
Job control	85.1 ± 7.7	55.3 ± 11.2	84.0 ± 8.0	50.5 ± 13.6	<0.001
Job support	25.2 ± 5.6	21.8 ± 6.1	25.2 ± 5.9	21.0 ± 6.7	<0.001
Age (years), n (%)					
<45	2829 (51.1%)	1024 (45.8%)	523 (56.4%)	293 (52.5%)	<0.001
≥45	2708 (48.9%)	1211 (54.2%)	404 (43.6%)	265 (47.5%)	
Seniority (years), n (%)					
<10	3229 (58.3%)	1238 (55.4%)	638 (68.8%)	369 (66.1%)	<0.001
≥10	2308 (41.7%)	997 (44.6%)	289 (31.2%)	189 (33.9%)	
Sex, n (%)					
Male	3349 (60.4%)	1302 (58.3%)	644 (69.5%)	364 (65.2%)	<0.001
Female	2191 (39.6%)	933 (41.7%)	283 (30.5%)	194 (34.8%)	
Position, n (%)					
Employee	5041 (91.0%)	2097 (93.8%)	826 (89.1%)	524 (93.8%)	<0.001
Manager	91 (9.0%)	93.8 (6.2%)	101 (10.9%)	34 (6.1%)	

Of the 9257 workers included in the study, 5537 (59.8%) had an “active” job (high demand, high control) according to the Karasek model, 558 (6.0%) a “passive” job (low demand, low control), 927 (10.0%) a “low strain” job (low demand, high control), and 2235 (24.1%) a “high strain” job and were in a job strain situation (high demand, low control). Further, among job strain workers, 1192 individuals (12.9% of the total population) were in an isostrain situation.

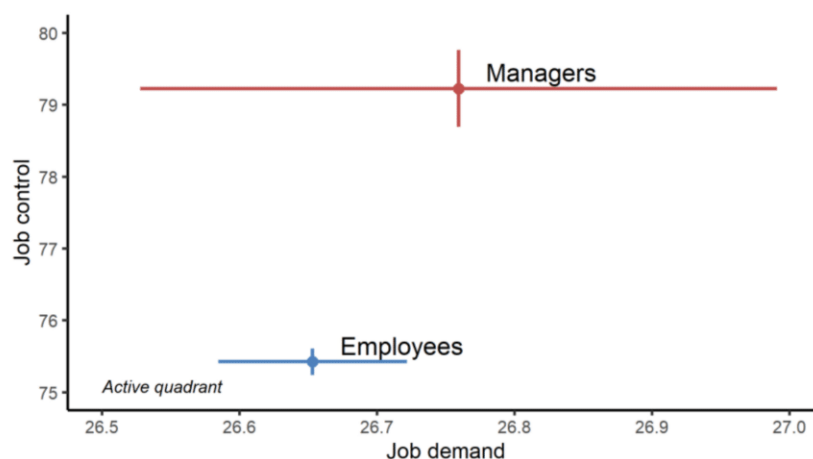
### 3.2. Impact of Sociodemographic Characteristics of Workers on Responses to the JDCS Model

#### 3.2.1. Job Demand, Control, and Support among Workers

On average, managers had significantly higher job control ( $79.2 \pm 14.9$  vs.  $75.4 \pm 16.9$ ,  $p < 0.001$ ) and job support ( $25.2 \pm 5.1$  vs.  $24.0 \pm 6.1$ ,  $p < 0.001$ ) than employees, but there was a non-significant difference in job demand ( $26.8 \pm 6.4$  vs.  $26.7 \pm 6.3$ ,  $p = 0.3$ ) (Figure 2). According to our results, both managers and employees using the Wittyfit software felt, on average, “active” at work (Figure 3).

**Figure 2.** Job demand, job control, and job support by position. Legend: \*\*\*  $p < 0.001$ .

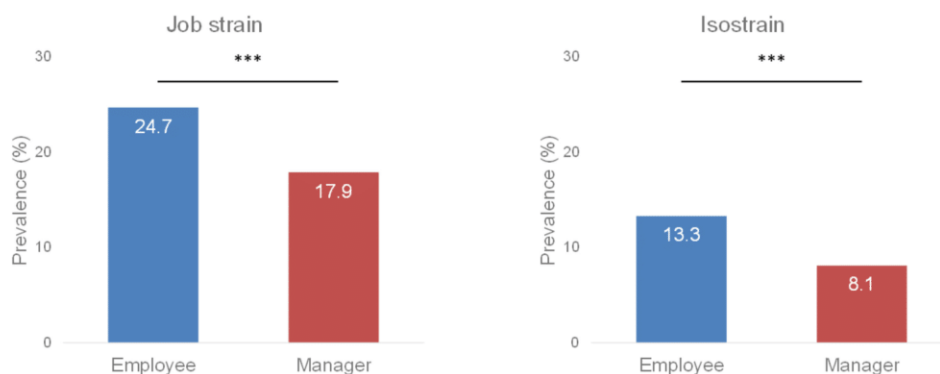




**Figure 3.** Karasek's Job Demand-Control-Support model. Point ranges are expressed as means and standard errors. Job demand varies from 9 to 36, job control from 24 to 96.

### 3.2.2. Job Strain and Isostrain Prevalence by Job Position

Managers had a lower risk of job strain (OR = 0.63, 95% CI 0.52 to 0.77,  $p < 0.001$ ) and isostrain (OR = 0.53, 95% CI 0.40 to 0.69,  $p < 0.001$ ) than employees. Thus, there was a lower prevalence of job strain (17.9% of managers vs. 24.7% of employees,  $p < 0.001$ ) and isostrain (8.1% of managers vs. 13.3% of employees,  $p < 0.001$ ) among managers (Figure 4).



**Figure 4.** Prevalence of job strain and isostrain by job position. Legend: \*\*\*  $p < 0.001$ .

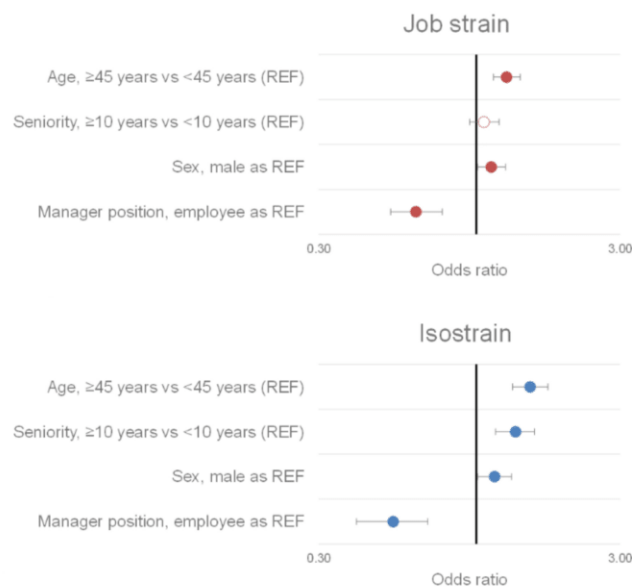
### 3.2.3. Sociodemographic Factors of Karasek's Model

Age emerged as a factor that influenced some dimensions of Karasek's model among all workers (both managers and employees). Indeed, workers over 45 had, on average, higher job demand (ES = 0.06, 95% CI 0.01 to 0.10,  $p = 0.009$ ), lower job control (ES =  $-0.06$ , 95% CI  $-0.11$  to  $-0.02$ ,  $p = 0.004$ ), and lower job support (ES =  $-0.19$ , 95% CI  $-0.24$  to  $-0.15$ ,  $p < 0.001$ ). Moreover, similar to age, workers with over 10 years of seniority reported having higher job demand (ES =  $-0.15$ , 95% CI  $-0.20$  to  $-0.10$ ,  $p = 0.02$ ) and less job support (ES =  $-0.21$ , 95% CI  $-0.26$  to  $-0.16$ ,  $p < 0.001$ ) (Figure 5).

Furthermore, women reported having higher job demand (ES = 0.05, 95% CI 0.01 to 0.09,  $p = 0.02$ ) than men but lower job support (ES =  $-0.10$ , 95% CI  $-0.14$  to  $-0.06$ ,  $p < 0.001$ ). Finally, our results show that the prevalence of job strain was higher among over-45 workers (OR = 1.26, 95% CI 1.14 to 1.40,  $p < 0.001$ ) and women (OR = 1.12, 95% CI 1.01 to 1.25,  $p = 0.03$ ). Likewise, the prevalence of isostrain was higher among over-45 workers (OR = 1.51, 95% CI 1.32 to 1.73,  $p < 0.001$ ), workers with 10 or more years of service (OR = 1.35, 95% CI 1.16 to 1.56,  $p < 0.001$ ), and women (OR = 1.15, 95% CI 1.00 to 1.31,  $p = 0.04$ ) (Figure 6).



**Figure 5.** Sociodemographic factors of the Karasek model dimensions. A colored circle means that the measured effect is significant. The reference classes (REF) of the different parameters are the following: age: workers under 45; seniority: workers with less than 10 years of service; sex: male; job position: employee.



**Figure 6.** Sociodemographic factors of job strain and isostrain. A colored circle means that the measured effect is significant. The reference classes (REF) of the different parameters are the following: age: workers under 45; seniority: workers with under less than 10 years of service; sex: male; job position: employee.

## 4. Discussion

The main results show that, due to higher job control and job support, managers were at less risk of job strain and isostrain than employees. However, other characteristics such as age, seniority, or sex were found to have an impact on the JDCS dimensions [5], sometimes protective, sometimes hazardous.

### 4.1. The Managers and Karasek's Model

Among workers who answered the JDCS questionnaire, we found that managers were less exposed to the risk of job strain and isostrain than employees. Indeed, only 17.9% of managers were in a job strain situation, compared to 24.7% of employees. Likewise, only 8.1% of managers were in an isostrain situation, compared to 13.3% of employees. In particular, the results highlight that the high levels of job control and job support perceived by the managers protect them from these risks. This suggests that the role of the manager, and more specifically the specificities of the manager's work, as opposed to that of the employee, acts as a protective factor against job strain and isostrain. These results confirm previous findings showing the impact of socioeconomic status on workers' type of job, on the one hand, and on the prevalence of job strain and isostrain, on the other, among workers, particularly the lower prevalence for workers with a higher socioeconomic position such as managers [25,26,37].

We did not show a difference in job demand between managers and employees. This result may seem surprising, especially since studies have shown higher job demand for managers [38]. Nevertheless, in our study, the three outcomes measured relate to the employee's perception of his or her own activity. In particular, a manager is generally expected to take on management tasks in addition to their own, unlike an employee. However, since these additional tasks are an integral part of a manager's job, they may factor into the perceived workload scale. This could explain the absence of a significant difference between employees and managers in their perceived job demand. However, managers reported having slightly higher job control and job support than employees. The strong empowerment (described in the literature as a form of control or autonomy) of managers has been shown previously in nurses [39,40], with a positive correlation between empowerment and positive work-related outcomes [39]. The difference in social support between managers and employees is coherent, as managers may receive support from their subordinates in addition to workers of the same or higher grade, improving the perception of their work [41]. Institutions should work to improve decision latitude and social support in the workplace. This should result in a consequent reduction in job strain and isostrain risks for workers, resulting in an improvement in their occupational well-being, especially for employees who are at greater risk.

### 4.2. The Other Factors of Job Strain and Isostrain

Sociodemographic factors may affect perceptions of work. In our study, we found a positive influence of age on job demand and a negative influence on job control and job support. This resulted in a higher risk of job strain and isostrain for people older than 45 years old. Previous studies have shown that, on the contrary, the risk of job strain seems to decrease with age [38] or has no effect [6]. Similarly, we found that seniority was positively associated with job demand and negatively associated with job support, resulting in a higher risk of isostrain for workers with more than 10 years of seniority. This suggests that, over the years, workers feel that the work they are asked to do becomes increasingly demanding, and that they lose the support from their colleagues or their management, thereby increasing their risk of strain.

According to our results, it seems that women perceived higher job demand than men while receiving less job support, which has previously been shown among managers [22]. Thus, women are slightly more at risk of job strain and isostrain than men, as already established in previous studies [24,38]. While the harmful effects of stress at work are well known among women [42], our study also shows that this is a population at risk.

Our results confirm that personal factors can alter the perception of work. This reinforces the idea that the Karasek model should not be applied without considering the sociodemographic attributes of workers. Institutions should strive to ensure that older (and more senior) workers feel as regarded as younger workers. They must also be aware of potential differences in the perception of work between men and women, in order to take all the measures to enable their employees, regardless of their sex, to feel on an equal footing.

#### 4.3. Limitations

We acknowledge some limitations in our study. Despite the fact several factors such as education or shift work are known to have an impact on job strain [24], they are not retrieved by Wittyfit. The use of self-reported responses to questionnaires could have led to over- or under-estimated values, which might lead to a measurement bias. Additionally, the use of a single VAS to assess each dimension of Karasek's model, instead of a multi-item questionnaire, may reinforce the idea of an affective bias. However, the quadrant classification of the JDCS model, based on historically studied thresholds, allowed us to obtain figures in agreement with the literature, with these values being very close to our first quartile (Q1) and tercile (T1) values (workload: Q1 = 22.5, T1 = 23.9; autonomy: Q1 = 65.8, T1 = 73.7; ambiance: Q1 = 20.0, T1 = 21.0), often used as cut-offs to detect job strain and isostrain risk in various studies [43]. With this construct, we found that 24% of the population was in a job strain situation and 13.1% in an isostrain situation, in line with the literature [9,11,24]. Further, as an ever-expanding, real-world database, we can expect further investigations and comparisons of these results, in the future, with a larger dataset, even with repeated measurements and thus with more measurements per individual. Numerous companies were included in the study, but they only represented a small sample of possible business types. Nevertheless, the use of mixed models allowed us to take into account a "company" effect and to separately estimate the effects of sociodemographic characteristics. As Wittyfit's clients did not provide further details about their employees, such as their job, our sample might be non-representative of French workers. However, the use of the "Wittyfit" database allowed us to work with a large sample representative of various management positions [30], which was our main focus. These considerations allow us to assume the generalizability of our results.

## 5. Conclusions

Managers have a lower risk of job strain and isostrain. Indeed, they have higher levels of control and support at work than employees. Older workers and those with higher seniority may also be slightly more at risk of job strain. Irrespective of job position, acting on the main dimensions of Karasek's model (job demand, job control, and job support) at work can be beneficial in improving the well-being of workers. Thus, implementing strategies to improve, among other things, decision latitude and social support in the workplace could help companies reduce the risk of job strain and isostrain.

**Author Contributions:** Study design: F.D., S.D. and T.C.; study conduct: R.C.-C., F.D., B.P. and S.D.; data collection: F.D. and S.D.; data analysis: R.C.-C. and B.P.; data interpretation: R.C.-C., F.D. and B.P.; drafting manuscript: R.C.-C.; revising manuscript content: R.C.-C., F.D., B.P. and A.C.B.; approving final version of manuscript: all authors. R.C.-C., B.P., S.D. and F.D. take responsibility for the integrity of the data analysis. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study received no external funding.

**Institutional Review Board Statement:** This study was approved by the National Commission for Data Protection and Liberties (CNIL), and the South-East VI ethics committee (clinicaltrials.gov identification number NCT02596737).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study. This information is set out in Wittyfit's terms and conditions of use.



**Data Availability Statement:** Data from Wittyfit cannot be transmitted without the prior consent of the company’s corporate clients, except to the University Hospital of Clermont-Ferrand, France, which may use the data for research purposes.

**Acknowledgments:** We express our sincere gratitude to all voluntary workers using Wittyfit, who participated in this study.

**Conflicts of Interest:** R.C.-C., S.D. and T.C. are employed by Wittyfit. The other authors declare that no competing interests exist (F.D. is responsible for the scientific accuracy of Wittyfit but is not paid by Wittyfit; as previously published, Wittyfit is a public–private partnership with the CHU Clermont-Ferrand).

## References

1. Karasek, R.; Brisson, C.; Kawakami, N.; Houtman, I.; Bongers, P.; Amick, B. The Job Content Questionnaire (JCQ): An instrument for internationally comparative assessments of psychosocial job characteristics. *J. Occup. Health Psychol.* **1998**, *3*, 322–355. [[CrossRef](#)] [[PubMed](#)]
2. Karasek, R.; Baker, D.; Marxer, F.; Ahlbom, A.; Theorell, T. Job decision latitude, job demands, and cardiovascular disease: A prospective study of Swedish men. *Am. J. Public Health* **1981**, *71*, 694–705. [[CrossRef](#)] [[PubMed](#)]
3. Karasek, R.A. Job Demands, Job Decision Latitude, and Mental Strain: Implications for Job Redesign. *Adm. Sci. Q.* **1979**, *24*, 285–308. [[CrossRef](#)]
4. Johnson, J.V.; Hall, E.M. Job strain, work place social support, and cardiovascular disease: A cross-sectional study of a random sample of the Swedish working population. *Am. J. Public Health* **1988**, *78*, 1336–1342. [[CrossRef](#)] [[PubMed](#)]
5. Karasek, R.; Theorell, T. *Healthy Work: Stress, Productivity, and the Reconstruction of Working Life*; Basic Books: New York, NY, USA, 1990.
6. Niedhammer, I.; Chastang, J.F.; Gendrey, L.; David, S.; Degioanni, S. Psychometric properties of the French version of Karasek’s «Job Content Questionnaire» and its scales measuring psychological pressures, decisional latitude and social support: The results of the SUMER. *Sante Publique* **2006**, *18*, 413–427. [[CrossRef](#)]
7. Niedhammer, I.; Chastang, J.-F.; Levy, D.; David, S.; Degioanni, S.; Theorell, T. Study of the validity of a job-exposure matrix for psychosocial work factors: Results from the national French SUMER survey. *Int. Arch. Occup. Environ. Health* **2008**, *82*, 87–97. [[CrossRef](#)]
8. Collins, S.M.; Karasek, R.A.; Costas, K. Job strain and autonomic indices of cardiovascular disease risk. *Am. J. Ind. Med.* **2005**, *48*, 182–193. [[CrossRef](#)]
9. Rugulies, R.; Krause, N. Job strain, iso-strain, and the incidence of low back and neck injuries. A 7.5-year prospective study of San Francisco transit operators. *Soc. Sci. Med.* **2005**, *61*, 27–39. [[CrossRef](#)] [[PubMed](#)]
10. Bugajska, J.; Jędryka-Góral, A.; Widerszal-Bazyl, M.; Orłowska-Baranowska, E.; Sagan, A.; Michalak, J.M.; Zuzewicz, K.; Konarska, M. Job strain, overtime, life style, and cardiovascular risk in managers and physical workers. *Int. J. Occup. Saf. Ergon.* **2011**, *17*, 25–32. [[CrossRef](#)] [[PubMed](#)]
11. Kivimäki, M.; Nyberg, S.T.; Batty, G.D.; Fransson, E.I.; Heikkilä, K.; Alfredsson, L.; Bjorner, J.B.; Borritz, M.; Burr, H.; Casini, A.; et al. Job strain as a risk factor for coronary heart disease: A collaborative meta-analysis of individual participant data. *Lancet* **2012**, *380*, 1491–1497. [[CrossRef](#)]
12. Steptoe, A.; Kivimäki, M. Stress and cardiovascular disease. *Nat. Rev. Cardiol.* **2012**, *9*, 360–370. [[CrossRef](#)]
13. Landsbergis, P.A.; Dobson, M.; Koutsouras, G.; Schnall, P. Job strain and ambulatory blood pressure: A meta-analysis and systematic review. *Am. J. Public Health* **2013**, *103*, e61–e71. [[CrossRef](#)]
14. Linton, S.J.; Kecklund, G.; Franklin, K.A.; Leissner, L.C.; Sivertsen, B.; Lindberg, E.; Svensson, A.C.; Hansson, S.O.; Sundin, Ö.; Hetta, J.; et al. The effect of the work environment on future sleep disturbances: A systematic review. *Sleep Med. Rev.* **2015**, *23*, 10–19. [[CrossRef](#)]
15. Wong, C.A.; Laschinger, H.K.S. The influence of frontline manager job strain on burnout, commitment and turnover intention: A cross-sectional study. *Int. J. Nurs. Stud.* **2015**, *52*, 1824–1833. [[CrossRef](#)]
16. Mutambudzi, M.; Javed, Z. Job strain as a risk factor for incident diabetes mellitus in middle and older age U.S. workers. *J. Gerontol. Ser. B* **2016**, *71*, 1089–1096. [[CrossRef](#)]
17. Madsen, I.E.H.; Nyberg, S.T.; Magnusson Hanson, L.L.; Ferrie, J.E.; Ahola, K.; Alfredsson, L.; Batty, G.; Bjorner, J.B.; Borritz, M.; Burr, H.; et al. Job strain as a risk factor for clinical depression: Systematic review and meta-analysis with additional individual participant data. *Psychol. Med.* **2017**, *47*, 1342–1356. [[CrossRef](#)] [[PubMed](#)]
18. Ferrario, M.M.; Veronesi, G.; Bertù, L.; Grassi, G.; Cesana, G. Job strain and the incidence of coronary heart diseases: Does the association differ among occupational classes? A contribution from a pooled analysis of Northern Italian cohorts. *BMJ Open* **2017**, *7*, e014119. [[CrossRef](#)] [[PubMed](#)]
19. Jun, D.; O’Leary, S.; McPhail, S.M.; Johnston, V. Job strain and psychological distress in office workers: The role of coping. *Work* **2019**, *64*, 55–65. [[CrossRef](#)] [[PubMed](#)]

20. Clinchamps, M.; Auclair, C.; Prunet, D.; Pfabigan, D.; Lesage, F.-X.; Baker, J.S.; Parreira, L.; Mermillod, M.; Gerbaud, L.; Dutheil, F. Burnout among hospital non-healthcare staff: Influence of job demand-control-support, and effort-reward imbalance. *J. Occup. Environ. Med.* **2021**, *63*, e13–e20. [[CrossRef](#)] [[PubMed](#)]
21. Wall, T.D.; Jackson, P.R.; Mullarkey, S.; Parker, S.K. The demands—Control model of job strain: A more specific test. *J. Occup. Organ. Psychol.* **1996**, *69*, 153–166. [[CrossRef](#)]
22. Gadinger, M.C.; Fischer, J.E.; Schneider, S.; Terris, D.D.; Krückeberg, K.; Yamamoto, S.; Frank, G.; Kromm, W. Gender moderates the health-effects of job strain in managers. *Int. Arch. Occup. Environ. Health* **2010**, *83*, 531–541. [[CrossRef](#)] [[PubMed](#)]
23. Rivera-Torres, P.; Araque-Padilla, R.A.; Montero-Simó, M.J. Job stress across gender: The importance of emotional and intellectual demands and social support in women. *Int. J. Environ. Res. Public Health* **2013**, *10*, 375–389. [[CrossRef](#)]
24. Bonsaksen, T.; Thørrisen, M.M.; Skogen, J.C.; Aas, R.W. Who reported having a high-strain job, low-strain job, active job and passive job? The WIRUS Screening study. *PLoS ONE* **2019**, *14*, e0227336. [[CrossRef](#)] [[PubMed](#)]
25. Bosma, H.; Peter, R.; Siegrist, J.; Marmot, M. Two alternative job stress models and the risk of coronary heart disease. *Am. J. Public Health* **1998**, *88*, 68–74. [[CrossRef](#)] [[PubMed](#)]
26. Wege, N.; Dragano, N.; Erbel, R.; Jockel, K.-H.; Moebus, S.; Stang, A.; Siegrist, J. When does work stress hurt? Testing the interaction with socioeconomic position in the Heinz Nixdorf Recall Study. *J. Epidemiol. Community Health* **2008**, *62*, 338–341. [[CrossRef](#)]
27. Backman, A.; Sjögren, K.; Lövheim, H.; Edvardsson, D. Job strain in nursing homes—Exploring the impact of leadership. *J. Clin. Nurs.* **2018**, *27*, 1552–1560. [[CrossRef](#)]
28. Kim, K.-W.; Cho, Y.-H. The Moderating Effect of Managerial Roles on Job Stress and Satisfaction by Employees' Employment Type. *Int. J. Environ. Res. Public Health* **2020**, *17*, 8259. [[CrossRef](#)]
29. Brett, J.M.; Stroh, L.K. Working 61 plus hours a week: Why do managers do it? *J. Appl. Psychol.* **2003**, *88*, 67–78. [[CrossRef](#)] [[PubMed](#)]
30. Dutheil, F.; Duclos, M.; Naughton, G.; Dewavrin, S.; Cornet, T.; Huguet, P.; Chatard, J.C.; Pereira, B. Wittyfit-live your work differently: Study protocol for a workplace-delivered health promotion. *JMIR Res. Protoc.* **2017**, *6*, e6267. [[CrossRef](#)]
31. Lesage, F.-X.; Berjot, S.; Deschamps, F. Clinical stress assessment using a visual analogue scale. *Occup. Med.* **2012**, *62*, 600–605. [[CrossRef](#)]
32. Dutheil, F.; Trousselard, M.; Perrier, C.; Lac, G.; Chamoux, A.; Duclos, M.; Naughton, G.; Mnatzaganian, G.; Schmidt, J. Urinary interleukin-8 is a biomarker of stress in emergency physicians, especially with advancing age—The JOBSTRESS\* randomized trial. *PLoS ONE* **2013**, *8*, e71658. [[CrossRef](#)]
33. Dutheil, F.; Pereira, B.; Moustafa, F.; Naughton, G.; Lesage, F.-X.; Lambert, C. At-risk and intervention thresholds of occupational stress using a visual analogue scale. *PLoS ONE* **2017**, *12*, e0178948. [[CrossRef](#)]
34. Ben-Shachar, M.S.; Lüdtke, D.; Makowski, D. effectsize: Estimation of effect size indices and standardized parameters. *J. Open Source Softw.* **2020**, *5*, 2815. [[CrossRef](#)]
35. Funder, D.C.; Ozer, D.J. Evaluating effect size in psychological research: Sense and nonsense. *Adv. Methods Pract. Psychol. Sci.* **2019**, *2*, 156–168. [[CrossRef](#)]
36. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021; Available online: <https://www.R-project.org/> (accessed on 14 May 2021).
37. Vanroelen, C.; Levecque, K.; Louckx, F. Differential exposure and differential vulnerability as counteracting forces linking the psychosocial work environment to socioeconomic health differences. *J. Epidemiol. Community Health* **2010**, *64*, 866–873. [[CrossRef](#)]
38. Niedhammer, I. Psychometric properties of the French version of the Karasek Job Content Questionnaire: A study of the scales of decision latitude, psychological demands, social support, and physical demands in the GAZEL cohort. *Int. Arch. Occup. Environ. Health* **2002**, *75*, 129–144. [[CrossRef](#)]
39. Trus, M.; Razbadauskas, A.; Doran, D.; Suominen, T. Work-related empowerment of nurse managers: A systematic review. *Nurs. Health Sci.* **2012**, *14*, 412–420. [[CrossRef](#)] [[PubMed](#)]
40. Cicolini, G.; Comparcini, D.; Simonetti, V. Workplace empowerment and nurses' job satisfaction: A systematic literature review. *J. Nurs. Manag.* **2014**, *22*, 855–871. [[CrossRef](#)] [[PubMed](#)]
41. Penconek, T.; Tate, K.; Bernardes, A.; Lee, S.; Micaroni, S.P.M.; Balsanelli, A.P.; de Moura, A.A.; Cummings, G.G. Determinants of nurse manager job satisfaction: A systematic review. *Int. J. Nurs. Stud.* **2021**, *118*, 103906. [[CrossRef](#)] [[PubMed](#)]
42. Padkapayeva, K.; Gilbert-Ouimet, M.; Bielecky, A.; Ibrahim, S.; Mustard, C.; Brisson, C.; Smith, P. Gender/sex differences in the relationship between psychosocial work exposures and work and life stress. *Ann. Work. Expo. Health* **2018**, *62*, 416–425. [[CrossRef](#)] [[PubMed](#)]
43. de Alves, M.G.M.; Hökerberg, Y.H.M.; Faerstein, E. Trends and diversity in the empirical use of Karasek's demand-control model (job strain): A systematic review. *Revista Brasileira de Epidemiologia* **2013**, *16*, 125–136.

## Article n°3

The managerial role and psychosocial  
factors of job satisfaction:  
a cross-sectional study among  
Wittyfit's users

### **Article n°3**

The managerial role and psychosocial factors of job satisfaction: a cross-sectional study among Wittyfit's users

Colin-Chevalier R, Pereira B, Dewavrin S, Cornet T, Zak M, Benson AC, Dutheil F

#### **Résumé**

Cet article s'inscrit dans la continuité de l'article précédent et constitue le deuxième article de l'étude sur les différences de perceptions du travail entre managers et employés. Dans cet article, nous souhaitons comparés les niveaux de satisfaction au travail et la prévalence du risque d'insatisfaction entre les deux populations de travailleurs. Confrontés aux mêmes défis méthodologiques que ceux rencontrés dans l'article précédent (présence de clusters de données naturels, à savoir les différentes entreprises, et absence de certaines données d'étude), nous avons adopté une approche similaire pour les résoudre avant toute analyse. Nos analyses ont révélé qu'en moyenne, le niveau de satisfaction au travail des managers était plus élevé que celui des employés. Nous avons également montré que les managers étaient moins à risque d'insatisfaction au travail que les employés. D'autres facteurs, notamment l'âge et l'ancienneté des travailleurs, se sont montrés être des facteurs de risque d'insatisfaction. Hormis celui d'équilibre entre vie personnelle et vie privée, les niveaux d'ambiance de travail, de reconnaissance, de sens, d'organisation du travail et de valeurs professionnelles étaient aussi plus élevés chez les managers. Enfin, nous avons mis en évidence que les facteurs de satisfaction n'étaient pas les mêmes selon la position occupée. Il est de la responsabilité des entreprises de veiller à garantir que les travailleurs soient satisfaits et d'améliorer leurs conditions de travail, en particulier les employés qui présentent des niveaux de satisfaction plus faibles.



# The Managerial Role and Psychosocial Factors of Job Satisfaction

## A Cross-sectional Study Among Wittyfit's Users

Rémi Colin-Chevalier, PhD, Bruno Pereira, PhD, Samuel Dewavrin, ATPL, MSc, Thomas Cornet, MBA, Marek Zak, MsPT, PhD, Amanda Clare Benson, PhD, AEP, and Frédéric Dutheil, MD, PhD

**Objective:** Job satisfaction is an emerging indicator for measuring workers' occupational well-being; however, this has been poorly studied in managers. We aimed to explore job satisfaction between managers and employees and assess its factors. **Methods:** Data from Wittyfit's users were collected between January 2018 and February 2020. Volunteers anonymously provided their sociodemographic profile, their sense of job satisfaction, and their psychosocial feelings (ambiance, meaning, organization, recognition, values, work-life balance). **Results:** Data of 10,484 employees and 836 managers were collected. Job satisfaction was higher in managers than employees. All psychosocial factors had an impact on job satisfaction in workers. There was a higher prevalence of satisfied managers than employees. Managerial position and other sociodemographic variables influenced this prevalence. **Conclusions:** Managers seem to be more satisfied than employees. Psychosocial and sociodemographic factors can influence workers' job satisfaction.

**Keywords:** job satisfaction, manager, sociodemographic factors, occupational health

### LEARNING OUTCOMES

After reading this article, readers will be able to:

- Outline the difference in job satisfaction between managers and employees
- Outline the influence of job position on job satisfaction with that of other sociodemographic factors (age, seniority, gender)
- Compare and discuss the influence of psychosocial factors on managers' and employees' job satisfaction

Job satisfaction is a widely studied concept, with positive benefits both for workers and companies.<sup>1-5</sup> It is a major and increasingly used indicator to assess occupational health of different aspects of work,

making it a global importance at the workplace.<sup>6</sup> One possible definition of job satisfaction is "a pleasurable or positive emotional state resulting from the appraisal of one's job or job experiences."<sup>7</sup> In addition to common models used to assess occupational well-being of workers,<sup>8-12</sup> job satisfaction is an important attribute in measuring quality of life and stress at work.<sup>1,6</sup> Therefore, understanding the predictors of job satisfaction allows for the development of effective strategies to improve occupational well-being and business performance.<sup>13</sup> Numerous studies have examined the determinants of job satisfaction, emphasizing the importance of psychosocial factors that can significantly impact an individual's overall satisfaction with their job. Extensive research has examined the difference in job satisfaction that workers may experience according to sociodemographic characteristics, such as age, seniority, or gender.<sup>14-16</sup> However, there is a gap in the literature when it comes to understanding the differences in job satisfaction between managers and employees, and how common psychosocial factors uniquely impact these two distinct job positions. Although some studies have assessed differences in job satisfaction between managerial positions, they were limited to a specific working population, and not to the general population.<sup>14,17</sup> While managers are indeed employees within an organization, their job satisfaction can vary compared with their subordinates due to distinctions in their roles and responsibilities. One notable distinction is that managers often possess greater decision-making authority and autonomy, factors that can contribute to elevated levels of job satisfaction.<sup>18,19</sup> However, this decision-making ability can backfire, leading to lower levels of satisfaction.<sup>20</sup> It is therefore essential for the organization to look after job satisfaction of managers, especially as it is reflected in that of employees.<sup>21</sup>

Although there may be differences in the factors influencing job satisfaction between managers and employees, the literature has identified a number of common factors, particularly those referred to as psychosocial factors. Such factors include support and team cohesion,<sup>16,22</sup> recognition,<sup>23,24</sup> work-life balance,<sup>23,25-27</sup> role meaningfulness,<sup>13,28,29</sup> organizational support,<sup>13,30,31</sup> or professional values.<sup>19,32</sup> To the best of our knowledge, to date, there are no comparative studies in the general population of the influence of these six dimensions on job

From the Université Clermont Auvergne, CNRS, LaPSCo, Physiological and Psychosocial Stress, CHU Clermont-Ferrand, University Hospital of Clermont-Ferrand, Preventive and Occupational Medicine, Cegid, Clermont-Ferrand, France (R.C.-C., F.D.); CHU Clermont-Ferrand, University Hospital of Clermont-Ferrand, Biostatistics Unit, the Clinical Research and Innovation Direction, Clermont-Ferrand, France (E.P.); Cegid Lyon, France (S.D., T.C.); Institute of Health Sciences, Collegium Medicum, The Jan Kochanowski University of Kielce, Kielce, Poland (M.Z.); and Swinburne University of Technology, Department of Health and Biostatistics, Hawthorn, Australia (A.C.B.).

Rémi Colin-Chevalier, 0000-0002-0531-8689

Bruno Pereira, 0000-0003-2460-6720

Marek Zak, 0000-0003-0881-9232

Amanda Clare Benson, 0000-0003-2959-8969

Frédéric Dutheil, 0000-0002-1468-6029

Funding sources: None to disclose.

This study was approved by the National Commission for Data Protection and Liberties and the South-East VI ethics committee (clinicaltrials.gov identification number NCT02596737).

Conflict of interest: R.C.-C., S.D., and T.C. are part of Cegid. The other authors have declared that no competing interests exist (F.D. is responsible for the scientific accuracy of Wittyfit but is not paid by Cegid; as previously published, Wittyfit is a public private partnership with the CHU Clermont-Ferrand).

Authors' contributions: F.D. and S.D. did the study design. R.C., F.D., B.P., S.D. did the study conduct. S.D. and T.C. did the data collection. R.C. and B.P. did the data analysis. R.C., F.D., and B.P. did the data interpretation. R.C. did the drafting manuscript. R.C., F.D., B.P., M.Z., and A.C.B. revising manuscript content. All authors did the approving final version of manuscript. R.C., B.P., S.D., and F.D. take responsibility for the integrity of the data analysis

Data availability: All relevant data can be found in the article and in the Supplementary Digital content.

Ethical considerations and disclosures: Informed consent was obtained from all subjects involved in the study. This information is set out in Wittyfit's terms and conditions of use.

EQUATER Network reporting guidelines: All authors adhered to the Strengthening the Reporting of Observational studies in Epidemiology guidelines.

Supplemental digital contents are available for this article. Direct URL citation appears in the printed text and is provided in the HTML and PDF versions of this article on the journal's Web site ([www.joem.org](http://www.joem.org)).

Address correspondence to: Rémi Colin-Chevalier, PhD, Occupational and Environmental Medicine, CHU Clermont-Ferrand, 58 rue Montalembert, F-63000 Clermont-Ferrand, France ([r.colin6374@gmail.com](mailto:r.colin6374@gmail.com)).

Copyright © 2023 American College of Occupational and Environmental Medicine  
DOI: 10.1097/JOM.0000000000000317

satisfaction by position. It remains necessary to carry out a comprehensive analysis that directly measures workers' feelings and the importance they attach to these different dimensions of work in their overall satisfaction, depending on their position.

Thus, the main objective of our study was to explore the influence of the job position (manager or employee) on job satisfaction. The secondary objectives were to assess the other psychosocial factors in relation to the job position and their influence on job satisfaction, as well as the putative influence of sociodemographic factors, such as age, seniority, and sex.

## MATERIALS AND METHODS

### Recruitment

This cross-sectional study was conducted among Wittyfit's users. Wittyfit is an online platform<sup>33</sup> designed by the Cegid company (Lyon, France) in partnership with the CHU of Clermont-Ferrand to improve the professional well-being of workers.<sup>34</sup> Through different questionnaires, volunteers are invited to anonymously report how they feel in several health-related areas when logging into the platform. The platform provides workers with individualized feedback, and managers with a report of their team's average health status and the issues affecting them. Because of the COVID-19 pandemic, data collection for this study was limited to the period between January 2018 and February 2020. All workers who completely or partially answered the Wittyfit's "job satisfaction" questionnaire at least once during this period were included in the study. Workers who did not fill in their personal information (age, seniority, sex, or position in the company) were excluded. Informed consent was obtained from all subjects involved in the study. This information is set out in Wittyfit's terms and conditions of use. This study was approved by the National Commission for Data Protection and Liberties and the South-East VI ethics committee (clinicaltrials.gov identification number NCT02596737). This study was designed as a combined cross-sectional study, so the Strengthening the Reporting of Observational studies in Epidemiology checklist was followed to ensure the quality of the study reporting (see Supplementary Table 1, <http://links.lww.com/JOM/B459>).

## OUTCOMES

### Primary Outcome

Job satisfaction of workers was assessed via the "job satisfaction" questionnaire of the platform by using the visual analogical scale (VAS; score range from 0 to 100). Wittyfit uses VAS to assess health outcomes in the same way as it does to measure stress, for example.<sup>35-37</sup> Workers could estimate their job satisfaction as many times as they wanted. For each indicator (job satisfaction and psychosocial feelings), we used the average score to evaluate the feeling of a worker. We considered a worker to be satisfied when their score was higher than 50 and dissatisfied less than 50.

### Secondary Outcomes

Psychosocial indicators included workers feelings of ambiance, work-life balance (or simply balance), meaning, organization, recognition, and values. Ambiance at work assesses the relationship with the manager and colleagues, the quality of the work environment, and the social support at work. Work-life balance measures the prioritization of personal and professional activities by the worker, meaning it evaluated the interest, usefulness, fulfillment, and motivation of the worker. Organization assessed the definition of worker's job description, resources, level of worker responsibility, their confidence in management and colleagues, and satisfaction with their salary. Recognition investigated the respect between management and colleagues and the appreciation and acknowledgment of the contributions of the individ-

ual to the company. Finally, values explored the balance between corporate and personal values. Similar to job satisfaction, all psychosocial indicators were evaluated using the associated VAS in the questionnaire. Workers could estimate their feeling in each category as many times as they wanted, and their average score was retained. Wittyfit's clients provided the age, seniority, sex, and position (manager or employee) of their workers who use the product. Workers were then divided into two age categories (older or younger than 45 years) and two seniority classes (more or less than 10 years of service).

## STATISTICS

Quantitative data (job satisfaction and psychosocial feelings) were expressed in terms of mean  $\pm$  SD. In absence of data, assumed to be missing completely at random,<sup>38</sup> a mean imputation by client was applied to avoid complete-cases analysis bias.<sup>39</sup> Qualitative data (worker's job satisfaction/dissatisfaction) were expressed using number of participants and associated percentage. First, comparisons between groups (e.g., by sex, by job position) were performed with analysis of variance for continuous variables (job satisfaction and psychosocial feelings) and  $\chi^2$  tests for categorical variables (satisfied/dissatisfied workers). Then, associations between endogenous variables (job satisfaction) and exogenous variables (psychosocial and sociodemographic factors) were assessed with linear and logistic mixed regressions, respectively. The company to which the worker belongs was considered a random effect. Variance influence factor values of exogenous variables were assessed and revealed no serious multicollinearity (a value less than 4 was considered acceptable). Coefficients of linear regression models have been expressed as effect sizes (ESs) and confidence intervals (95% CI) using a "refit" method<sup>40</sup> and interpreted according to Funder and Ozer.<sup>41</sup> Coefficients of logistic regression models are expressed as standardized odds ratio (OR) and 95% CI and interpreted according to Chen et al.<sup>42</sup> Subgroup analyses were conducted by job position to compare the primary drivers of satisfaction in workers. Statistical analyses were performed using R (version 4.1.1) (R Foundation for Statistical Computing, Vienna, Austria)<sup>43</sup> in the RStudio (version 1.3.1056) (R Foundation for Statistical Computing, Boston, MA) platform. Unless specified, we considered a *P* value less than 0.05 as statistically significant.

## RESULTS AND DISCUSSION

### Participants

Of the 41,856 workers using Wittyfit in 2019, 11,320 answered the "job satisfaction" questionnaire at least once during the period. We obtained the data from 6753 men and 4567 women, separated as 10,484 employees and 836 managers, with 49.6% of the population younger than 45 and 56.0% with less than 10 years of service, distributed across 17 French companies (Fig. 1).

### Job Satisfaction Among Managers

Managers reported having a higher job satisfaction than employees (68.1  $\pm$  20.4 vs. 57.8  $\pm$  24.2, *P* < 0.001). Indeed, the position within the company had an influence on job satisfaction (ES = 0.46, 95% CI: 0.39 to 0.53, *P* < 0.001). Managers also tended to be more satisfied with their jobs, with a higher prevalence of managers highly satisfied with their jobs than employees (84.7 vs. 68.8%, *P* < 0.001) (Fig. 2). Managers were twice as likely to be more satisfied than employees (OR = 2.65, 95% CI: 2.18 to 3.23, *P* < 0.001) (Fig. 3).

### Results of the "Job Satisfaction" Questionnaire

Overall, the psychosocial feelings were higher in managers than in employees (Fig. 4). Ambiance (mean  $\pm$  SD, 71.2  $\pm$  20.9 vs. 66.1  $\pm$  24.2), meaning (66.9  $\pm$  21.8 vs. 56.1  $\pm$  23.1), organization (55.3  $\pm$  23.6 vs. 46.6  $\pm$  24.6), recognition (62.8  $\pm$  23.9 vs.



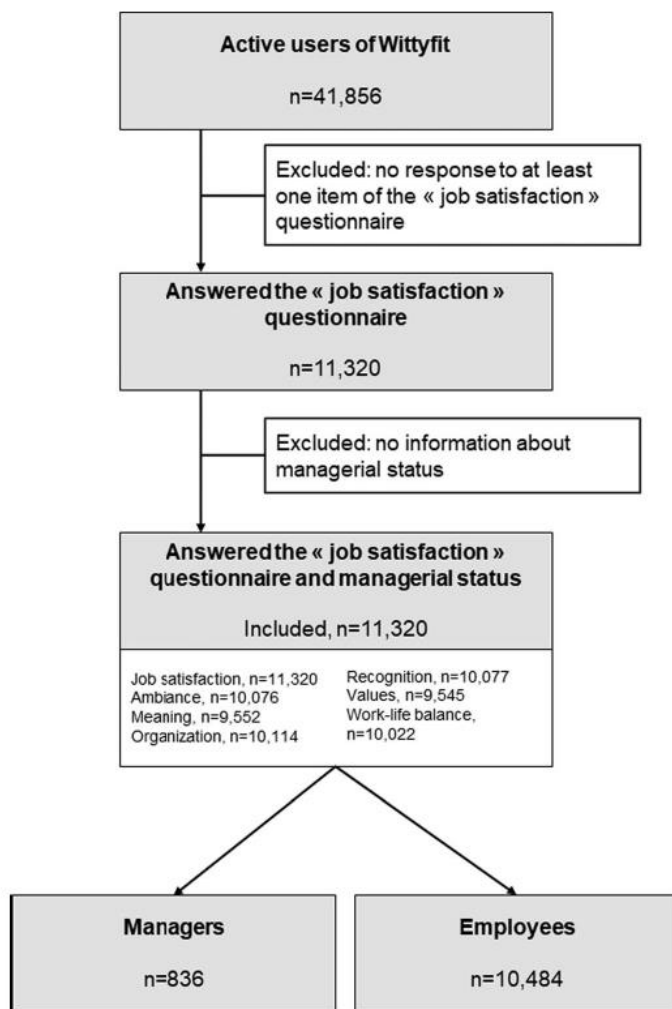


FIGURE 1. Flow chart of Wittyfit’s users.

48.3 ± 26.7), and values (66.3 ± 21.7 vs. 56.5 ± 23.1) were higher among managers ( $P < 0.001$ ). Only the work-life balance (balance) did not seem to be impacted by the position within the company (58.1 ± 23.9 vs. 59.2 ± 23.4,  $P = 0.2$ ).

Each psychosocial factor was found to have positively influenced job satisfaction regardless of the workers job position (Fig. 5). For managers, job satisfaction was mainly driven by recognition (ES = 0.48, 95% CI: 0.33 to 0.58) followed by values (0.43, 0.30 to 0.57) and organization (0.36, 0.23 to 0.49) ( $P < 0.001$ ). Among

employees, recognition (0.45, 0.41 to 0.48), meaning (0.41, 0.37 to 0.45), and ambiance (0.32, 0.29 to 0.36) were the psychosocial factors that most influenced level of job satisfaction ( $P < 0.001$ ). The balance was found to have the least contribution for both managers (0.18, 0.07 to 0.29) and employees (0.11, 0.07 to 0.14) ( $P < 0.001$ ).

The job satisfaction indicators of recognition (OR = 3.84, 95% CI: 2.11 to 6.99,  $P < 0.001$ ), values (2.23, 2.20 to 8.14,  $P < 0.001$ ) and organization (2.66, 1.30 to 5.45,  $P = 0.008$ ) remained the primary drivers of a high job satisfaction in managers. Balance (2.32, 1.29 to

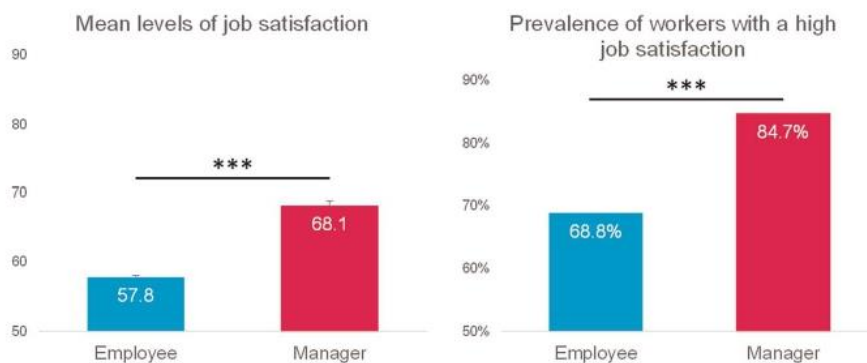
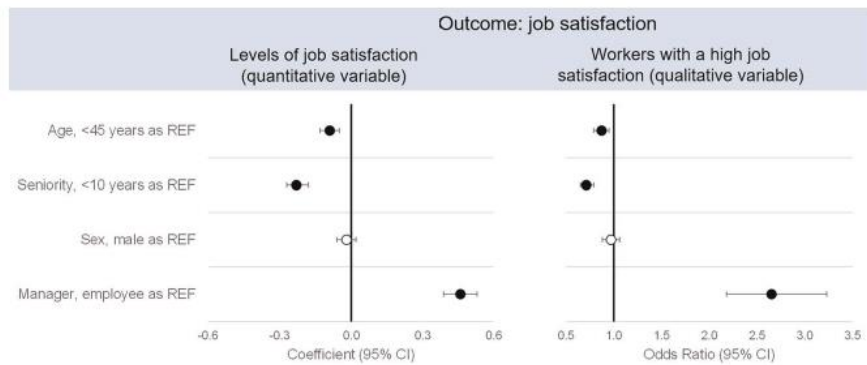


FIGURE 2. Job satisfaction by job position. \*\*\* $P < 0.001$ .



**FIGURE 3.** Influence of sociodemographic factors of job satisfaction. A colored circle means that the measured effect is significant, an empty circle means no effect. The reference classes of the different parameters are the following: age: workers younger than 45 years; seniority: workers with less than 10 years of service; sex: male; and job position: employee.

4.15,  $P = 0.005$ ) and meaning (1.87, 1.00 to 3.49,  $P = 0.049$ ) were also shown to be a driver of a high job satisfaction (2.05, 1.11 to 3.79,  $P = 0.02$ ), but not the ambiance (1.45, 0.85 to 2.49,  $P = 0.2$ ). In employees, recognition (3.14, 2.74 to 3.60) and meaning (2.70, 2.33 to 3.13) were the primary drivers of a high job satisfaction, followed by organization (2.34, 2.02 to 2.71) ( $P < 0.001$ ). Other psychosocial factors, namely, ambiance (2.02, 1.78 to 2.29), values (1.86, 1.59 to 2.17) and balance (1.34, 1.19 to 1.51), were also found to be drivers of a high job satisfaction in employees ( $P < 0.001$ ) (Fig. 5).

job satisfaction (0.97, 0.88 to 1.06,  $P = 0.5$ ) did not seem to differ by sex (Fig. 3).

**DISCUSSION**

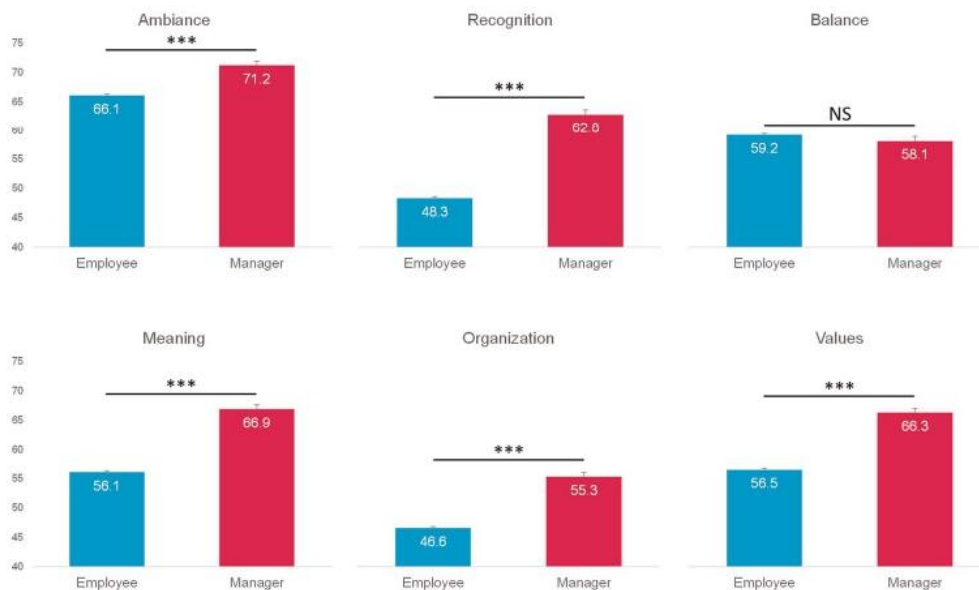
The main results show that in terms of job satisfaction and psychosocial feelings, the position within the company plays a crucial role. Psychosocial factors can influence the job satisfaction of managers and employees differently. Other sociodemographic factors were also found to have an impact on job satisfaction among workers.

**The Job Satisfaction and the Sociodemographics of Workers**

In addition to the major effect of the managerial position, some sociodemographic variables have been revealed as factors of job satisfaction, although their effect was much smaller than that of the job position. Workers older than 45 years ( $ES = -0.09$ , 95% CI:  $-0.13$  to  $-0.05$ ) and those with more than 10 years of seniority ( $-0.23$ ,  $-0.27$  to  $-0.18$ ) reported having a lower level of job satisfaction ( $P < 0.001$ ) (Fig. 3). Being older than 45 years old decreases the chances of being satisfied at work by 13% ( $OR = 0.87$ , 95% CI: 0.79 to 0.95,  $P = 0.002$ ) and having 10 or more years of seniority by 19% (0.71, 0.65 to 0.79,  $P < 0.001$ ). Finally, both job satisfaction level ( $-0.02$ ,  $-0.06$  to 0.02,  $P = 0.97$ ) and prevalence of workers with a high

**The Influence of the Manager’s Role on Job Satisfaction**

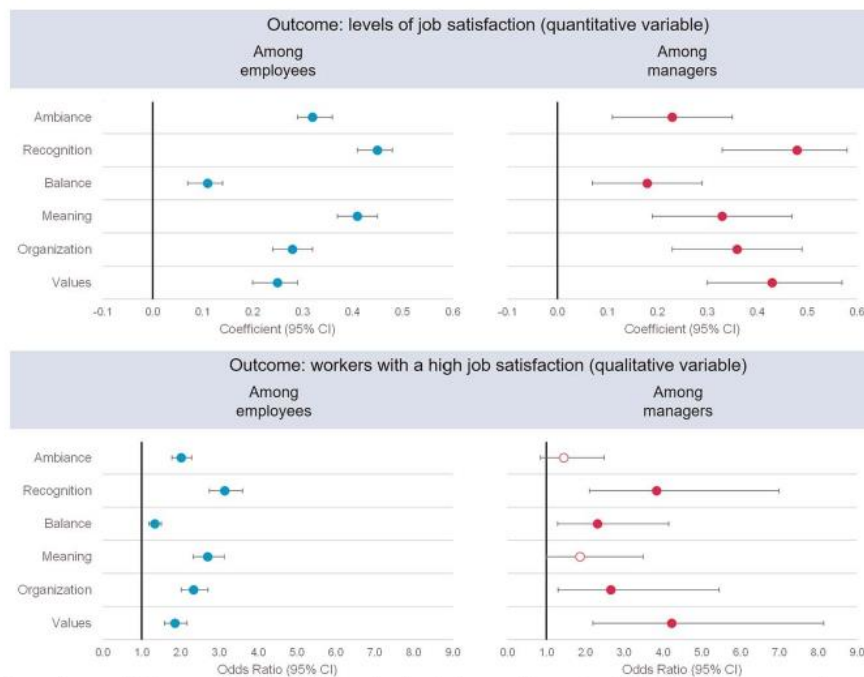
The role of the manager is crucial in the professional environment. Indeed, the manager has a heavy responsibility to achieve successful outcomes while ensuring that the expectations of their staff are met, both human and organizational.<sup>44,45</sup> Therefore, the benefits of good leadership are numerous.<sup>3,46,47</sup> Thus, if the role of the manager is to ensure the well-being and job satisfaction of their employees,<sup>48</sup> it is also essential to ensure that they are satisfied themselves. In this study, we aimed to assess the influence of the managerial position on both job satisfaction and psychosocial feelings. Our results show that on average, managers have higher levels of job satisfaction and are more satisfied with their job than employees. Previously,



**FIGURE 4.** Psychosocial feelings by job position. \*\*\* $P < 0.001$ , NS, not significant.

Downloaded from http://journals.lww.com/jeem by BIDMIBEPHKAVI2E0UMT10M4+KULHEZGSH04XMI0H0YWCX1AAW on 02/05/2024





**FIGURE 5.** Influence of psychosocial factors of job satisfaction by job position. A colored circle means that the measured effect is significant, an empty circle means no effect.

differences in perceived job satisfaction according to position in a healthcare environment have shown that physicians had the highest job satisfaction,<sup>14</sup> with registered nurses the lowest<sup>49</sup> among several healthcare professions. This may suggest that greater responsibilities attributed to managers, often associated with greater latitude in decision making, make them more satisfied than employees, who are themselves often the subject of management policies.<sup>50,51</sup> And yet, it remains important to note that these same responsibilities are likely to turn against managers, and so on employees.<sup>20,52</sup> Knowing that job satisfaction is linked to health benefits,<sup>1,2,4,5</sup> organizations should strive to improve the job satisfaction of employees, who seem to be less satisfied than managers.

### Psychosocial Feelings as Factors of Job Satisfaction

All considered psychosocial factors were found to influence job satisfaction for both employees and managers. Our findings show that job satisfaction is positively influenced by ambiance at work, which is in line with literature that found that it can be improved by encouraging social support and team cohesion.<sup>49</sup> Recognition at work can take several forms including notably recognition by patients or clients, by colleagues, or by management. A study among doctors in China found that doctors with higher patient recognition presented higher level of job satisfaction,<sup>53</sup> as observed in our study, with managers/employees reporting higher recognition having greater job satisfaction. While it may seem obvious that recognition is essential to employee satisfaction, it is just as much for managers.<sup>23</sup> Work-life balance was also found to be a factor of job satisfaction, which is in line with literature.<sup>26,54</sup> Even if it emerged as the least important dimension, as previously found,<sup>23</sup> ensuring a safe balance between professional and private life for managers (but also for employees) remains essential to protect from dissatisfaction. Meaning at work is defined as “the discovery of existential meaning from work experience, work itself and work purpose/goals.”<sup>55</sup> Thus, it is not surprising to find a direct link between meaning and job satisfaction as it translates among other things the motivation and the fulfillment of the worker. Organization commitment, known for being related to job satisfaction among nurses,<sup>13,56</sup> was also found to be positively related to job satisfaction.

Lastly, work values were positively related to job satisfaction, which is consistent with the findings of previous studies on the subject.<sup>19,57</sup> It has also been highlighted among doctors that those with higher work values scores presented higher job satisfaction.<sup>53</sup> Recommendations should be made for organizations to strive for improvement of workers’ psychosocial feelings, to consequently improve their job satisfaction.

### The Sociodemographic Factors of Job Satisfaction

Age and seniority also emerged as factors that could affect job satisfaction, with lower levels on average and even a lower prevalence of satisfied workers among the older workers (45 years or older) and more senior workers (10 years of service or more). The link between age and job satisfaction has already been demonstrated, but the results are equivocal. Sometimes job satisfaction increased with age,<sup>58</sup> sometimes it decreased<sup>59</sup> or presented a U-shape effect.<sup>60</sup> For the seniority, its negative influence on job satisfaction has already been reported in the literature.<sup>58</sup> No sex differences could be found in our study, although previous studies have shown that women pharmacists were more satisfied than men at work,<sup>61</sup> while the opposite relationship has been found in physicians.<sup>62,63</sup> Therefore, an absence of effect may not seem so surprising and suggests that, among Wittyfit’s users, the drivers of job satisfaction are the same for both men and women. Organizations must be aware that sociodemographic characteristics can influence job satisfaction to ensure that policies to improve well-being in the workplace are effective.

### LIMITATIONS

We acknowledge some limitations in our study. Self-reported feelings of workers (job satisfaction and psychosocial factors) were assessed using a single VAS instead of complete questionnaire as often used.<sup>64</sup> This may lead to overestimated or underestimated measurements that may cause a measurement bias as well as an affective bias.<sup>65</sup> Nevertheless, with a Cronbach’s  $\alpha$  of 0.84, we can assume that our questionnaire is reliable. Because of its size, our sample might not be representative of French workers. Indeed, it is not possible to state that it is representative in terms of jobs, nor even in terms of sectors

of activity, as this information is not provided by Wittyfit's clients. However, on the one hand, with the inclusion of 17 different companies, it can be assumed that numerous jobs were represented. On the other hand, the inclusion of a random effect in our regression models, namely, the "company" effect, allowed us to measure the influence of the company on the outcome independently of the fixed effects, namely, the psychosocial and sociodemographic factors, including the influence of the position held by the employee. Even taking these considerations into account, we believe generalizability of the findings from our study is possible. Age and seniority were also available only as categorical data, precluding further analyses.

## CONCLUSIONS

Managers have higher job satisfaction and psychosocial feelings at work than employees. Several psychosocial factors can influence workers' job satisfaction. Associated ESs may vary according to the level of position within the company. Sociodemographic factors, such as age and seniority, may influence job satisfaction as well. Companies should implement policies aimed at improving the well-being of their employees to improve their job satisfaction, regardless of their position.

## ACKNOWLEDGMENTS

The authors thank all voluntary workers using Cegid Wittyfit, who participated in this study.

## REFERENCES

- Rahman M, Sen AK. Effect of job satisfaction on stress, performance and health in self-paced repetitive work. *Int Arch Occup Environ Health* 1987;59:115–121.
- Bogler R, Somech A. Influence of teacher empowerment on teachers' organizational commitment, professional commitment and organizational citizenship behavior in schools. *Teach Teacher Educ* 2004;20:277–289.
- Han GH, Jekel M. The mediating role of job satisfaction between leader-member exchange and turnover intentions. *J Nurs Manag* 2011;19:41–49.
- Satuf C, Monteiro S, Pereira H, Esgalhado G, Marina Afonso R, Loureiro M. The protective effect of job satisfaction in health, happiness, well-being and self-esteem. *Int J Occup Saf Ergon* 2018;24:181–189.
- Hou J, He Y, Zhao X, et al. The effects of job satisfaction and psychological resilience on job performance among residents of the standardized residency training: a nationwide study in China. *Psychol Health Med* 2020;25:1106–1118.
- Faragher EB, Cass M, Cooper CL. The relationship between job satisfaction and health: a meta-analysis. *Occup Environ Med* 2005;62:105–112.
- Locke EA. The nature and causes of job satisfaction. In: Dunnette MD (ed.). *Handbook of Industrial and Organizational Psychology*. Chicago, IL, USA: Rand McNally; 1297–1349.
- Karasek RA. Job demands, job decision latitude, and mental strain: implications for job redesign. *Adm Sci Q* 1979;24:285–308.
- Karasek R, Theorell T. *Healthy Work: Stress, Productivity, and the Reconstruction of Working Life*. New York: Basic Books; 1990.
- Siegrist J. Adverse health effects of high-effort/low-reward conditions. *J Occup Health Psychol* 1996;1:27–41.
- de Jonge J, Bosma H, Peter R, Siegrist J. Job strain, effort-reward imbalance and employee well-being: a large-scale cross-sectional study. *Soc Sci Med* 2000;50:1317–1327.
- Siegrist J, Starke D, Chandola T, et al. The measurement of effort-reward imbalance at work: European comparisons. *Soc Sci Med* 2004;58:1483–1499.
- Lu H, Zhao Y, While A. Job satisfaction among hospital nurses: a literature review. *Int J Nurs Stud* 2019;94:21–31.
- Tarcan M, Hikmet N, Schooley B, Top M, Tarcan GY. An analysis of the relationship between burnout, socio-demographic and workplace factors and job satisfaction among emergency department health professionals. *Appl Nurs Res* 2017;34:40–47.
- Gazioglu S, Tansel A. Job satisfaction in Britain: individual and job related factors. *Applied Econ* 2006;38:1163–1171.
- Penconek T, Tate K, Bernardes A, et al. Determinants of nurse manager job satisfaction: a systematic review. *Int J Nurs Stud* 2021;118:103906.
- Krugman M, Smith V. Charge nurse leadership development and evaluation. *J Nurs Adm* 2003;33:284–292.
- Klopper HC, Coetzee SK, Pretorius R, Bester P. Practice environment, job satisfaction and burnout of critical care nurses in South Africa. *J Nurs Manag* 2012;20:685–695.
- Kantek F, Kaya A. Professional values, job satisfaction, and intent to leave among nursing managers. *J Nurs Res* 2017;25:319–325.
- Byrne A, Dionisi AM, Barling J, et al. The depleted leader: the influence of leaders' diminished psychological resources on leadership behaviors. *Leadersh Q* 2014;25:344–357.
- Bachman J, Henry R, Jackson C, Mitchell T, Crocco OS. Must it be lonely at the top? Developing leader well-being in organizations. *Adv Dev Hum Res* 2023;25:73–94.
- Soto-Rubio A, Giménez-Espert MDC, Prado-Gascó V. Effect of emotional intelligence and psychosocial risks on burnout, job satisfaction, and nurses' health during the COVID-19 pandemic. *Int J Environ Res Public Health* 2020;17:E7998.
- Rama-Maceiras P, Parente S, Kranke P. Job satisfaction, stress and burnout in anaesthesia: relevant topics for anaesthesiologists and healthcare managers? *Eur J Anaesthesiol* 2012;29:311–319.
- Weziak-Białowolska D, Białowolski P. Associations of recognition at work with subsequent health and quality of life among older working adults. *Int Arch Occup Environ Health* 2022;95:835–847.
- Susanto P, Hoque ME, Jannat T, Emely B, Zona MA, Islam MA. Work-life balance, job satisfaction, and job performance of SMEs employees: the moderating role of family-supportive supervisor behaviors. *Front Psychol* 2022;13:906876.
- Orkibi H, Brandt YI. How positivity links with job satisfaction: preliminary findings on the mediating role of work-life balance. *Eur J Psychol* 2015;11:406–418.
- Choi E, Kim J. The association between work-life balance and health status among Korean workers. *Work* 2017;58:509–517.
- Van Bogaert P, Adriaenssens J, Dilles T, Martens D, Van Rompaey B, Timmermans O. Impact of role-, job- and organizational characteristics on nursing unit managers' work related stress and well-being. *J Adv Nurs* 2014;70:2622–2633.
- Gatti P, Ghislieri C, Cortese CG. Relationships between followers' behaviors and job satisfaction in a sample of nurses. *PLoS One* 2017;12:e0185905.
- Keith AC, Washawsky N, Neff D, Loerzel V, Parchment J. Factors that influence nurse manager job satisfaction: an integrated literature review. *J Nurs Manag* 2021;29:373–384.
- Gillet N, Colombat P, Michinov E, Pronost AM, Fouquereau E. Procedural justice, supervisor autonomy support, work satisfaction, organizational identification and job performance: the mediating role of need satisfaction and perceived organizational support. *J Adv Nurs* 2013;69:2560–2571.
- Dempsey J. Nurses values, attitudes and behaviour related to falls prevention. *J Clin Nurs* 2009;18:838–848.
- HR engagement and transformation platform—Wittyfit [Internet]. 2021. Available at: <https://wittyfit.com/en/>. Accessed July 5, 2021.
- Dutheil F, Duclos M, Naughton G, et al. Wittyfit-live your work differently: study protocol for a workplace-delivered health promotion. *JMIR Res Protoc* 2017;6:e26267.
- Lesage FX, Berjot S, Deschamps F. Clinical stress assessment using a visual analogue scale. *Occup Med (Lond)* 2012;62:600–605.
- Dutheil F, Trousselard M, Perrier C, et al. Urinary interleukin-8 is a biomarker of stress in emergency physicians, especially with advancing age—the JOBSTRESS\* randomized trial. *PLoS One* 2013;8:e71658.
- Dutheil F, Pereira B, Moustafa F, Naughton G, Lesage FX, Lambert C. At-risk and intervention thresholds of occupational stress using a visual analogue scale. *PLoS One* 2017;12:e0178948.
- Little RJA, Rubin DB. *Statistical Analysis With Missing Data*. 2nd ed. New York: John Wiley & Sons; 2002.
- Li P, Stuart EA, Allison DB. Multiple imputation: a flexible tool for handling missing data. *JAMA* 2015;314:1966–1967.
- Ben-Shachar MS, Lüdtke D, Makowski D. Effectsize: estimation of effect size indices and standardized parameters. *J Open Source Softw* 2020;5:2815.
- Funder DC, Ozer DJ. Evaluating effect size in psychological research: sense and nonsense. *Adv Methods Pract Psychol Sci* 2019;2:156–168.
- Chen H, Cohen P, Chen S. How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Commun Stat* 2010;39:860–864.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2021.
- Birken SA, Lee SY, Weiner BJ. Uncovering middle managers' role in healthcare innovation implementation. *Implement Sci* 2012;7:28.
- Birken SA, DiMartino LD, Kirk MA, Lee SY, McClelland M, Albert NM. Elaborating on theory with middle managers' experience implementing healthcare innovations in practice. *Implementation Sci* 2016;11:2.

46. Bianchi M, Bagnasco A, Bressan V, et al. A review of the role of nurse leadership in promoting and sustaining evidence-based practice. *J Nurs Manag* 2018;26:918–932.
47. Moon SE, Van Dam PJ, Kitsos A. Measuring transformational leadership in establishing nursing care excellence. *Healthcare (Basel)* 2019;7:132.
48. Tang JHC, Hudson P. Evidence-based practice guideline: nurse retention for nurse managers. *J Gerontol Nurs* 2019;45:11–19.
49. Cheung K, Ching SSY. Job satisfaction among nursing personnel in Hong Kong: a questionnaire survey. *J Nurs Manag* 2014;22:664–675.
50. Kane-Urrabazo C. Management's role in shaping organizational culture. *J Nurs Manag* 2006;14:188–194.
51. Colin-Chevalier R, Pereira B, Benson AC, Dewavrin S, Cornet T, Dutheil F. The protective role of job control/autonomy on mental strain of managers: a cross-sectional study among Wittyfit's users. *Int J Environ Res Public Health* 2022;19:2153.
52. Kim KW, Cho YH. The moderating effect of managerial roles on job stress and satisfaction by employees' employment type. *Int J Environ Res Public Health* 2020;17:E8259.
53. Zhang Z, Shi G, Li L, Bian Y. Job satisfaction among primary care physicians in western China. *BMC Fam Pract* 2020;21:219.
54. Allen TD, Herst DE, Bruck CS, Sutton M. Consequences associated with work-to-family conflict: a review and agenda for future research. *J Occup Health Psychol* 2000;5:278–308.
55. Lee S. A concept analysis of “meaning in work” and its implications for nursing. *J Adv Nurs* 2015;71:2258–2267.
56. Peng J, Li D, Zhang Z, et al. How can core self-evaluations influence job burnout? The key roles of organizational commitment and job satisfaction. *J Health Psychol* 2016;21:50–59.
57. Kaya A, Boz İ. The development of the professional values model in nursing. *Nurs Ethics* 2019;26:914–923.
58. Dobrow Riza S, Ganzach Y, Liu Y. Time and job satisfaction: a longitudinal study of the differential roles of age and tenure. *J Manag* 2018;44:2558–2579.
59. Anser MK, Ali M, Anwar F, Usman M. Subjective age and job satisfaction: a moderated mediation model of job burnout and chronological age. *Front Public Health* 2020;8:62.
60. Clark AE, Oswald AJ. Satisfaction and comparison income. *J Public Econ* 1996;61:359–381.
61. Carvajal MI, Popovici I, Hardigan PC. Gender and age variations in pharmacists' job satisfaction in the United States. *Pharmacy (Basel)* 2019;7:46.
62. Saperstein AK, Viera AJ, Fimhaber GC. Mentorship and job satisfaction among Navy family physicians. *Mil Med* 2012;177:883–8, 888.
63. Miao Y, Li L, Bian Y. Gender differences in job quality and job satisfaction among doctors in rural western China. *BMC Health Serv Res* 2017;17:848.
64. Weiss DJ, Dawis RV, England GW. Manual for the Minnesota Satisfaction Questionnaire. *Minn Stud Vocational Rehabil* 1967;22:120–120.
65. Devaux M, Sassi F. Social disparities in hazardous alcohol use: self-report bias may lead to incorrect estimates. *Eur J Public Health* 2016;26:129–134.



## Article n°4

Stress and job satisfaction over time,  
the influence of the managerial position:  
a bivariate longitudinal modelling of  
Wittyfit data

## **Article n°4**

Stress and job satisfaction over time, the influence of the managerial position: a bivariate longitudinal modelling of Wittyfit data

Colin-Chevalier R, Dutheil F, Benson AC, Dewavrin S, Cornet T, Lambert C, Pereira B

### **Résumé**

Cet article marque la conclusion de la série d'études portant sur les différences de perception du travail entre managers et employés. Dans cet article, nous souhaitons étudier les disparités annuelles de stress et de satisfaction au travail sur quatre années en utilisant des modèles bivariés longitudinaux. Par passage d'un cadre transversal à un cadre longitudinal, de nombreux défis méthodologiques supplémentaires se sont posés à nous, comme par exemple la nécessité de tenir compte de l'« effet individu » comme effet de cluster en plus de l'« effet entreprise ». Le recours à des modèles longitudinaux nous a également conduit à valider certaines hypothèses. Ainsi, les critères de Nagin ont été respectés pour garantir la validité des trajectoires latentes identifiés, et les hypothèses de stabilité et de synchronicité du modèle de médiation par panel à décalage croisé ont été validées. Enfin, en raison d'un phénomène de données manquantes dont la nature a été identifiée comme complexe, nous avons jugé pertinent de conduire des analyses de sensibilité tout au long de nos analyses, garantissant la robustesse de nos résultats. Dès lors, dans un premier temps, nous avons observés au moyen de modèle mixtes, transversaux et longitudinaux, les différences de stress et de satisfaction perçues selon la position managériale des travailleurs. En moyenne, si leur niveau de stress n'était pas différent, les managers avaient une satisfaction plus élevée que les salariés. Dans un second temps, en utilisant un modèle de trajectoire basé sur le groupe sur l'ensemble de la population, nous avons mis en évidence trois groupes de travailleurs ayant suivi des évolutions différentes de stress et de satisfaction au travail dans le temps. Nous avons pu constater que la proportion de managers était plus élevée dans le groupe qui présentait le plus haut niveau moyen de

satisfaction au travail, et inversement. Enfin, dans un troisième et dernier temps, nous avons analysé les influences croisées des deux indicateurs dans le temps en ayant recours à un modèle de médiation par panel à décalage croisé. Au cours du temps, les deux indicateurs avaient tendance à s'influencer l'un et l'autre. Par exemple, si la satisfaction d'un travailleur avait augmenté d'une année à l'autre, son stress avait baissé. Garantir que tous les travailleurs soient satisfaits, en particulier les employés qui présentent des niveaux plus faibles, est un moyen pour les entreprises de diminuer leur stress, ce qui contribue à améliorer leur QVT.

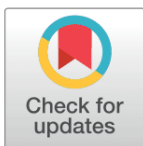
## RESEARCH ARTICLE

# Stress and job satisfaction over time, the influence of the managerial position: A bivariate longitudinal modelling of Wittyfit data

Rémi Colin-Chevalier<sup>1\*</sup>, Frédéric Dutheil<sup>1</sup>, Amanda Clare Benson<sup>2</sup>, Samuel Dewavrin<sup>3</sup>, Thomas Cornet<sup>3</sup>, Céline Lambert<sup>4</sup>, Bruno Pereira<sup>4</sup>

**1** Université Clermont Auvergne, CNRS, LaPSCo, CHU Clermont-Ferrand, Cegid, Clermont-Ferrand, France, **2** Sport Innovation Research Group, Department of Health and Biostatistics, Swinburne University of Technology, Hawthorn, Victoria, Australia, **3** Cegid, Lyon, France, **4** Biostatistics Unit, DRCl, CHU Clermont-Ferrand, Clermont-Ferrand, France

\* [r.colin6374@gmail.com](mailto:r.colin6374@gmail.com)



## OPEN ACCESS

**Citation:** Colin-Chevalier R, Dutheil F, Benson AC, Dewavrin S, Cornet T, Lambert C, et al. (2024) Stress and job satisfaction over time, the influence of the managerial position: A bivariate longitudinal modelling of Wittyfit data. PLoS ONE 19(3): e0298126. <https://doi.org/10.1371/journal.pone.0298126>

**Editor:** Prakash K.C., Tampere University, FINLAND

**Received:** May 23, 2023

**Accepted:** January 19, 2024

**Published:** March 4, 2024

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0298126>

**Copyright:** © 2024 Colin-Chevalier et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its [Supporting Information files](#).

## Abstract

### Background

The managerial position affects stress and job satisfaction of workers, but these influences have always been studied separately.

### Objective

We aimed to assess bivariate influence of the managerial position on workers' stress and job satisfaction and the inter-relationship of these indicators over time.

### Methods

We have analyzed data from workers who use the Wittyfit software, collected annually between 2018 and 2021. Stress and job satisfaction were evaluated by self-report questionnaires. Job position (manager or employee) was provided by the software's client companies.

### Results

Data of 704 workers were included in the study. Cross-sectional and longitudinal multivariate analyses revealed that managerial position improves job satisfaction ( $p < 0.001$ ), but not stress ( $p = 0.4$ ). Overall, while workers' job satisfaction has improved ( $p < 0.001$ ), stress has remained stable over time ( $p = 0.3$ ). Three latent groups, with specific evolutionary multi-trajectory of stress and job satisfaction were identified in the sample (entropy = 0.80). Age and seniority, but not gender tended to influence managers' and employees' indicators. Over time, stress and job satisfaction have tended to negatively interconnect, in cross-section and in a cross-lagged manner ( $p < 0.001$ ).

**Funding:** The authors received no specific funding for this work.

**Competing interests:** RC, SD, and TC are part of Wittyfit. Other authors have declared that no competing interests exist (FD is responsible for the scientific accuracy of Wittyfit but is not paid by Cegid; as previously published, Wittyfit is a public private partnership with the CHU of Clermont-Ferrand, France). This does not alter our adherence to PLOS ONE policies on sharing data and materials.

## Conclusions

The managerial position improves workers' job satisfaction but has no effect on stress. Sociodemographics including age and seniority, but not gender, can affect this relationship. Stress and job satisfaction can influence each other, both cross-sectionally and over time. To be more effective, organizations should implement holistic strategies targeting multiple indicators.

## Trial registration

Clinicaltrials.gov: [NCT02596737](https://clinicaltrials.gov/ct2/show/study/NCT02596737).

## Introduction

Stress refers to the inability of an organism to face a threat, whether physical or emotional [1], and its physiological and pathophysiological response [2]. Known to increase the risk of cardiovascular disease [3,4], multiple factors, mostly work-related, have an influence on stress [5,6]. Stress has become a major problem for organizations [7]. Job satisfaction is a state of balance between a worker's expectations for his or her job and the professional values that the job enables him or her to achieve [8]. As an emerging and predictive factor of good occupational health [9], several studies have highlighted its main work-related drivers [10,11]. In particular, stress and job satisfaction were found to have a negative influence on each other, in other words, stress is a factor of job dissatisfaction [12–14].

The managerial position was found to be a common factor in stress [15] and job satisfaction [16], as were sociodemographic factors such as age, seniority and gender [16–18]. To the best of our knowledge, no evaluation of the putative short- and long-term influences of managerial status on occupational health indicators such as stress and job satisfaction have been conducted yet. The managerial role being associated with greater duties and responsibilities [19], but also more risk for health [20,21], it seems relevant to understand whether managers may be affected differently from employees, how this persists over time, and how age, seniority or gender may affect workers' stress and job satisfaction depending on their job position.

Thus, the main objective of our study was to evaluate the influence of the job position, notably the managerial role, on stress and job satisfaction. The secondary objectives were to explore the simultaneous evolution over time of the two work-related indicators and to assess their reciprocal influences, in all workers and as a function of the job position (manager or employee), and to measure the influence of age, seniority and gender on workers' stress and job satisfaction according to their job position.

## Materials and methods

### Recruitment

Wittyfit is a digital web software designed by Cegid (Lyon, France) in partnership with the university hospital of Clermont-Ferrand, implemented in numerous French companies from any industry sector [22]. Wittyfit enables the well-being of a company's worker to be assessed by collecting data on their feelings to improve the collective performance. Based on voluntary participation, the software offers its users the opportunity to express themselves anonymously by answering various health-related questionnaires to evaluate their overall health status and provides personalized feedback. This study received approval from both the National

Commission for Data Protection and Liberties (CNIL) and the ethics committee of South-East VI (identified on [clinicaltrials.gov](https://clinicaltrials.gov) with the registration number NCT02596737).

## Material

**Stress and job satisfaction.** During the data collection period, users of the Wittyfit software could assess for the first time or update their stress and job satisfaction levels using the relative analog scales, as often as they wished [23–25]. The software's visual analog scales go from 0 (low level) to 100 (high level). For both parameters, the average annual score of a worker was computed and retained so that each worker was defined by only two records per year, one for stress and one for job satisfaction. If a worker did not complete a questionnaire for a full year, it was considered a missing value.

**Time.** Time was defined as the year of the measurements, ranging from 2018 to 2021, taken in their entirety. The first year of the study, 2018, was considered as the reference point for comparing changes in stress and job satisfaction in the longitudinal analyses.

**Job position and other sociodemographics.** Wittyfit collects sociodemographic data from its users. This includes information about their job position, which can only have two states: manager or employee. Managers can be defined as workers in charge of an organization, with one or more individuals under their supervision, namely the employees. In addition to job position, Wittyfit also collects the age, seniority, gender (male or female) and company of its users. Age has been transformed into a two-modality class depending on whether the worker was above or below the age of 40-years old. Similarly, seniority has been transformed into a two-modality class depending on whether the worker was in a position of seniority for greater or less than five years. These thresholds were chosen in accordance with the median of these indicators in the sample.

**Inclusion/exclusion criteria.** Workers from multiple French companies using Wittyfit were included in the study. Data collection began in 2018 and continued through 2022 (excluded). All companies that were not yet Wittyfit customers in 2018 have been excluded, as well as those with no managers, as it was impossible to measure the effect of job position in these companies. To be included, a worker had to have at least one record between stress and job satisfaction in time one. Workers with two or fewer annual records of stress or job satisfaction were excluded because it is impossible to estimate individual change with too few measures [26]. Users with no sociodemographic data were excluded.

## Statistical analysis

Statistical analyses were performed using R (version 4.2.1) and Stata (version 11.2) software. Quantitative data (stress and job satisfaction per year), expressed as mean±SD, were compared between managers and employees (i) in cross-section using first a univariate and then a bivariate linear mixed model with only company effect as random and (ii) longitudinally, still bivariate, adding individual and time random effects. Subgroup analyses were conducted to assess the influence of other sociodemographic data on both outcomes by job position. Then, longitudinal analysis was conducted using multivariate generalized linear mixed models to evaluate the workers' average changes in stress and job satisfaction over time. Afterwards, the same analysis was conducted for each job position, and stress and job satisfaction were compared at each time. A group-based multi-trajectory model [27], using the “traj, multgroups” function of the “traj” package of Stata, was used to highlight latent groups within the sample, according to their changes in stress and job satisfaction. Model selection has been made by following multiple rules: (1) each trajectory had to have an average posterior probability of assignment greater than 70%, (2) a probability of correct classification greater than five, (3) include more than 5%



of individuals in the total sample, and (4) have an entropy greater than or equal to 0.80 [28,29]. The model was fitted for several classes ranging from one to 10, and the one with the largest number of groups and which complied with the rules was selected. Individuals were then assigned to the group to which they were most likely to belong, that is the group whose trajectory was most like their own individual trajectory. A multinomial logistic mixed model with company effect as random enabled comparison of the proportion of managers and employees into the different groups. Finally, we used a cross-lagged panel model to assess the simultaneous relationships between stress and job satisfaction over time, among all workers and then by job position.

All models' estimates were turned into Hedges'  $g$  (also called  $g$ ) effect sizes [30] and 95% confidence intervals (95CI) and interpreted according to Cohen's rules [31]. Multicollinearity of outcomes was checked prior to analysis. Linear model assumptions (residual independence and normality, and variance homogeneity) were verified post hoc. Missing data phenomenon has occurred over the course of the study. As the type of missing data was identified as not missing at random, and the structure of missing data as non-monotonic, a sensitivity analysis was conducted. Five different datasets were used to assess the sensitivity of the results: (i) an available-cases dataset, (ii) a complete-cases dataset (i.e., the dataset that includes workers who reported their stress and job satisfaction levels each year), two single-imputed datasets imputed with (iii) the linear interpolation and (iv) the "last observation carried forward" methods, and (v) a multiple-imputed dataset. Single imputations were performed with the command "imputation" of the "longitudinalData" package of R. Multiple imputation were performed with the "mice" command of the eponymous package of R. Only estimates ( $g$  and 95CI) obtained with the original dataset (the dataset of available cases) were included. Unless specified, a  $p$ -value  $< 0.05$  was considered as statistically significant.

Further information on the conduct of the study and the quality of reporting was provided in the distribution of outcomes (S1 Appendix), the statistical analysis plan (S2 Appendix), the sensitivity analysis results (S3 Appendix), and the reporting guidelines (S4 Appendix).

## Results

### Participants

Data were collected from 704 workers (362 men, 342 women), with 381 workers over 40, and 434 with more than 5 years' seniority at the start of the study. Workers included 81 managers and 623 employees from three different companies, with 143, 259 and 302 workers (Fig 1). The three companies belonged to commercial activities and services sector.

Table 1 shows workers' stress and job satisfaction over time by job position with the difference between the two populations, and the distribution of missing values. Additional information on the distribution of variables can be found in Multimedia Appendix 1 (Fig S1 in S1 Appendix, Table S1 in S1 Appendix).

### Influence of job position on average stress and job satisfaction

Univariate analyses revealed that stress was not different between job position ( $g = 0.13$ , 95CI  $-0.10$  to  $0.36$ ,  $p = 0.3$ ), but that job satisfaction was lower for employees than for managers ( $-0.65$ ,  $-0.89$  to  $-0.42$ ,  $p < 0.001$ ) (Fig 2). In managers, neither age, seniority nor gender influenced the two indicators. Employees with more than five years of seniority had higher stress ( $0.26$ ,  $0.10$  to  $0.43$ ,  $p = 0.004$ ) and lower job satisfaction ( $-0.33$ ,  $-0.50$  to  $-0.16$ ,  $p = 0.002$ ), while age and gender did not show any influence on either indicator. The sensitivity analysis confirmed these results (S3 Table 1 in S3 Appendix, S3 Table 2 in S3 Appendix).



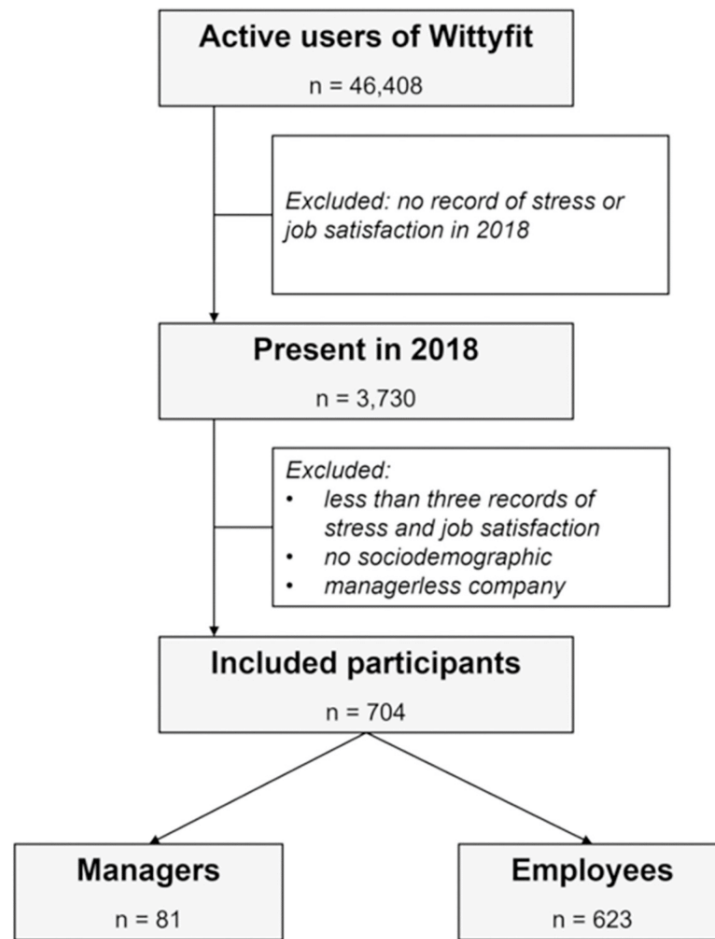


Fig 1. Flowchart of the study. Legend: ‘n’: sample size.

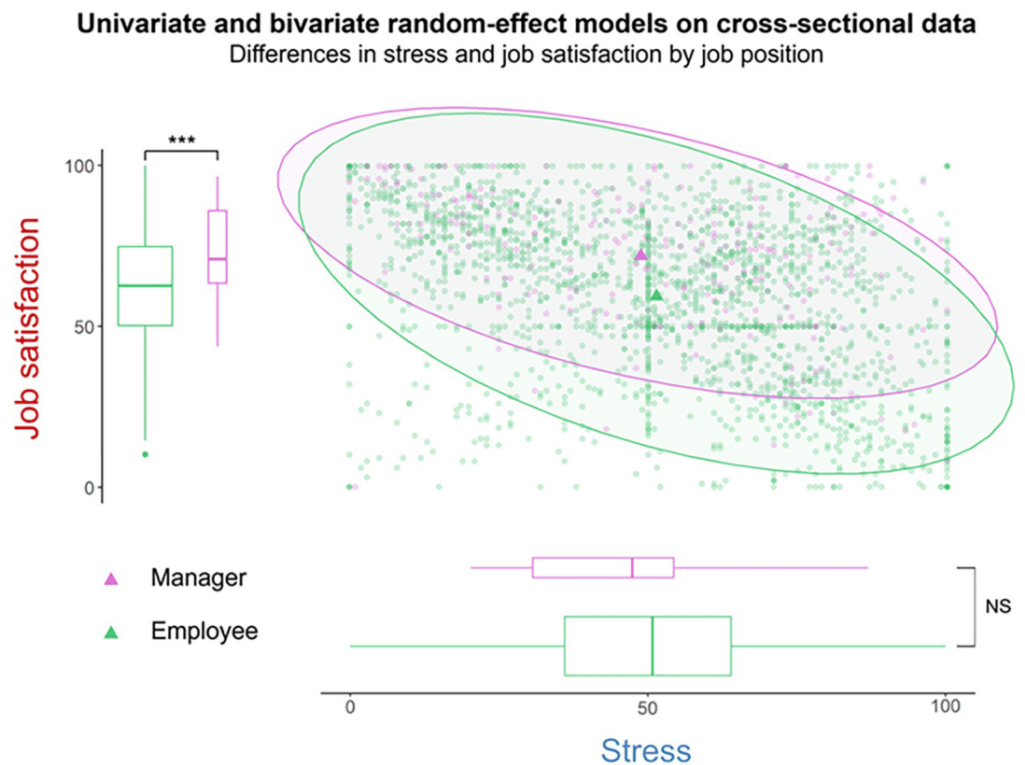
<https://doi.org/10.1371/journal.pone.0298126.g001>

Table 1. Stress and job satisfaction by job position over time.

Outcome (by time)	Manager Mean±SD (NA/n)	Employee Mean±SD (NA/n)	Difference between groups (g and 95CI)
Stress (/100)			
2018	53.7±26.3 (1/81)	51.6±27.7 (8/623)	-0.05 (-0.61 to 0.32)
2019	49.3±23.5 (17/81)	52.8±27.8 (140/623)	0.06 (-0.44 to 0.83)
2020	45.9±29.0 (24/81)	50.5±25.2 (157/623)	0.13 (-0.19 to 1.09)
2021	44.2±27.4 (20/81)	50.7±26.7 (149/623)	0.17 (-0.04 to 1.24)
Job satisfaction (/100)			
2018	69.3±22.0 (0/81)	57.5±25.2 (1/623)	-0.28 (-1.26 to -0.46)
2019	70.4±19.5 (16/81)	58.3±26.7 (122/623)	-0.32 (-1.55 to -0.46)
2020	72.0±24.4 (24/81)	61.5±25.2 (158/623)	-0.29 (-1.50 to -0.32)
2021	76.6±18.5 (19/81)	61.2±26.3 (148/623)	-0.42 (-1.87 to -0.82)

Legend: Difference in stress and job satisfaction are calculated between managers and employees at each time using univariate linear random-effect model with “company” effect as random. A bold result means that the difference is significant in terms of effect size. ‘SD’: standard deviation, ‘NA’: missing data (not available), ‘n’: group size (number), ‘g’: Hedges’ g effect size, ‘95CI’: 95% confidence interval.

<https://doi.org/10.1371/journal.pone.0298126.t001>



The bivariate model revealed that managerial position can influence workers' job satisfaction, but not their stress. Effect size measuring the difference in job satisfaction by job position, calculated using Hedge's  $g$ , was found to be moderate (Hedges'  $g = -0.65$ , 95CI  $-0.89$  to  $-0.42$ ,  $P < .001$ ). No differences in mean stress level were found between groups ( $0.13$ ,  $-0.10$  to  $0.36$ ,  $P = .3$ ).

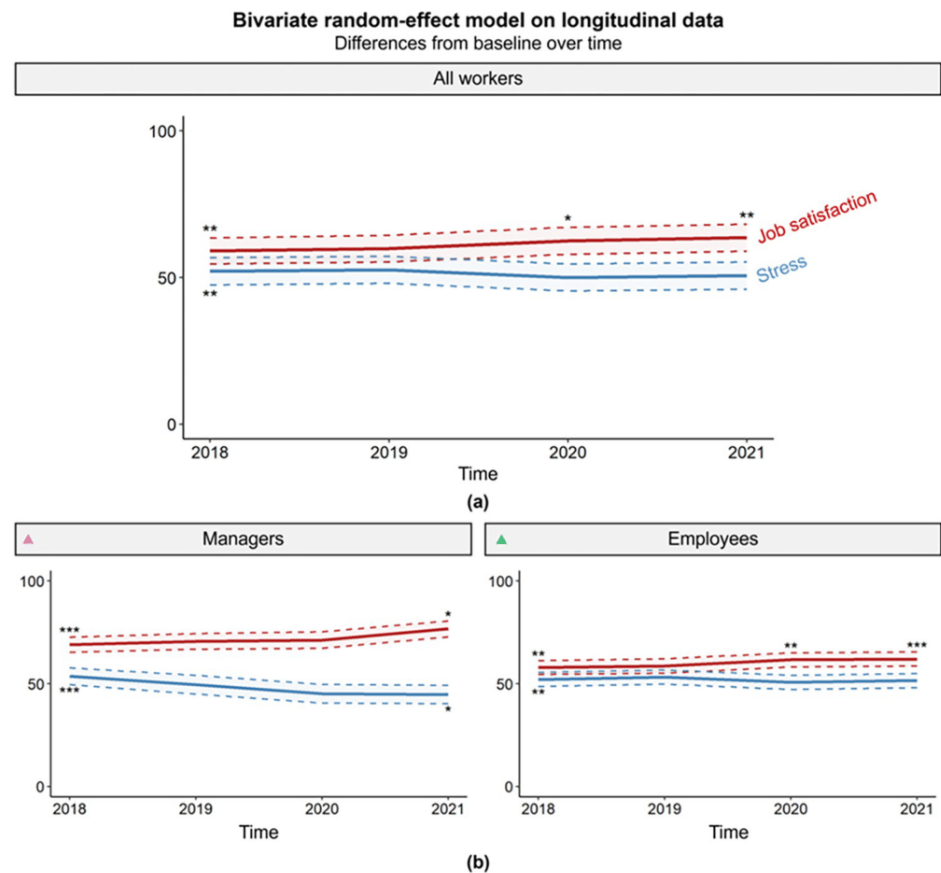
**Fig 2. Stress and job satisfaction by job position across all years (cross-sectional univariate and bivariate analyses).**

<https://doi.org/10.1371/journal.pone.0298126.g002>

Overall, bivariate cross-sectional analysis revealed that the job satisfaction was lower in employees ( $-0.85$ ,  $-1.10$  to  $-0.62$ ,  $p < 0.001$ ), and stress not different ( $0.14$ ,  $-0.11$  to  $0.36$ ,  $p = 0.2$ ). Age was found to reduce job satisfaction of managers ( $-0.44$ ,  $-0.89$  to  $-0.01$ ,  $p = 0.05$ ), but not seniority or gender. In employees, after cross-sectional analysis, workers with greater than five-years seniority were found to be more stressed ( $0.34$ ,  $0.18$  to  $0.52$ ,  $p < 0.001$ ) and less satisfied ( $-0.47$ ,  $-0.63$  to  $-0.30$ ,  $p < 0.001$ ). Age and gender had no effect on indicators. The results of the sensitivity analyses, except for some obtained with the complete cases analysis were in line with these results (S3 Table 3 in S3 Appendix).

### Average changes in stress and job position over time

Over time, while workers' stress has not changed ( $p = 0.3$ ), their job satisfaction did. After one year of stability, it improved slightly in 2020 ( $g = 0.13$ , 95CI  $0.02$  to  $0.24$ ,  $p = 0.02$ ) and 2021 ( $0.17$ ,  $0.07$  to  $0.28$ ,  $p = 0.002$ ) (Fig 3A). In managers, while there was no change in 2019 and 2020, stress was found to be lower in 2021 ( $-0.31$ ,  $-0.63$  to  $-0.01$ ,  $p = 0.05$ ), and job satisfaction higher ( $0.34$ ,  $0.05$  to  $0.66$ ,  $p = 0.03$ ). Only the results obtained with the multiple-imputed dataset can confirm these results. For the entire sample, employees' job satisfaction improved very slightly in 2020 ( $0.14$ ,  $0.04$  to  $0.27$ ,  $p = 0.006$ ) and 2021 ( $0.16$ ,  $0.05$  to  $0.29$ ,  $p < 0.001$ ) (Fig 3B).



Estimates and confidence intervals were computed from longitudinal multivariate generalized linear mixed models. Baseline levels (2018) and differences from baseline (2019 to 2021), were measured in Hedges'  $g$  effect size and 95% confidence interval (Legend: \*\*:  $P < 0.05$ , \*\*\*:  $P < 0.01$ , \*\*\*\*:  $P < 0.001$ ).

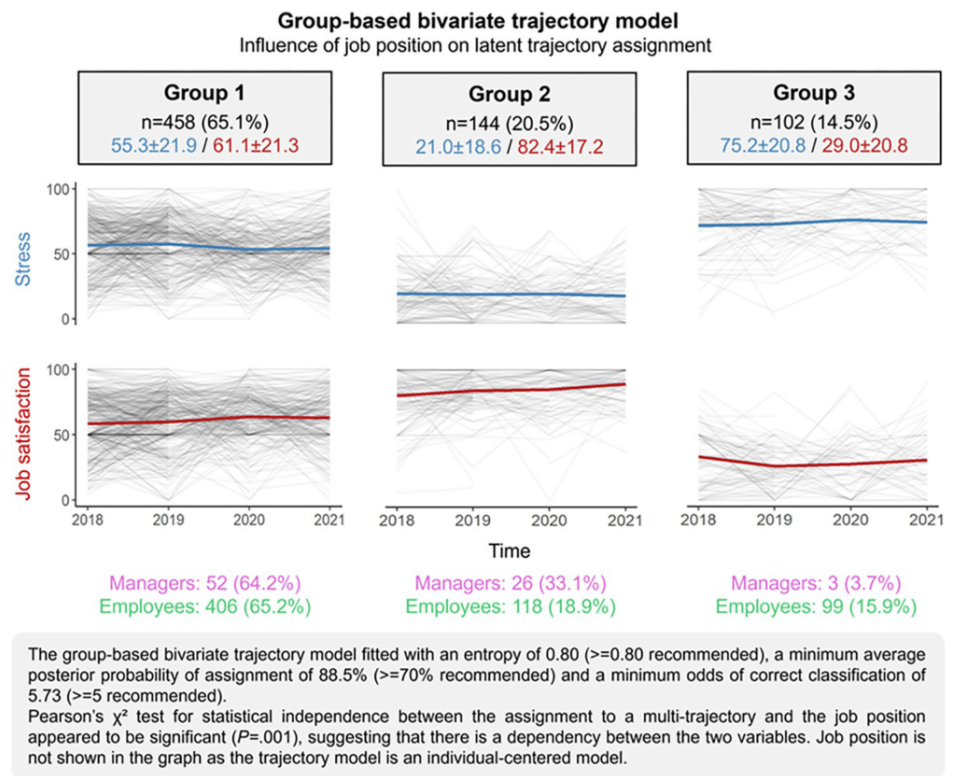
On average, while stress did not change over time, job satisfaction has very slightly increased in all workers in 2020 (Hedges'  $g = 0.13$ , 95CI 0.02 to 0.24,  $P = .02$ ) and 2021 (0.17, 0.07 to 0.28,  $P = .002$ ) (a). The subgroup analysis revealed that stress and job satisfaction developed differently according to job position, especially among managers. Differences from baseline were only found in 2021, namely a reduction in stress ( $-0.31$ ,  $-0.63$  to  $-0.01$ ,  $P = .05$ ) and an increase in job satisfaction (0.34, 0.05 to 0.66,  $P = .03$ ) (b).

**Fig 3. Evolution of stress and job satisfaction over time by job position (longitudinal bivariate analysis).** Legend: (a) for all workers, (b) by job position.

<https://doi.org/10.1371/journal.pone.0298126.g003>

Furthermore, the comparison of average changes by job position revealed that stress was not different in any year ( $p = 0.3$ ), but that job satisfaction for employees remained lower than for managers over time ( $p < 0.001$ ). Indeed, a small difference was found in 2020 ( $-0.47$ ,  $-0.79$  to  $-0.14$ ,  $p = 0.002$ ), and a moderate difference in 2018 ( $-0.68$ ,  $-0.99$  to  $-0.38$ ,  $p < 0.001$ ), 2019 ( $-0.63$ ,  $-0.95$  to  $-0.32$ ,  $p < 0.001$ ) and 2021 ( $-0.74$ ,  $-1.05$  to  $-0.41$ ,  $p < 0.001$ ). The results of the sensitivity analyses, except for some obtained with the complete cases analysis, were in line with these results (S3 Table 4 in S3 Appendix).

Managers over 40 reported being more stressed ( $g = 0.45$ , 95CI 0.02 to 0.88,  $p = 0.04$ ) and less satisfied ( $-0.44$ ,  $-0.90$  to  $-0.02$ ,  $p = 0.04$ ) with their job. Employees with more than five years' seniority were also found to be more stressed (0.36, 0.20 to 0.52,  $p < 0.001$ ) and less satisfied ( $-0.46$ ,  $-0.62$  to  $-0.28$ ,  $p < 0.001$ ). Overall, the results of the sensitivity analyses, except for some obtained with the complete cases analysis, confirmed these results (S3 Table 5 in S3 Appendix).



**Fig 4. Group-based multi-trajectories of workers' stress and job satisfaction.**

<https://doi.org/10.1371/journal.pone.0298126.g004>

### Individual multi-trajectory of stress and job satisfaction among workers

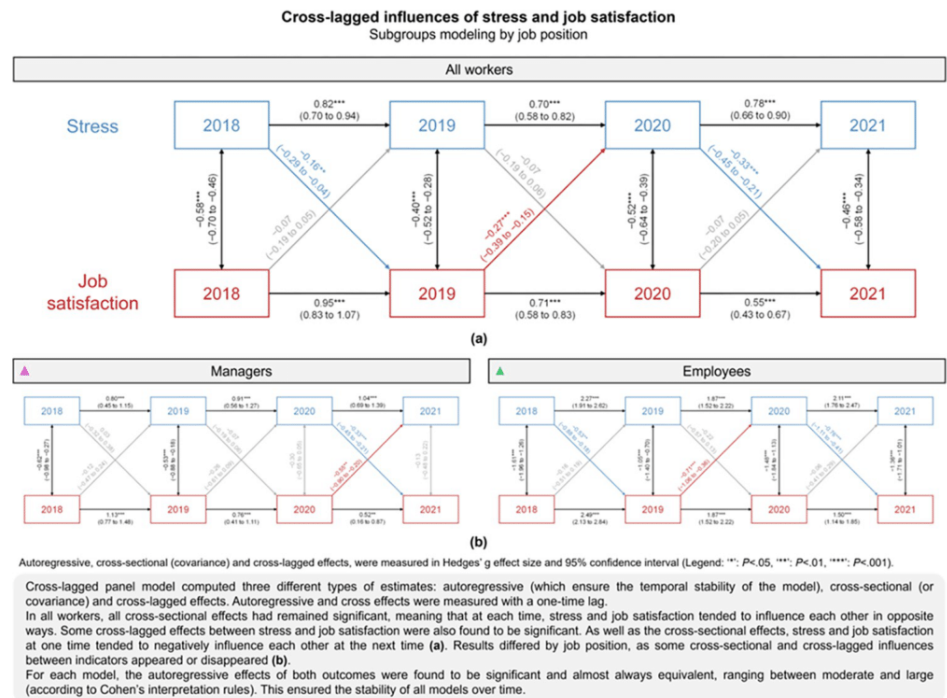
Three latent groups of workers were found in the global sample (entropy = 0.80) (Fig 4). The results of the sensitivity analysis, except for the complete-cases analysis, confirmed this optimal number of latent groups (S3 Table 6 in S3 Appendix). Group 1 ( $n = 458$  individuals, 65.1% of the sample) included workers whose stress ( $55.3 \pm 21.9$ ) and job satisfaction ( $61.1 \pm 21.3$ ) were moderate over time, on average, with a slight reduction in stress in 2020 ( $g = -0.19$ , 95CI  $-0.33$  to  $-0.05$ ,  $p = 0.008$ ), and an increase in job satisfaction in 2020 ( $0.26$ ,  $0.11$  to  $0.39$ ,  $p < 0.001$ ) and 2021 ( $0.25$ ,  $0.11$  to  $0.40$ ,  $p = 0.002$ ) compared to 2018. Group 2 ( $n = 144$ , 20.5%) included workers whose stress remained low and stable over time ( $21.0 \pm 18.6$ ) and job satisfaction was high ( $82.4 \pm 17.2$ ), on average, with an increase in 2021 ( $0.40$ ,  $0.13$  to  $0.63$ ,  $p = 0.004$ ). Group 3 ( $n = 102$ , 14.5%), which included workers whose stress was high over time ( $75.2 \pm 20.8$ ) while job satisfaction was low ( $29.0 \pm 20.8$ ), on average, with a decrease in 2019 ( $-0.45$ ,  $-0.75$  to  $-0.16$ ,  $p < 0.001$ ) and 2020 ( $-0.29$ ,  $-0.58$  to  $0.00$ ,  $p = 0.05$ ).

The proportion of employees and managers in each latent group was found to be different ( $p < 0.001$ ). Indeed, although it seemed equivalent in group 1 (managers 64.2% vs. employees 65.2%), the proportion of managers appeared to be higher than that of employees in group 2 (33.1% vs. 18.9%), and lower in group 3 (3.7% vs. 15.9%). The results of sensitivity analysis confirmed these results (S3 Table 6 in S3 Appendix).

### Cross-lagged effects of stress and job satisfaction by job position

Overall, stress and job satisfaction of workers influenced each other over time ( $p < 0.001$ ). The stress of workers in 2018 seemed to have slightly reduced their job satisfaction in 2019 ( $g = -$





**Fig 5. Cross-lagged effects of stress and job satisfaction by job position.** Legend: (a) for all workers, (b) by job position.

<https://doi.org/10.1371/journal.pone.0298126.g005>

0.16, 95CI -0.29 to -0.04,  $p = 0.006$ ). This was also found from 2020 to 2021 (-0.33, -0.45 to -0.21,  $p < 0.001$ ). Simultaneously, workers' job satisfaction did slightly reduce their stress from 2019 to 2020 (-0.27, -0.39 to -0.15,  $p < 0.001$ ). In addition to cross-lagged effects, all cross-sectional influences between stress and job satisfaction were found to be significant and negative ( $p < 0.001$ ) (Fig 5A). The sensitivity analysis confirmed these results (S3 Table 7 in S3 Appendix).

When analyzing by job position, similar results were found in employees, but not in managers. Indeed, in managers, only the last two cross-lagged effects were significant, namely that of stress and job satisfaction (-0.33, -0.45 to -0.21,  $p < 0.001$ ), and conversely (-0.55, -0.90 to -0.20,  $p = 0.002$ ). Moreover, only the first two covariances between stress and job satisfaction appeared to be significant (2018: -0.62, -0.98 to -0.27,  $p < 0.001$ ; 2019: -0.53, -0.88 to -0.18,  $p = 0.003$ ) (Fig 5B). The results of the sensitivity analyses confirmed these results (S3 Table 7 in S3 Appendix).

## Discussion

The main results showed that the managerial position was mainly a factor in improving job satisfaction among workers, with almost no effect on stress. Socio-demographic factors such as age and seniority may have influenced these two indicators, but not gender. Overall, workers' job satisfaction improved and in parallel, their stress did not change, even though the two indicators had a negative influence on each other over time, however, these results were different according to job position.

## Stress, job satisfaction and the managerial role

Our findings suggest that the managerial position is mainly a factor in improving job satisfaction, but not stress. This suggests that the responsibilities associated with the position do not

add stress for managers. A study among Norwegian employees found that the job strain risk, which can cause stress to workers [32,33], was not different between regular employees and workers with managerial roles [34]. Despite higher demands, greater control protects managers from this risk [18,35]. A study in nurses showed that managers may face higher occupational stressors [36]. Like previous research, the role within the working environment makes managers more satisfied with their work than employees. A study in the emergency department health professionals highlighted that job position played a role in workers' holistic satisfaction [16]. Another study in nursing personnel also found that job satisfaction may differ by job position, with nurse managers having the highest job satisfaction of all groups [37]. Organizations should strive to enhance the job satisfaction of their workers, especially employees, as a factor in job well-being [12,38–40].

### **Mean changes and group-based multi-trajectories of stress and job satisfaction over time**

On average, while their stress remained stable, the job satisfaction of workers was above the baseline in the last two years of the study. While employees followed the same trend, we found that managers experienced different changes for both outcomes. The last year, at the same time as their job satisfaction improved compared to baseline, their stress also decreased. Three latent groups with specific evolutionary trajectories of stress and job satisfaction were identified in the sample. Groups can be named according to the shape of their associated multi-trajectory as follows: the average (group 1), the good (group 2) and the poor (group 3) occupational health groups. The proportion of managers and employees was different across groups, with notably a higher proportion within the good occupational health group. These two results suggest that the managerial position may act as a long-term factor in reducing stress and improving job satisfaction. Further analyses of the average changes in managers and the individual multi-trajectories of group 2 and group 3 especially, including the use of static covariates, could reveal the protective and risk factors associated with good occupational health of workers [41].

### **Concurrent influences between stress and job satisfaction**

The negative reciprocal influences between stress and job satisfaction have been studied and identified many times, both transversally [11,12,14] and longitudinally [13]. In our study, in addition to cross-sectional reciprocal influences, cross-lagged influences were also highlighted. This suggests that stress and job satisfaction can influence each indicator over time. For example, a satisfied worker might experience a reduction in stress because of the benefits of good job satisfaction [6]. In contrast, a stressed worker may experience a decrease in job satisfaction [10,11]. Among managers, we found that indicators were interconnected in the last two years. This could explain the changes observed on average in this population, where stress and job satisfaction have evolved in opposite directions over the final year. These results are similar to a previous study in nurses [42]. This suggests that to improve occupational health through stress and job satisfaction, organizations should strive to implement holistic interventions to make them more effective.

### **Sociodemographic factors of occupational health among managers and employees**

Sociodemographics factors were found to influence managers' and employee's indicators. The results of the longitudinal analysis suggest that managers over 40 years old had poorer

occupational health, reporting higher stress and lower job satisfaction than others. This finding is consistent with literature that the risk of stress is greater among older workers [43], and that age and job satisfaction are negatively related [44]. The same result was found for employees with more than five years of seniority. Regardless of the job position, no influence of gender was found, suggesting that male and female workers face the same putative occupational health risk factors. Organizations must therefore consider the personal characteristics of workers to ensure the effectiveness of interventions and programs related to improving workplace health.

### Limitations

We acknowledge some limitations in our study. Workers' self-reported responses to questionnaires may have resulted in self-report bias [45]. This could also have led to an affective bias caused by social desirability [46]. However, Wittyfit allows its users to express themselves freely by anonymizing all user data, thus reducing this bias. The process of averaging the records may have hidden a potential seasonal bias. Nevertheless, the distribution of records, reported in S4 Table 1 in S4 Appendix, appeared very arbitrary, revealing no seasonal effect. The COVID-19 pandemic in France may have affected the responses given by workers in 2020 (or even 2021). On the one hand, we found, using the linear mixed model with time effect treated as random, that the differences in stress and job satisfaction between managers and employees remained consistent before, during, and after 2020. On the other hand, the temporal stability provided by the cross-sectional panel analysis tends to prove that the influence of COVID-19 pandemic was negligible. Although stress and job satisfaction are known to be associated with well-being [47] and may therefore appear as an important covariate, we did not measure it. The study suffered from a lack of information on workers' socio-demographic characteristics, such as employment, level of education or income. It cannot be ruled out that workers with a low level of education or limited computer skills may have been excluded. As it is well documented that the level of job stress and job satisfaction varies by the work sector for both employees and managers [48], our results could not be generalized to whole working population. The low sample size also precluded generalizability. However, the database grows over the years [22], and further studies with larger populations and longer time periods could be designed to confirm or add to these initial results. It would also be very valuable to compare the results of this study, based on a real-world database and therefore constitutes real evidence, with those of a randomized clinical trial [41].

### Conclusions

Although they do not perceive a difference in stress, managers have higher job satisfaction than employees. The changes in both indicators over time, about job position and other socio-demographic factors, may have long-term effects on occupational health of workers. Moreover, the latent evolutionary trajectories followed by some of the workers suggest that by understanding their putative drivers, organizations may be able to improve the occupational health of all their workers in a holistic way.

### Supporting information

**S1 Appendix. Distribution of outcomes.**  
(DOCX)

**S2 Appendix. Statistical analysis plan.**  
(DOCX)



**S3 Appendix. Sensitivity analysis results.**

(DOCX)

**S4 Appendix. Reporting guidelines.**

(DOCX)

**S1 Data.**

(ZIP)

**Acknowledgments**

We express our sincere gratitude to all voluntary workers using Wittyfit, who participated in this study.

**Author Contributions**

**Conceptualization:** Rémi Colin-Chevalier, Frédéric Dutheil, Samuel Dewavrin, Thomas Cornet.

**Data curation:** Samuel Dewavrin, Thomas Cornet.

**Formal analysis:** Rémi Colin-Chevalier.

**Investigation:** Rémi Colin-Chevalier.

**Methodology:** Rémi Colin-Chevalier, Frédéric Dutheil, Bruno Pereira.

**Project administration:** Rémi Colin-Chevalier.

**Resources:** Rémi Colin-Chevalier.

**Software:** Rémi Colin-Chevalier.

**Supervision:** Frédéric Dutheil, Bruno Pereira.

**Validation:** Rémi Colin-Chevalier, Frédéric Dutheil, Amanda Clare Benson, Samuel Dewavrin, Thomas Cornet, Céline Lambert, Bruno Pereira.

**Visualization:** Rémi Colin-Chevalier.

**Writing – original draft:** Rémi Colin-Chevalier.

**Writing – review & editing:** Rémi Colin-Chevalier, Frédéric Dutheil, Amanda Clare Benson, Samuel Dewavrin, Thomas Cornet, Céline Lambert, Bruno Pereira.

**References**

1. Chrousos GP. Stress and disorders of the stress system. *Nat Rev Endocrinol.* 2009; 5(7):374–81. <https://doi.org/10.1038/nrendo.2009.106> PMID: 19488073
2. Chrousos GP, Gold PW. The concepts of stress and stress system disorders: overview of physical and behavioral homeostasis. *JAMA.* 1992; 267(9):1244–52.
3. Steptoe A, Kivimäki M. Stress and cardiovascular disease. *Nat Rev Cardiol.* 2012; 9(6):360–70. <https://doi.org/10.1038/nrcardio.2012.45> PMID: 22473079
4. Santosa A, Rosengren A, Ramasundarahettige C, Rangarajan S, Gulec S, Chifamba J, et al. Psychosocial risk factors and cardiovascular disease and death in a population-based cohort from 21 low-, middle-, and high-income countries. *JAMA Netw Open.* 2021; 4(12):e2138920. <https://doi.org/10.1001/jamanetworkopen.2021.38920> PMID: 34910150
5. Zhou AY, Panagioti M, Esmail A, Agius R, Van Tongeren M, Bower P. Factors associated with burnout and stress in trainee physicians: a systematic review and meta-analysis. *JAMA Netw Open.* 2020; 3(8):e2013761. <https://doi.org/10.1001/jamanetworkopen.2020.13761> PMID: 32809031

6. Lim JY, Kim GM, Kim EJ. Factors associated with job stress among hospital nurses: a meta-correlation analysis. *Int J Environ Res Public Health*. 2022; 19(10):5792. <https://doi.org/10.3390/ijerph19105792> PMID: 35627328
7. Backé EM, Seidler A, Latza U, Rossnagel K, Schumann B. The role of psychosocial stress at work for the development of cardiovascular diseases: a systematic review. *Int Arch Occup Environ Health*. 2012; 85(1):67–79. <https://doi.org/10.1007/s00420-011-0643-6> PMID: 21584721
8. Locke EA. The nature and causes of job satisfaction. *Handbook of Industrial and Organizational Psychology*. 1976; 1:1297–343.
9. Faragher EB, Cass M, Cooper CL. The relationship between job satisfaction and health: a meta-analysis. *Occup Environ Med*. 2005; 62(2):105–12. <https://doi.org/10.1136/oem.2002.006734> PMID: 15657192
10. Lu H, Zhao Y, While A. Job satisfaction among hospital nurses: a literature review. *International Journal of Nursing Studies*. 2019; 94:21–31. <https://doi.org/10.1016/j.ijnurstu.2019.01.011> PMID: 30928718
11. Penconek T, Tate K, Bernardes A, Lee S, Micaroni SPM, Balsanelli AP, et al. Determinants of nurse manager job satisfaction: A systematic review. *Int J Nurs Stud*. 2021; 118:103906. <https://doi.org/10.1016/j.ijnurstu.2021.103906> PMID: 33765624
12. Rahman M, Sen AK. Effect of job satisfaction on stress, performance and health in self-paced repetitive work. *Int Arch Occup Environ Health*. 1987; 59(2):115–21. <https://doi.org/10.1007/BF00378489> PMID: 3557622
13. Amati M, Tomasetti M, Ciuccarelli M, Mariotti L, Tarquini LM, Bracci M, et al. Relationship of job satisfaction, psychological distress and stress-related biological parameters among healthy nurses: a longitudinal study. *J Occup Health*. 2010; 52(1):31–8. <https://doi.org/10.1539/joh.19042> PMID: 20032591
14. Pecino V, Mañas MA, Díaz-Fúnez PA, Aguilar-Parra JM, Padilla-Góngora D, López-Liria R. Organisational climate, role stress, and public employees' job satisfaction. *Int J Environ Res Public Health*. 2019; 16(10):E1792. <https://doi.org/10.3390/ijerph16101792> PMID: 31117168
15. Kim KW, Cho YH. The Moderating Effect of Managerial Roles on Job Stress and Satisfaction by Employees' Employment Type. *Int J Environ Res Public Health*. 2020; 17(21):E8259. <https://doi.org/10.3390/ijerph17218259> PMID: 33182275
16. Tarcan M, Hikmet N, Schooley B, Top M, Tarcan GY. An analysis of the relationship between burnout, socio-demographic and workplace factors and job satisfaction among emergency department health professionals. *Applied Nursing Research*. 2017; 34:40–7. <https://doi.org/10.1016/j.apnr.2017.02.011> PMID: 28342622
17. Gazioglu S, Tansel A. Job satisfaction in Britain: individual and job related factors. *Applied Economics*. 2006; 38(10):1163–71.
18. Colin-Chevalier R, Pereira B, Benson AC, Dewavrin S, Cornet T, Dutheil F. The protective role of job control/autonomy on mental strain of managers: a cross-sectional study among Wittyfit's users. *International Journal of Environmental Research and Public Health*. 2022; 19(4):2153. <https://doi.org/10.3390/ijerph19042153> PMID: 35206335
19. Brett JM, Stroh LK. Working 61 plus hours a week: why do managers do it? *J Appl Psychol*. 2003; 88(1):67–78. <https://doi.org/10.1037/0021-9010.88.1.67> PMID: 12675395
20. Bugajska J, Jędryka-Góral A, Widerszal-Bazyl M, Orłowska-Baranowska E, Sagan A, Michalak JM, et al. Job strain, overtime, life style, and cardiovascular risk in managers and physical workers. *Int J Occup Saf Ergon*. 2011; 17(1):25–32. <https://doi.org/10.1080/10803548.2011.11076867> PMID: 21375951
21. Ikesu R, Miyawaki A, Svensson AK, Svensson T, Kobayashi Y, Chung UI. Association of managerial position with cardiovascular risk factors: A fixed-effects analysis for Japanese employees. *Scand J Work Environ Health*. 2021; 47(6):425–34. <https://doi.org/10.5271/sjweh.3966> PMID: 34013355
22. Dutheil F, Duclos M, Naughton G, Dewavrin S, Cornet T, Huguet P, et al. Wittyfit-live your work differently: study protocol for a workplace-delivered health promotion. *JMIR Research Protocols*. 2017; 6(4):e6267. <https://doi.org/10.2196/resprot.6267> PMID: 28408363
23. Lesage FX, Berjot S, Deschamps F. Clinical stress assessment using a visual analogue scale. *Occup Med (Lond)*. 2012; 62(8):600–5. <https://doi.org/10.1093/occmed/kqs140> PMID: 22965867
24. Dutheil F, Trousselard M, Perrier C, Lac G, Chamoux A, Duclos M, et al. Urinary interleukin-8 is a biomarker of stress in emergency physicians, especially with advancing age—the JOBSTRESS\* randomized trial. *PLoS One*. 2013; 8(8):e71658. <https://doi.org/10.1371/journal.pone.0071658> PMID: 23977105
25. Dutheil F, Pereira B, Moustafa F, Naughton G, Lesage FX, Lambert C. At-risk and intervention thresholds of occupational stress using a visual analogue scale. *PLoS One*. 2017; 12(6):e0178948. <https://doi.org/10.1371/journal.pone.0178948> PMID: 28586383
26. Rogosa D, Brandt D, Zimowski M. A growth curve approach to the measurement of change. *Psychological Bulletin*. 1982; 92(3):726–48.

27. Nagin DS, Jones BL, Passos VL, Tremblay RE. Group-based multi-trajectory modeling. *Stat Methods Med Res*. 2018; 27(7):2015–23. <https://doi.org/10.1177/0962280216673085> PMID: 29846144
28. Nagin DS. *Group-Based Modeling of Development*. Harvard University Press; 2005. 226 p.
29. Nguena Nguetack HL, Pagé MG, Katz J, Choinière M, Vanasse A, Dorais M, et al. Trajectory Modelling Techniques Useful to Epidemiological Research: A Comparative Narrative Review of Approaches. *Clin Epidemiol*. 2020; 12:1205–22. <https://doi.org/10.2147/CLEP.S265287> PMID: 33154677
30. Hedges LV. Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*. 1981; 6(2):107–28.
31. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. L. Erlbaum Associates; 1988. 567 p.
32. Karasek RA. Job Demands, Job Decision Latitude, and Mental Strain: Implications for Job Redesign. *Administrative Science Quarterly*. 1979; 24(2):285–308.
33. Karasek R, Theorell T. *Healthy work: stress, productivity, and the reconstruction of working life*. New York: Basic Books; 1990.
34. Bonsaksen T, Thørrisen MM, Skogen JC, Aas RW. Who reported having a high-strain job, low-strain job, active job and passive job? The WIRUS Screening study. *PLoS One*. 2019; 14(12):e0227336. <https://doi.org/10.1371/journal.pone.0227336> PMID: 31887201
35. Lundqvist D. Psychosocial work environment and health when entering or leaving a managerial position. *Work*. 2022; <https://doi.org/10.3233/WOR-210469> PMID: 35938267
36. Jenaro C, Flores N, Orgaz MB, Cruz M. Vigour and dedication in nursing professionals: towards a better understanding of work engagement. *J Adv Nurs*. 2011; 67(4):865–75. <https://doi.org/10.1111/j.1365-2648.2010.05526.x> PMID: 21129008
37. Cheung K, Ching SSY. Job satisfaction among nursing personnel in Hong Kong: a questionnaire survey. *Journal of Nursing Management*. 2014; 22(5):664–75. <https://doi.org/10.1111/j.1365-2834.2012.01475.x> PMID: 25041805
38. Han GH, Jekel M. The mediating role of job satisfaction between leader-member exchange and turnover intentions. *J Nurs Manag*. 2011; 19(1):41–9. <https://doi.org/10.1111/j.1365-2834.2010.01184.x> PMID: 21223404
39. Satuf C, Monteiro S, Pereira H, Esgalhado G, Marina Afonso R, Loureiro M. The protective effect of job satisfaction in health, happiness, well-being and self-esteem. *Int J Occup Saf Ergon*. 2018; 24(2):181–9. <https://doi.org/10.1080/10803548.2016.1216365> PMID: 27560543
40. Hou J, He Y, Zhao X, Thai J, Fan M, Feng Y, et al. The effects of job satisfaction and psychological resilience on job performance among residents of the standardized residency training: a nationwide study in China. *Psychol Health Med*. 2020; 25(9):1106–18. <https://doi.org/10.1080/13548506.2019.1709652> PMID: 31992067
41. Colin-Chevalier R, Dutheil F, Cambier S, Dewavrin S, Cornet T, Baker JS, et al. Methodological issues in analyzing real-world longitudinal occupational health data: a useful guide to approaching the topic. *International Journal of Environmental Research and Public Health*. 2022; 19(12):7023. <https://doi.org/10.3390/ijerph19127023> PMID: 35742269
42. Cheng CY, Liou SR, Tsai HM, Chang CH. Job stress and job satisfaction among new graduate nurses during the first year of employment in Taiwan. *Int J Nurs Pract*. 2015; 21(4):410–8. <https://doi.org/10.1111/ijn.12281> PMID: 24666722
43. Jones MK, Latreille PL, Sloane PJ, Staneva AV. Work-related health risks in Europe: Are older workers more vulnerable? *Social Science & Medicine*. 2013; 88:18–29. <https://doi.org/10.1016/j.socscimed.2013.03.027> PMID: 23702206
44. Anser MK, Ali M, Anwar F, Usman M. Subjective Age and Job Satisfaction: A Moderated Mediation Model of Job Burnout and Chronological Age. *Front Public Health*. 2020; 8:62. <https://doi.org/10.3389/fpubh.2020.00062> PMID: 32211366
45. Devaux M, Sassi F. Social disparities in hazardous alcohol use: self-report bias may lead to incorrect estimates. *European Journal of Public Health*. 2016; 26(1):129–34. <https://doi.org/10.1093/eurpub/ckv190> PMID: 26585784
46. Crowne DP, Marlowe D. A new scale of social desirability independent of psychopathology. *J Consult Psychol*. 1960; 24:349–54. <https://doi.org/10.1037/h0047358> PMID: 13813058
47. Locker D. Work stress, job satisfaction and emotional well-being among Canadian dental assistants. *Community Dent Oral Epidemiol*. 1996; 24(2):133–7. <https://doi.org/10.1111/j.1600-0528.1996.tb00830.x> PMID: 8654035
48. Brown A, Kitchell M, O'Neill T, Locklear J, Vosler A, Kubek D, et al. Identifying meaning and perceived level of satisfaction within the context of work. *Work*. 2001; 16(3):219–26. PMID: 12441451

## Article n°5

Psychosocial well-being index and sick leave in the workplace: a structural equation modelling of Wittyfit data

## **Article n°5**

### Psychosocial well-being index and sick leave in the workplace: a structural equation modelling of Wittyfit data

Colin-Chevalier R, Pereira B, Dewavrin S, Cornet T, Baker JS, Dutheil F

#### **Résumé**

Dans ce cinquième article, nous avons étudié les liens entre le bien-être psychosocial des travailleurs et le risque d'absentéisme pour cause de maladie sur le lieu de travail. Or, il n'existe dans Wittyfit aucun indicateur de bien-être psychosocial, et avons donc pour objectif d'en construire un. De retour dans un cadre d'étude transversal, nous faisons face aux deux mêmes défis méthodologiques que dans les articles 2 et 3, et avons donc tenu compte de l'« effet entreprise » comme effet cluster tant dans l'ajustement de nos modèles qu'au moment de l'imputation des données. Dans un premier temps, nous avons donc construit un modèle d'équation structurelle capable de fournir une mesure du bien-être psychosocial à partir des dix indicateurs de la plateforme. Les indicateurs utilisés étaient la satisfaction au travail, l'ambiance de travail, la reconnaissance, l'équilibre entre vie personnelle et vie privée, la reconnaissance, le sens, l'organisation du travail, les valeurs professionnelles, la charge de travail, l'autonomie et le stress. Une fois construit, nous avons utilisé le modèle d'équation structurelle pour calculer l'état de bien-être des travailleurs. Dans un second temps, nous avons alors pu étudier l'effet protecteur du bien-être psychosocial sur l'absentéisme pour maladie, d'une part sur le risque d'absence, et d'autre part sur la durée des absences, en utilisant des modèles de régression logistique et négative-binomiale, respectivement. Nous avons observé que plus le bien-être d'un travailleur était élevé, plus son risque d'être absent et sa durée moyenne d'absence étaient faibles. En travaillant à l'amélioration du bien-être psychosocial de leurs salariés, les entreprises pourraient être en mesure de réduire l'absentéisme sur le lieu de travail, le tout en préservant la santé des travailleurs.

# **Psychosocial well-being index and sick leave in the workplace: a structural equation modelling of Wittyfit data**

Rémi Colin-Chevalier,<sup>1\*</sup> Bruno Pereira,<sup>2</sup> Samuel Dewavrin,<sup>3</sup> Thomas Cornet,<sup>3</sup> Julien Steven Baker,<sup>4</sup> Frédéric Dutheil<sup>1</sup>

<sup>1</sup> CNRS, LaPSCo, Physiological and Psychosocial Stress, Université Clermont Auvergne, Occupational and Environmental Medicine, CHU Clermont-Ferrand, F-63000 Clermont-Ferrand, France

<sup>2</sup> Biostatistics Unit, DRCI, CHU Clermont-Ferrand, F-63000 Clermont-Ferrand, France

<sup>3</sup> Cegid, F-69009 Lyon, France

<sup>4</sup> Centre for Health and Exercise Science Research, Hong Kong Baptist University, Kowloon Tong, Hong Kong 999077, China

## **\*Corresponding author**

Rémi Colin, PhD Student

Occupational and Environmental Medicine, CHU Clermont-Ferrand, 58 rue Montalembert, F-63000 Clermont-Ferrand, France

Phone: +33 (0)4 73 26 56 75

Email: [r.colin6374@gmail.com](mailto:r.colin6374@gmail.com)

Short title: Psychosocial well-being index and sick leave

Word count for abstract: 261; Number of tables 1; Number of figures: 5

## Abstract

**Background:** Psychosocial well-being, which assesses emotional, psychological, social, and collective well-being, could help measure risk and duration of sick leave in workers.

**Objective:** This study aims to build a structural equation model of a psychosocial well-being index based on 10 psychosocial factors and investigate its association with sick leave.

**Methods:** Data of workers using Wittyfit was collected in 2018. Psychosocial factors (job satisfaction, atmosphere, recognition, work-life balance, meaning, work organization, values, workload, autonomy, and stress) were self-assessed using health-related surveys, while sick leave records were provided by volunteer companies.

**Results:** A total of 1,399 workers were included in the study (mean age:  $39.4 \pm 9.4$ , mean seniority:  $9.2 \pm 7.7$ , 49.8% of women, 12.0% managers). The prevalence of absenteeism was 34.5%, with an average of  $8.48 \pm 28.7$  days of sick leave per worker. Structural equation modelling facilitated computation of workers' psychosocial well-being index (AIC: 123,016.2, BIC: 123,231.2, RMSEA: 0.03). All factors, except workload ( $p=0.9$ ), were influential, with meaning ( $\beta=0.72$ , 95% CI 0.69 to 0.74), values (0.69, 0.67 to 0.70) and job satisfaction (0.64, 0.61 to 0.66) being the main drivers ( $p<0.001$ ). Overall, psychosocial well-being was found to be a protective factor for sick leave, with a 2% decreased risk (OR=0.98, 95% CI 0.98 to 0.99,  $p<0.001$ ) and duration (IRR=0.98, 95% CI 0.97 to 0.99,  $p<0.001$ ) per psychosocial well-being index point.

**Conclusions:** The psychosocial well-being index provides a measure of psychosocial well-being and helps predict sick leave in the workplace. This new indicator could be used to analyze the association between psychosocial well-being and other health outcomes.

**Trial Registration:** [Clinicaltrials.gov: NCT02596737](https://clinicaltrials.gov/ct2/show/study/NCT02596737).

**Keywords:** psychosocial well-being; psychosocial factors; sick leave; structural equation modelling



## **Introduction**

Psychosocial well-being is a multidimensional concept that encompasses emotional, psychological, social and collective well-being [1], which relates to the concept of “quality of life” [2]. Conversely, psychosocial factors are workplace factors that can affect psychological and social well-being of employees, namely their psychosocial well-being, and potentially harm their physical health [3]. Such factors include job satisfaction, psychological and social support, recognition and rewards, work-life balance, work environment, organizational change and restructuring, job demands/workload, job control/autonomy, and even stress.

Sick leave refers to a period during which a worker is absent from work due to illness and granted a leave of absence. The phenomenon of sick leave is multifaceted and influenced by a variety of personal, social, and demographic factors, as well as organizational factors such as work-related factors both within and outside of the work environment. Such work-related factors include, non-exhaustively, poor work-life balance [4], poor psychosocial work environment [5], low social support [6,7], job dissatisfaction [8,9], job stress [9], and low job control [10,11]. In addition to having a significant impact on the financial health and productivity of an organization [12,13], sick leave appears above all to be a considerable risk for the health of workers [14,15].

Psychosocial well-being is an inclusive and complex concept and seems relevant for investigation via its multiple dimensions before assessing its putative effect on sick leave in the workplace. Thus, the main objective of the study was to build a structural equation model of a psychosocial well-being index (PWI) which embraces multiple psychosocial factors: job satisfaction, atmosphere, recognition, work-life balance, meaning, work organization, values, workload, autonomy, and stress. A secondary objective was to measure the influence of the PWI on both risk and duration of sick leave.

## **Materials and methods**

### **Recruitment**

The software Wittyfit was created by Cegid (Lyon, France) in collaboration with the university hospital of Clermont-Ferrand. It is a web-based application used by many companies across different industries in France [16]. Wittyfit allows companies to evaluate the well-being and psychosocial risks of their employees by gathering data on their feelings, with the goal of enhancing overall performance. Participation is voluntary, and the software provides a way for users to express themselves anonymously by responding to a variety of health-related surveys, which can help assess their general health status. The software then provides personalized feedback based on this information. Data collection for this study was limited to the year 2018. The study included all employees who had completed surveys using Wittyfit at least once during the year, whether they had completed fully or only partially completed the surveys. Sick leave data for 2018 was provided by two clients of Wittyfit. Before participating in the study, all subjects gave their informed consent, as outlined in the terms and conditions of Wittyfit. The study itself was approved by two regulatory bodies: the French data protection authority (CNIL), as well as the Ethics Committee of Clermont-Ferrand (South-East IV).

### **Measures**

#### **Psychosocial factors**

Psychosocial factors include job satisfaction, atmosphere, recognition, work-life balance, meaning, work organization, values, workload, autonomy, and stress. To evaluate all factors, Wittyfit uses visual analog scales with a score range of zero to 100 that are specific to these factors. For all factors except stress, a score of zero indicates a poor score in that factor, while a score of 100 indicates a good score. For stress, the lower the score, the better the worker's feeling. The use of visual analog scales is common to measure health indicators, such as stress

for example [17–19]. Workers were able to estimate their feeling related to each factor multiple times throughout the year. The average score for each factor was retained. If a worker expressed no feeling for a factor, we considered it a missing value.

### **Sick leave risk and duration**

If a worker was absent due to illness during the year, his or her period of unavailability was measured in days. The total duration of sick leave was then determined by adding up all the periods of unavailability and rounding up to the nearest whole day. If a worker had never been absent due to illness, his or her total duration was considered zero. Finally, we considered a worker to be absent if he or she counted one or more sick leave days, i.e., if his or her total duration was non-zero.

### **Sociodemographics**

Companies using Wittyfit provided age, seniority, gender, and job position (manager or employee) of their workers. For the purposes of this study, a manager was defined as an individual who is responsible for overseeing and managing an organization or a group of workers, namely the employees.

### **Statistical analysis**

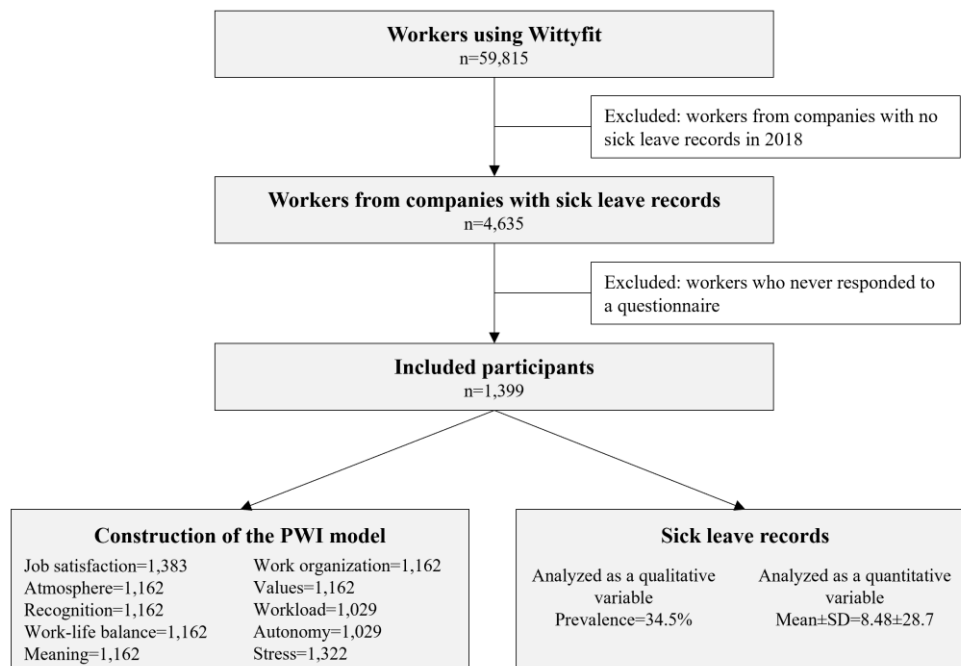
Quantitative data (age, seniority, psychosocial factors, duration of sick leave, PWI) were expressed in terms of mean $\pm$ SD. Qualitative data (share of women, share of managers, prevalence of the risk of sick leave) were expressed in terms of number of individuals and associated percentage. To construct the SEM of PWI, a two-step approach was performed. Firstly, we fitted several confirmatory factor analysis models to highlight the best structure of psychosocial factors to explain the PWI. Each psychosocial factor has been considered as an item. From a general model, we then made step by step modifications to improve the outcomes.

Then, at each step, modification indices were calculated to observe whether the fitted model would improve by adding a path or releasing a constraint. If the change made improved the model, the new model was retained, otherwise a different change was made. Metrics such as the root mean square errors of approximation, Akaike's Information Criteria and Bayesian Information Criteria were used to compare nested models. Whatever the metric considered, the lower the value, the better the model fit. Regarding the RMSEA, a value close to zero indicated a good fit. A value  $p < 0.08$  is considered satisfactory, or otherwise poor [20]. When no further indices could be calculated, or when the model could no longer be improved, we considered the model to be optimal. Secondly, the optimal SEM was fitted and used to assess the influence of the PWI on sick leave risk. CFA and SEM were performed using the "cfa" and "sem" commands of the "lavaan" package of R, respectively. Once constructed, PWI was computed for each worker. Univariate linear mixed-effect models with company effect as random were fitted to assess the influence of each psychosocial factor on PWI to validate its construction. Finally, a generalized linear mixed-effect model with company effect as random was fitted to assess the influence of PWI on sick leave risk, measured in odds ratio and 95% confidence interval, and a negative binomial generalized linear mixed-effect model with company effect as random was fitted to assess the influence of PWI on sick leave duration, measured in incidence rate ratio and 95% CI. In the absence of data for psychosocial factors, the type of which we have identified as being missing completely at random [21], hierarchical multiple imputation with company as a cluster was performed using the "mice" command of the eponymous package of R with age, seniority, gender and job position as predictive factors to avoid complete-cases analysis bias [22,23]. Multicollinearity of outcomes was checked prior to analysis. Statistical analyses were performed using R (version 4.2.2) [24]. Unless specified, a p-value  $< 0.05$  was considered as statistically significant.

# Results

## Participants

Data was collected from 1,399 workers, with a mean age of  $39.4 \pm 9.4$  years, a mean seniority of  $9.2 \pm 7.7$  years, divided into 702 men (50.2%) and 697 (49.8%) women. The workers' subdivisions consisted of 1231 being employees (88.0%) and 168 being managers (12.0%). Among these workers, 483 (34.5%) were absent due to illness at least once during 2018, and the average duration of sick leave was  $8.5 \pm 28.7$  (estimated kurtosis: 58.9) (Figure 1). Further information on statistics and correlations between psychosocial factors are provided in Appendix (Table S1, Table S2).



**Figure 1.** Flowchart of the study.

Legend: The numbers associated with each psychosocial factor correspond to the number of records (i.e., the number of non-missing values).

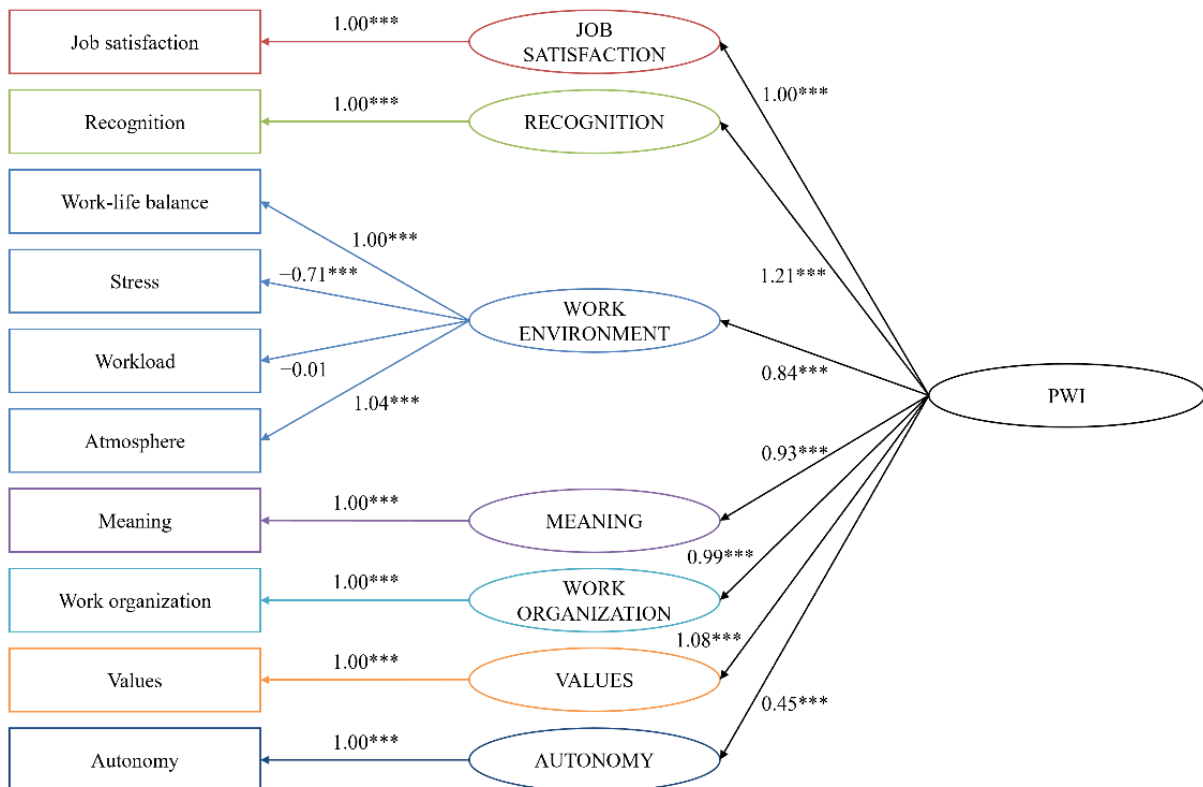
## Construction of the psychosocial well-being index

The first model (model 1) tested was the one in which each psychosocial factor was considered a measure of a distinct latent class, all of which were related to a PWI latent class, with no covariance between items. Except for the latent class related to the workload item ( $p=0.9$ ), all

latent classes showed significant loading on the PWI ( $p < 0.001$ ). This result remained valid in all models. The latent class related to the recognition item was found to have the highest absolute loading on the PWI ( $\beta = 1.17$ , 95% CI 1.00 to 1.34,  $p < 0.001$ ) among all classes. With a RMSEA value of 0.10, the goodness-of-fit was found to be poor. The first change made to the model, leading to model 2, was the addition of the covariance measurement between workload and stress items. This model appeared to be better than model 1 ( $\chi^2$ -test  $p < 0.001$ ). Despite lower AIC (123,327.6 vs. 123,456.7) and BIC (123,490.1 vs. 123,614.0), goodness-of-fit of the model remained poor (RMSEA=0.09). In model 3, covariance between workload and autonomy items was added ( $p < 0.001$ ). In addition to being better than the previous model (AIC=123,214.8, BIC=123,382.6), the goodness-of-fit of the model appeared to be satisfactory (RMSEA=0.07) but still improvable. This result remained valid from this model to the optimal and last one. Model 4 was built by gathering work-life balance and stress items into the same latent class (the latent class related to the work-life balance item). Next, model 5 was built by gathering the latent class related to work-life balance and stress and the one related to workload, but its BIC was found to be higher than model 4 (123,339.1 vs. 123,337.6), which led us to test another change. Thus, model 5.2 was built by gathering work-life balance, stress, and workload items in the same latent class. This time, the BIC was found to be lower (123,336.5 vs. 123,337.6) so the model was kept. In the same way, the atmosphere item was added to the latent class which included work-life balance, stress, and workload to build model 6, resulting in a huge increase in goodness-of-fit (RMSEA=0.06). From this model, the number of latent classes in the models that haven been preserved step by step remained stable at seven: six measured by a single item (job satisfaction, recognition, meaning, work organization, values, and autonomy) and the last measured by the four gathered items (work-life balance, stress, workload, and atmosphere). For ease of understanding, each one-item latent class was named by the item that measured it. For example, the latent class measured by the job satisfaction item was called JOB SATISFACTION. In view of the elements that measure it, the four-items latent class was called



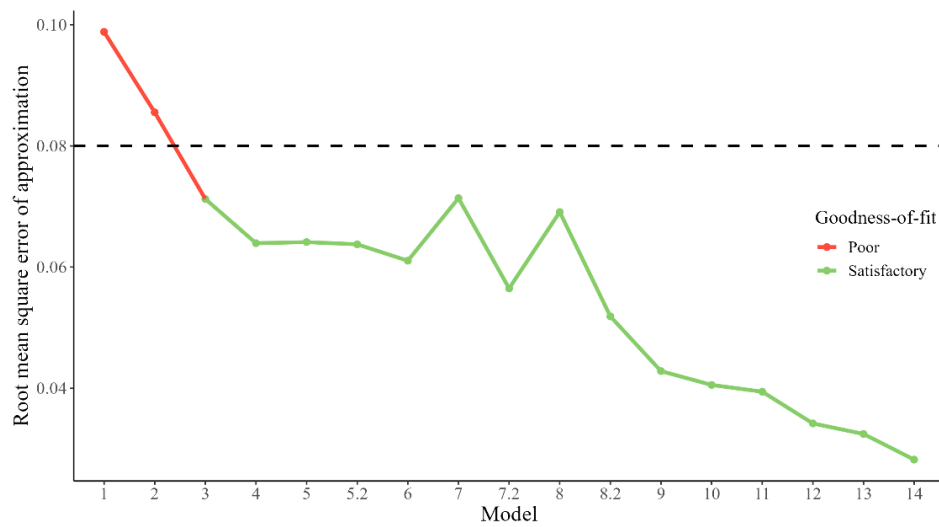
the WORK ENVIRONMENT class. As model 7 which consisted of adding the recognition item to the WORK ENVIRONMENT latent class was found to be less well-fitting than the model 6 (BIC: 123,382.6 vs. 123,321.4), model 7.2 was built by adding the covariance between stress and work organization items, which fitted better (BIC: 123,299.1). Similarly, model 8 which consisted of adding the recognition item to the largest latent class proved to be less well-adjusted than model 7, so model 8.2 which consisted of adding covariance between the RECOGNITION and WORK ENVIRONMENT latent classes was built and was better adjusted in terms of goodness-of-fit. Finally, models 9 to 14 were built by adding covariances between meaning and values items, WORK ENVIRONMENT and VALUES latent classes, recognition and workload items followed by work-life balance and stress, job satisfaction and recognition, and lastly work-life balance and values. After model 14, no further improvements could have been made. With a RMSEA of 0.03, the model 14 was considered as optimal (Figure 2).



**Figure 2.** Illustration of the structural equation model of the psychosocial well-being index.

Legend: All measured variables are presented in squares, all latent variables in circles. Covariances between items and variables are not shown but present. ‘\*\*\*’:  $p < 0.001$ .

Summary of step-by-step approach to the construction of the optimal model can be found in Figure 3 and Table 1.



**Figure 3.** Evolution of goodness-of-fit across models.

Legend: A point of green color indicates a satisfactory goodness-of-fit of the model (i.e., a RMSEA <0.8), a red color a poor one.

**Table 1.** Fit indices of models.

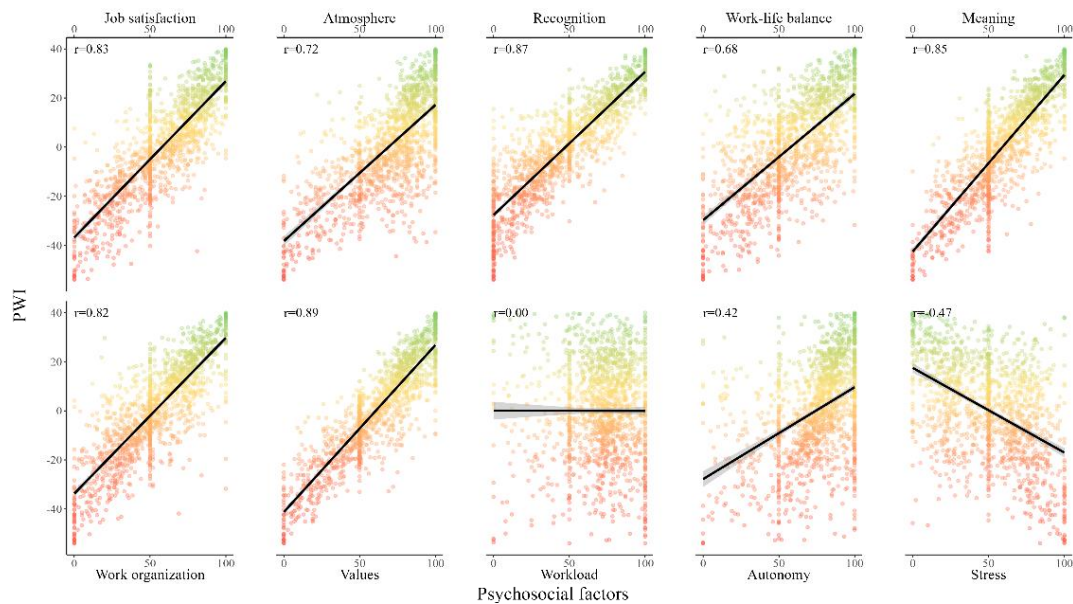
Model	Specifications	df	AIC	BIC	RMSEA
Model 1	Each variable as a latent class	35	123,456.7	123,614.0	0.10
Model 2	Covariance between workload and stress items	34	123,327.6	123,490.1	0.09
Model 3	Covariance between workload and autonomy items	33	123,214.8	123,382.6	0.07
Model 4	Work-life balance and stress in the same latent class	32	123,164.5	123,337.6	0.06
Model 5	Covariance between latent classes 4 (work-life balance/stress) and 8 (workload)	31	123,160.8	123,339.1	0.06
Model 5.2	Work-life balance, stress, and workload in the same latent class	32	123,163.5	123,336.5	0.06
Model 6	Work-life balance, stress, workload, and atmosphere in the same latent class	32	123,148.3	123,321.4	0.06
Model 7	Work-life balance, stress, workload, atmosphere, and recognition in the same latent class	32	123,209.5	123,382.6	0.07

Model	Specifications	df	AIC	BIC	RMSEA
Model 7.2	Covariance between stress and work organization items	31	123,120.9	123,299.1	0.06
Model 8	Work-life balance, stress, workload, atmosphere, and recognition in the same latent class	31	123,189.5	123,367.8	0.07
Model 8.2	Covariance between latent classes 3 (recognition) and (work-life balance/stress/workload/atmosphere)	30	123,096.4	123,279.9	0.05
Model 9	Covariance between meaning and values items	29	123,058.9	123,247.7	0.04
Model 10	Covariance between latent classes 4 (work-life balance/stress/workload/atmosphere) and 7 (values)	28	123,049.9	123,243.9	0.04
Model 11	Covariance between recognition and workload items	27	123,045.2	123,244.5	0.04
Model 12	Covariance between work-life balance and stress items	26	123,030.0	123,234.5	0.03
Model 13	Covariance between job satisfaction and recognition items	25	123,025.3	123,235.1	0.03
Model 14	Covariance between work-life balance and values items	24	123,016.2	123,231.2	0.03

Legend: 'df': degrees of freedom, 'AIC': Akaike's Information Criteria, 'BIC': Bayesian Information Criteria, 'RMSEA': root mean square error of approximation.

### Calculation of the psychosocial well-being index and its influence on sick leave

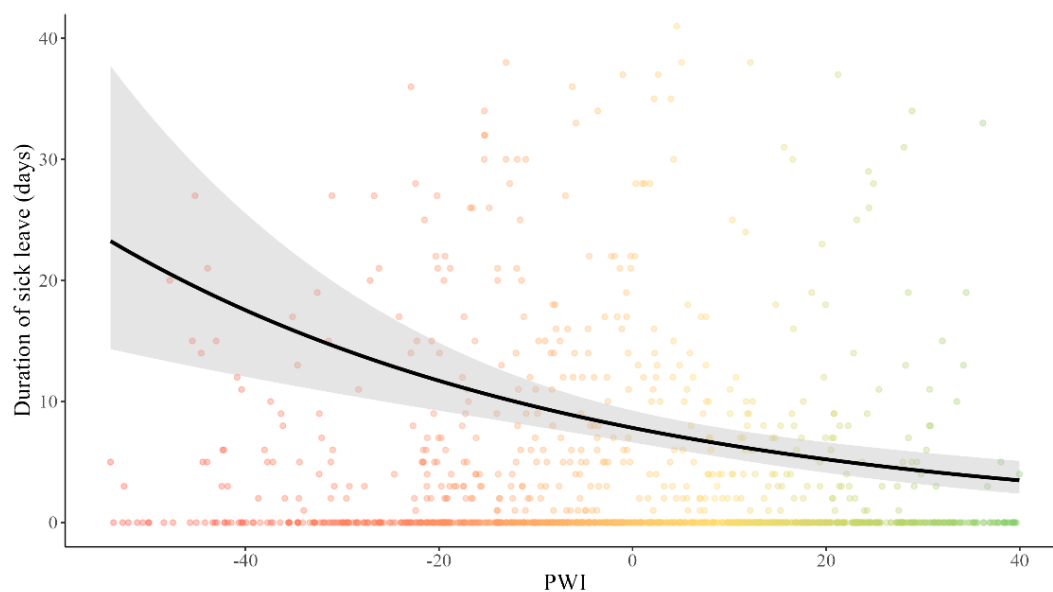
Using the optimal SEM, the PWI of each worker was computed ( $0.00 \pm 19.8$ ,  $\text{MIN} = -53.9$ ,  $\text{MAX} = 40.0$ ). Univariate linear mixed-effect models revealed that, except for workload ( $0.00$ ,  $95\% \text{ CI } -0.05 \text{ to } 0.05$ ,  $p = 0.9$ ), all psychosocial factors had an impact on the PWI ( $p < 0.001$ ). In influential order, meaning was the factor with the highest influence across all factors ( $0.72$ ,  $0.69$  to  $0.74$ ), followed by values ( $0.69$ ,  $0.67$  to  $0.70$ ), job satisfaction ( $0.64$ ,  $0.61$  to  $0.66$ ), work organization ( $0.63$ ,  $0.61$  to  $0.66$ ), recognition ( $0.59$ ,  $0.57$  to  $0.60$ ), atmosphere ( $0.57$ ,  $0.54$  to  $0.59$ ), autonomy ( $0.37$ ,  $0.33$  to  $0.41$ ) and stress ( $-0.34$ ,  $-0.38$  to  $-0.31$ ) (Figure 4).



**Figure 4.** Influence of psychosocial factors on psychosocial well-being index.

Legend: ‘r’: Pearson’s correlation between the psychosocial well-being index and the psychosocial factor.

Overall, the PWI presents a positive influence on sick leave, and a greater influence than all psychosocial factors. Indeed, for each PWI point, the risk of sick leave was decreased by 2% (OR=0.98, 95% CI 0.98 to 0.99,  $p < 0.001$ ). The duration of sick leave also showed a significantly declining trend (IRR=0.98, 95% CI 0.97 to 0.99,  $p < 0.001$ ) with an expected duration of 7.8 days of sick leave for an individual with zero PWI. These results remained true with the addition of age, seniority, gender, and job position as confounding variables (Figure 5).



**Figure 5.** Influence of psychosocial well-being index on duration of sick leave.

Legend: The y-axis, “duration of sick leave”, is the total duration of sick leave during the year (in days).

## **Discussion**

The main results showed that in addition to synthesizing and embracing all considered psychosocial factors, the PWI appeared as a notable factor of sick leave in the workplace. Compared to all other psychosocial factors, PWI demonstrated the greatest influence on sick leave, for both risk and duration.

### **Principal results**

SEM is a highly effective multivariate technique that is increasingly used in scientific research to evaluate complex causal relationships among variables. Unlike other modelling approaches, SEMs allow to measure supposed variables that could not be measured previously, also called latent variables. This feature offers researchers the opportunity to create measurement scales, which is particularly useful in the health sector [25,26]. In this study, we aimed to create an index able to measure workers' psychosocial well-being (or quality of life), called PWI. A total of 10 psychosocial factors were selected, combined, and weighted in a SEM to allow identification and calculation of a conceptualized index. Except for workload (even if covariances between workload and recognition, autonomy, and stress are computed in the model), all psychosocial factors were found to be significantly related to psychosocial well-being. This result suggests that, as desired, PWI provides a holistic understanding of worker health. In their desire to promote health at work, organizations should strive to work on all dimensions of psychosocial well-being since each contributes to its improvement. Following the construction of the PWI, we analyzed how psychosocial well-being could increase the risk of sick leave in workers. Our findings suggests that workers who may experience discomfort due to their working conditions are at greater risk of deteriorating health, resulting in an increased risk of sick leave. In addition to reducing the risk of absence due to illness, we found that workers' average duration of sick leave over the year was lower as their PWI increased, and vice versa. In view of all these results, it appears that the improvement of psychosocial

working conditions could thus contribute significantly to reducing absenteeism due to illness [27]. In theory, interventions to promote psychosocial well-being in the workplace should lead to a decrease in sick leave in the workplace, in addition to, and most importantly, preserving workers' health.

## **Limitations**

We acknowledged some limitations within the study. The possibility of self-report bias and affective bias due to social desirability may exist in workers' self-reported responses to surveys. However, Wittyfit minimizes bias by anonymizing all user data, which enables users to express themselves freely [16]. The study population was also limited to French workers, although differences in sick leave risk exist between countries. However, the prevalence of sick leave in our study, measured at 34.5%, is consistent with the literature [15], which supports our results. Also, the limited data sample used in the study may not be a true representation of the entire population of French workers, notably in terms of companies, where differences in the prevalence of sick leave may exist. To reduce this bias, all models were transformed into mixed models in which the firm effect was considered as a cluster. As a cross-sectional study, causality cannot be established. It is possible that workers who have better psychosocial well-being are less likely to take sick leave [28,29], but it is also possible that taking sick leave affects psychosocial well-being and intentions to go back to work [30]. No information underlying the illnesses that caused workers to stop working was provided by the client companies included in the study. However, it is known that shorter absences tend to be caused by illnesses with short recovery time (cold, cough, and influenza), and longer absences by illnesses with long recovery time, mainly associated with mental health issues [31]. Nevertheless, even low-recovery-time illnesses such as cold [32] or high-recovery-time illnesses as depression [33], can be caused by psychosocial factors. Future studies that examine the reasons for sick leave could strengthen the extent of the association between psychosocial well-being and sick leave.



## **Comparison with prior work**

The holistic aspect of PWI reinforces the idea that understanding worker well-being cannot be done without a comprehensive approach, which is the stated goal of the Wittyfit platform [16]. Among the factors included in the model, meaning proved to be the most relevant in explaining a worker's psychosocial well-being, which suggests that meaning at work might be the main driver of workers' well-being, which is not surprising since this concept is known to offer an existential justification for work and gives meaning to life [34], which goes beyond professional life, as well as increasing performance [35]. Slightly behind meaning, values and job satisfaction emerged as factors for improving workers well-being as previously observed [36]. The lack of effect of workload may be surprising when it is known that it can predict mental strain, but only when combined with job control (e.g. autonomy) [37–39]. As they are not associated in our model, this result is consistent with the literature, where it has been noted that workload may not have an effect on psychosocial well-being [40]. The association between psychosocial well-being and sick leave is well-known [27,30]. Our results confirm the role of psychosocial factors in the risk of sick leave. For example, a study in the Norwegian armed forces found that psychosocial resilience could predict the likelihood of absence due to illness [28]. More generally, a study in pregnant employees found that poor health was associated with more sick leave [41]. Similarly, we found that psychosocial well-being could reduce the average length of sick leave for workers. A study in a Dutch telecommunications company showed that coping, which refers to how individuals deal with stressors, is an important factor in promoting psychosocial well-being and reducing duration of sickness absence [42]. The duration of sickness absenteeism has also been observed to increase in response to adverse psychosocial job conditions [43].

## **Conclusions**

This study highlighted the importance of psychosocial well-being, but most importantly of the PWI, and its association with sick leave in the workplace. The findings suggest that interventions aimed at improving psychosocial well-being factors could help reduce sick leave risk and mean duration. This study provides valuable insights for employers and policymakers to implement measures to improve workers' psychosocial well-being, which could lead to a healthier and more productive workforce. Further research could investigate the effectiveness of interventions to improve psychosocial well-being in reducing sick leave and improving overall health outcomes.

## **Authors' Contributions**

Study design: FD, SD.

Study conduct: RC, FD, BP.

Data collection: SD, TC.

Data analysis: RC, BP.

Data interpretation: RC, FD, BP.

Drafting manuscript: RC.

Revising manuscript content: RC, FD, JSB, BP.

Approving final version of manuscript: all authors.

RC, BP, FD had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

## **Acknowledgements**

We express our sincere gratitude to all voluntary workers using Wittyfit, who participated in this study.

## Disclosure of potential conflicts of interest

RC, SD, and TC are part of Cegid. Other authors have declared that no competing interests exist (FD is responsible for the scientific accuracy of Wittyfit but is not paid by Cegid; as previously published, Wittyfit is a public private partnership with the CHU Clermont-Ferrand).

## Sources of funding and support

This research received no external funding.

## Data sharing statement

Data cannot be transmitted without the prior consent of the company's corporate clients, except to the University Hospital of Clermont-Ferrand, France, which may use the data for research purposes.

## Supplementary materials

**Table S1. Statistics on psychosocial factors.**

Variable	n	MIN	MAX	Mean±SD	Shapiro-Wilk test (statistic and p-value)
Job satisfaction	1399	0	100	57.8±25.7	0.97***
Atmosphere	1399	0	100	69.0±25.9	0.91***
Recognition	1399	0	100	47.4±29.5	0.96***
Work-life balance	1399	0	100	57.8±26.3	0.97***
Meaning	1399	0	100	59.1±23.4	0.97***
Work organization	1399	0	100	53.2±25.7	0.98***
Values	1399	0	100	60.6±26.0	0.96***
Workload	1399	0	100	70.9±21.2	0.93***
Autonomy	1399	0	100	74.3±22.5	0.90***
Stress	1399	0	100	50.7±26.9	0.97***

Legend: 'n': number of individuals, 'MIN': minimum value, 'MAX': maximum value, 'SD': standard deviation  
\*\*\*: p<0.001.

**Table S2. Matrix of Pearson's correlation coefficients between psychosocial factors.**

Variable	1	2	3	4	5	6	7	8	9	10
1. Job satisfaction	--									
2. Atmosphere	0.57***	--								
3. Recognition	0.70***	0.53***	--							
4. Work-life balance	0.52***	0.51***	0.49***	--						
5. Meaning	0.65***	0.56***	0.69***	0.54***	--					
6. Work organization	0.62***	0.54***	0.67***	0.54***	0.66***	--				
7. Values	0.69***	0.58***	0.72***	0.50***	0.75***	0.68***	--			
8. Workload	-0.03	0.03	-0.04	-0.06*	0.03	-0.02	0.02	--		
9. Autonomy	0.34***	0.30***	0.33***	0.24***	0.35***	0.33***	0.36***	0.26***	--	
10. Stress	-0.40***	-0.35***	-0.33***	-0.43***	-0.35***	-0.44***	-0.36***	0.27***	-0.14***	--

Legend: '\*': p<0.05, '\*\*\*': p<0.001.

## References

1. Larson JS. The World Health Organization's definition of health: social versus spiritual health. *Soc Indic Res.* 1996;38(2):181–92.
2. Eiroa-Orosa FJ. Understanding psychosocial wellbeing in the context of complex and multidimensional problems. *Int J Environ Res Public Health.* 2020;17(16):5937.
3. Thomas K, Nilsson E, Festin K, Henriksson P, Lowén M, Löf M, et al. Associations of psychosocial factors with multiple health behaviors: a population-based study of middle-aged men and women. *Int J Environ Res Public Health.* 2020;17(4):1239.
4. Antai D, Oke A, Braithwaite P, Anthony DS. A “balanced” life: work-life balance and sickness absence in four Nordic countries. *Int J Occup Environ Med.* 2015;6(4):205–22.
5. Catalina-Romero C, Sainz JC, Pastrana-Jiménez JI, García-Diéguez N, Irizar-Muñoz I, Aleixandre-Chiva JL, et al. The impact of poor psychosocial work environment on non-work-related sickness absence. *Soc Sci Med.* 2015;138:210–6.
6. Undén AL. Social support at work and its relationship to absenteeism. *Work & Stress.* 1996;10(1):46–61.
7. Silva-Junior JS, Fischer FM. Sickness absence due to mental disorders and psychosocial stressors at work. *Rev Bras Epidemiol.* 2015;18(4):735–44.
8. Lardon A, Dubois JD, Cantin V, Piché M, Descarreaux M. Predictors of disability and absenteeism in workers with non-specific low back pain: a longitudinal 15-month study. *Appl Ergon.* 2018;68:176–85.
9. Tadesse S, Ebrahim K, Gizaw Z. Sickness absenteeism and associated factors among horticulture employees in lume district, southeast Ethiopia. *J Occup Med Toxicol.* 2015;10:32.
10. Gimeno D, Benavides F, Amick B, Benach J, Martinez JM. Psychosocial factors and work related sickness absence among permanent and non-permanent employees. *J Epidemiol Community Health.* 2004;58(10):870–6.

11. Kottwitz MU, Schade V, Burger C, Radlinger L, Elfering A. Time pressure, time autonomy, and sickness absenteeism in hospital employees: a longitudinal study on organizational absenteeism records. *Saf Health Work*. 2018;9(1):109–14.
12. Asay GRB, Roy K, Lang JE, Payne RL, Howard DH. Absenteeism and Employer Costs Associated With Chronic Diseases and Health Risk Factors in the US Workforce. *Prev Chronic Dis*. 2016;13:E141.
13. Nagata T, Mori K, Ohtani M, Nagata M, Kajiki S, Fujino Y, et al. Total Health-Related Costs Due to Absenteeism, Presenteeism, and Medical and Pharmaceutical Expenses in Japanese Employers. *J Occup Environ Med*. 2018;60(5):e273–80.
14. Johnston DA, Harvey SB, Glozier N, Calvo RA, Christensen H, Deady M. The relationship between depression symptoms, absenteeism and presenteeism. *Journal of Affective Disorders*. 2019;256:536–40.
15. Antczak E, Miszczyńska KM. Causes of sickness absenteeism in Europe—analysis from an intercountry and gender perspective. *Int J Environ Res Public Health*. 2021;18(22):11823.
16. Dutheil F, Duclos M, Naughton G, Dewavrin S, Cornet T, Huguet P, et al. Wittyfit-live your work differently: study protocol for a workplace-delivered health promotion. *JMIR Research Protocols*. 2017;6(4):e6267.
17. Lesage FX, Berjot S, Deschamps F. Clinical stress assessment using a visual analogue scale. *Occup Med (Lond)*. 2012;62(8):600–5.
18. Dutheil F, Trousselard M, Perrier C, Lac G, Chamoux A, Duclos M, et al. Urinary interleukin-8 is a biomarker of stress in emergency physicians, especially with advancing age—the JOBSTRESS\* randomized trial. *PLoS One*. 2013;8(8):e71658.
19. Dutheil F, Pereira B, Moustafa F, Naughton G, Lesage FX, Lambert C. At-risk and intervention thresholds of occupational stress using a visual analogue scale. *PLoS One*. 2017;12(6):e0178948.



20. Hooper D, Coughlan J, Mullen M. Structural equation modelling: guidelines for determining model fit. *Electron J Bus Res Methods*. 2008;(6):53–60.
21. Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. Second Edition. New York: John Wiley & Sons; 2002.
22. Li P, Stuart EA, Allison DB. Multiple Imputation: A Flexible Tool for Handling Missing Data. *JAMA*. 2015;314(18):1966–7.
23. Colin-Chevalier R, Dutheil F, Cambier S, Dewavrin S, Cornet T, Baker JS, et al. Methodological Issues in Analyzing Real-World Longitudinal Occupational Health Data: A Useful Guide to Approaching the Topic. *International Journal of Environmental Research and Public Health*. 2022;19(12):7023.
24. R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing; 2021.
25. Yazdanirad S, Golbabaei F, Foroushani AR, Monazzam MR, Dehghan H. Development and validation of an environmental heat strain risk assessment (EHSRA) index using structural equation modeling based on empirical relations. *Environ Health Prev Med*. 2020;25(1):63.
26. Maier T, Kugelmann M, Rhee DS, Brill S, Gündel H, Friemert B, et al. Structural equation modeling of a global stress index in healthy soldiers. *J Clin Med*. 2021;10(8):1799.
27. Mathisen J, Nguyen TL, Jensen JH, Mehta AJ, Rugulies R, Rod NH. Impact of hypothetical improvements in the psychosocial work environment on sickness absence rates: a simulation study. *Eur J Public Health*. 2022;32(5):716–22.
28. Hystad SW, Eid J, Brevik JI. Effects of psychological hardiness, job demands, and job control on sickness absence: a prospective study. *J Occup Health Psychol*. 2011;16(3):265–78.
29. North FM, Syme SL, Feeney A, Shipley M, Marmot M. Psychosocial work environment and sickness absence among British civil servants: the Whitehall II study. *Am J Public Health*. 1996;86(3):332–40.

30. Hedlund Å, Kristofferzon ML, Boman E, Nilsson A. Are return to work beliefs, psychological well-being and perceived health related to return-to-work intentions among women on long-term sick leave for common mental disorders? A cross-sectional study based on the theory of planned behaviour. *BMC Public Health*. 2021;21:535.
31. Demou E, Smith S, Bhaskar A, Mackay DF, Brown J, Hunt K, et al. Evaluating sickness absence duration by musculoskeletal and mental health issues: a retrospective cohort study of Scottish healthcare workers. *BMJ Open*. 2018;8(1):e018085.
32. Bagwath Persad LA, Kamerman PR, Wadley AL. Predictors of Cold and Pressure Pain Tolerance in Healthy South African Adults. *Pain Med*. 2017;18(11):2126–37.
33. Bonde JPE. Psychosocial factors at work and risk of depression: a systematic review of the epidemiological evidence. *Occup Environ Med*. 2008;65(7):438–45.
34. Lee S. A concept analysis of “Meaning in work” and its implications for nursing. *J Adv Nurs*. 2015;71(10):2258–67.
35. Wingerden JV, Stoep JV der. The motivational potential of meaningful work: Relationships with strengths use, work engagement, and performance. *PLOS ONE*. 2018;13(6):e0197599.
36. Kantek F, Kaya A. Professional values, job satisfaction, and intent to leave among nursing managers. *J Nurs Res*. 2017;25(4):319–25.
37. Karasek RA. Job Demands, Job Decision Latitude, and Mental Strain: Implications for Job Redesign. *Administrative Science Quarterly*. 1979;24(2):285–308.
38. Karasek R, Baker D, Marxer F, Ahlbom A, Theorell T. Job decision latitude, job demands, and cardiovascular disease: a prospective study of Swedish men. *Am J Public Health*. 1981;71(7):694–705.
39. Theorell T, Karasek RA, Eneroth P. Job strain variations in relation to plasma testosterone fluctuations in working men - a longitudinal study. *Journal of Internal Medicine*. 1990;227(1):31–6.

40. Deng G, Huang C, Cheung SP, Zhu S. Job demands and resources and employee well-being in the Chinese nonprofit sector. *Front Psychol.* 2021;12:780718.
41. Pedersen P, Momsen AMH, Andersen DR, Nielsen CV, Nohr EA, Maimburg RD. Associations between work environment, health status and sick leave among pregnant employees. *Scand J Public Health.* 2021;49(2):149–58.
42. van Rhenen W, Schaufeli WB, van Dijk FJH, Blonk RWB. Coping and sickness absence. *Int Arch Occup Environ Health.* 2008;81(4):461–72.
43. Leach L, Milner A, Too LS, Butterworth P. Poor psychosocial job conditions increase sickness absence: evidence from the PATH Through Life Mid-Aged Cohort. *BMJ Open.* 2022;12(9):e059572.

## Article n°6

Psychosocial well-being as a protective factor against short- and long-term work absenteeism: from censored-data survival models to multistate analysis

## **Article n°6**

### Psychosocial well-being as a protective factor against short- and long-term work absenteeism: from censored-data survival models to multistate analysis

Colin-Chevalier R, Dutheil F, Cambier S, Dewavrin S, Cornet T, Baker JS, Pereira B

#### **Résumé**

Dans ce sixième et dernier article, nous avons souhaité poursuivre l'étude des liens entre bien-être psychosocial et absentéisme, en différenciant les absences de courte durée (sept jours ou moins) de celles de longue durée (plus de sept jours), dans une approche non plus transversale mais longitudinale. Parmi les approches possibles, le modèle de survie l'un des plus prisé des chercheurs. Le modèle permet d'évaluer à la fois le temps avant qu'un individu ne s'absente du travail et les facteurs pouvant influencer l'absence. Bien que très attrayant, le modèle présente néanmoins plusieurs contraintes telles que la limitation à deux états et à une unique transition (d'un état de vie à un état de mort) ou la non-prise en compte de la durée passée dans l'état de fin. Le modèle multi-états, qui répond à toutes les limites du modèle de survie, apparaît comme une alternative ou comme un complément idéal pour étudier le cas particulier de l'absentéisme au travail. Une fois de plus confrontés à des défis méthodologiques, comme la présence de clusters naturels de données et la nécessité de prendre en compte des effets de clusters (tels que l'« effet entreprise » et l'« effet individu »), ainsi que l'absence de certaines données nécessitant l'application d'une méthode d'imputation multiple prenant en compte ces effets de clusters, nous nous sommes employés à les résoudre avant toute analyse. Dès lors, dans une première approche, nous avons ajusté des modèles de survie, à évènement unique se limitant à la première absence uniquement et à évènements multiples basé sur toutes les absences des travailleurs, pour mesurer l'évolution de risque d'absentéisme de court ou long terme au cours du temps et l'influence du ressenti psychosocial et du bien-être psychosocial sur ce risque. Dans une seconde approche, nous avons ajusté un modèle multi-états pour mesurer la probabilité des

travailleurs de transiter entre les trois états « en poste », « absent pour une courte période », et « absent pour une longue période » et l'effet putatif du ressenti/bien-être psychosocial sur le passage d'un état à un autre. Les deux approches ont permis de mettre en évidence l'influence du bien-être psychosocial sur l'absentéisme au travail et de nous renseigner sur le risque d'absentéisme encourus par les travailleurs, de courte comme de longue durée. Nos résultats démontrent l'importance cruciale du bien-être psychosocial sur la santé des travailleurs et la nécessité pour les employeurs de veiller à le garantir à leurs salariés. Si le modèle de survie demeure un outil pertinent pour étudier l'absentéisme au travail et son évolution dans le temps, le modèle multi-états devrait être plus largement considéré par les chercheurs car il modélise bien mieux le phénomène.



# **Psychosocial well-being as a protective factor against short- and long-term work absenteeism: from censored-data survival models to multistate analysis**

Rémi Colin-Chevalier,<sup>1\*</sup> Frédéric Dutheil,<sup>1</sup> Sébastien Cambier,<sup>2</sup> Samuel Dewavrin,<sup>3</sup> Thomas Cornet,<sup>3</sup> Julien Steven Baker,<sup>4</sup> Bruno Pereira<sup>2</sup>

<sup>1</sup> Université Clermont Auvergne, CNRS, LaPSCo, Physiological and Psychosocial Stress, CHU Clermont-Ferrand, Occupational and Environmental Medicine, Wittyfit, F-63000 Clermont-Ferrand, France

<sup>2</sup> CHU Clermont-Ferrand, Biostatistics Unit, DRCI, F-63000 Clermont-Ferrand, France

<sup>3</sup> Wittyfit, Cegid, F-69009 Lyon, France

<sup>4</sup> Hong Kong Baptist University, Centre for Health and Exercise Science Research, Kowloon Tong, Hong Kong 999077, China

## **\* Corresponding author**

Rémi Colin, PhD Student

Occupational and Environmental Medicine, CHU Clermont-Ferrand, 58 rue Montalembert, F-63000 Clermont-Ferrand, France

Phone: +33 (0)4 73 26 56 75

Email: [r.colin6374@gmail.com](mailto:r.colin6374@gmail.com)

Word count for abstract: 266; Number of tables: 1; Number of figures: 5

## Abstract

**Background:** The influence of psychosocial well-being on work absenteeism remains inadequately explored. Using survival models to study this relationship over time is an effective but limited approach given the complex nature of absenteeism, particularly its duration. The multi-state model seems to be an interesting alternative.

**Objective:** To assess the influence of workers psychosocial feelings/well-being on short- and long-term work absenteeism using a survival then multi-state approach.

**Methods:** Worker absenteeism records were collected from three companies and categorized into whether it was short ( $\leq 7$  days) or long ( $> 7$  days). Psychosocial feelings were assessed using the Psychosocial Well-being Index based on 10 Wittyfit indicators. Univariate and adjusted survival (single- and multiple-event) and multi-state models were fitted to examine the putative effect of psychosocial well-being on work absenteeism.

**Results:** 5,745 workers using Wittyfit (mean age:  $40.5 \pm 10.6$  years, mean seniority:  $9.0 \pm 8.9$  years, 54.0% of women, 7.9% of managers) were included. In total, 45.6% of the working population experienced at least one episode of short-term absenteeism and 19.5% a long-term episode. On average, each day, 92.2% of workers were at work, 5.6% were absent for a short period, and 2.2% were absent for a long period. High psychosocial feelings, as well as psychosocial well-being, were associated with lower risks of short- and long-term absenteeism. The multi-state model provided a more in-depth understanding of absenteeism, notably by measuring more transitions between states than the survival model.

**Conclusions:** Psychosocial well-being appears to protect workers from absenteeism. Survival and multi-state analysis make it possible to study absenteeism, but the multi-state model seems more complete and more suitable, overcoming the constraints of the survival model.

**Trial registration:** Clinicaltrials.gov: NCT02596737.

**Keywords:** psychosocial well-being; work absenteeism, survival model; multistate model

## Introduction

Workplace absenteeism refers to the “unplanned nonattendance of an individual at his or her workstation” [1]. Absenteeism quickly became a major problem for organizations due to the significant risk it poses to employee health and loss of productivity [2]. Protecting workers from the risk of work absenteeism is essential to keeping them healthy [3], which is an employer’s duty. Most of the factors contributing to absenteeism from work belong to the family of psychosocial factors [4–6]. Yet, psychosocial well-being, which is a unique and inclusive concept encompassing multiple dimensions of well-being [7], remains poorly studied in the general working population [8,9]. It could therefore be worthwhile to observe the impact of psychosocial well-being both on the onset of an absence and on its duration. As it is known that certain factors tend to cause both short- and long-term absences [10], our hypothesis is that workers expressing high psychosocial feelings, or in other words a state of psychosocial well-being rather than discomfort, should be less exposed to the risk of absence, whatever the duration. Other sociodemographic factors known to influence absenteeism, such as age [11], seniority [12], and gender [13], could also mediate this relationship.

Because of its complex nature, the literature on work absenteeism is replete with different approaches [14]. Among these approaches, survival analysis has emerged as one of the most efficient, as it enables to study the occurrence of an event over time, such as the absences of a worker [15–18]. However, there are three major problems with using a survival model to study absenteeism at work. First, the model is limited to only two states (life and death), which requires either applying a model per end state or turning to a more complex model such as the competing risks survival model. Secondly, the model is limited to the study of a single event or transition (from a state of life to a state of death), which does not allow multiple events per individual to be considered. Unlike a singular event like death, for example, workers may experience multiple episodes of work absenteeism over time. Thirdly, while the model measures the probability of an event occurring, it attaches no importance to the duration of that

event. For absenteeism, this constraint appears restrictive since workers may experience longer or shorter interruptions in their activity. To overcome these issues, the survival model would then have to be based on all worker absences and no longer be limited to individual's first absence, but this does make the interpretation of results more complex, even impossible. The multi-state model approach then appears to be the perfect tool, as a replacement or a complement to survival analysis, for analyzing absenteeism data [19,20]. Indeed, the approach makes it possible to model a process in which individuals can move from one state to another, for example, between periods when workers are at work and periods when they are absent, and vice versa, in addition to measuring the time spent by individuals in each state.

Thus, the main objective of this study was to evaluate the influence of psychosocial well-being (or discomfort) on both short- and long-term work absenteeism using two different approaches, first with survival models, one based on the worker's first absenteeism event only (single-event approach) and the other with all the worker's absenteeism periods (multiple-event approach), and then with a multi-state.

## **Methods**

### **Study design and recruitment**

The web-based application Wittyfit (Cegid) was developed in collaboration with the university hospital of Clermont-Ferrand. It is widely used by companies in various industries in France to assess the well-being and psychosocial risks of their employees, aiming to enhance overall performance [21]. Participation in Wittyfit is voluntary, and it allows users to anonymously express their thoughts and feelings by completing health-related surveys, which contribute to evaluating their general health status. The software provides personalized feedback based on the gathered information. Data collection for this study was conducted in the year 2022. Absenteeism data was obtained from employees of three companies using Wittyfit who volunteered to take part in the study and who had participated in Wittyfit surveys at least once

during that year, whether the surveys were completed fully or partially. Prior to their participation, all subjects provided informed consent, following the terms and conditions of Wittyfit. The study received approval from two regulatory bodies, namely the French data protection authority (CNIL) and the Ethics Committee of Clermont-Ferrand (South-East IV) (Clinicaltrials.gov: NCT02596737).

## **Outcomes**

**Work absenteeism:** Over the year, when an employee was absent from work for a reason not related to work (i.e., other than scheduled vacation, a training period, a day off or an occupational illness), this was considered a period of absenteeism from work. When a worker was absent for a period of seven days or less, we considered this to be a short-term period of absenteeism, otherwise a long-term period. This threshold was chosen in accordance with French law that permit employees to get daily allowances for work stoppage higher than seven days of sick leave. This duration of absenteeism also previously required an occupational health consultation for occupational accidents. Outside his periods of absenteeism, a worker was supposed to be at work. If a worker left his or her company before the end of the year, we considered the data right censored at the time of his or her departure.

**Psychosocial feelings, well-being, and discomfort:** Workers' psychosocial feelings was measured using the Psychosocial Well-being Index and normalized, ranging from zero to one [8]. All psychosocial factors necessary for its assessment, including job satisfaction, atmosphere, recognition, work-life balance, meaning, work organization, values, workload, autonomy, and stress, were assessed in the Wittyfit platform using visual analog scales (VAS) [22–24]. The VAS of Wittyfit are graduated from zero to 100, with zero indicating a low level and 100 a high level. The average score over the year was retained for each indicator. If a worker has never evaluated an indicator during the year, it was considered a missing value. Initially ranging from -54.6 to 40.0, the psychosocial well-being index has been rescaled from

zero to one. To determine whether a worker was in a state of psychosocial well-being or discomfort, we considered that if his/her psychosocial feelings was above the median level of the population, he/she was in a state of psychosocial well-being, otherwise in a state of discomfort.

**Sociodemographic:** Age, seniority, gender, and job position of workers were given by companies using Wittyfit. The job position of a worker could be either employee or manager, where a manager is defined by someone who must assume the overseeing of other workers, namely the employees.

## **Statistics**

Quantitative data such as age, seniority, and psychosocial feelings of workers were presented as mean  $\pm$  standard deviation. Qualitative data, including the proportion of men and women, the proportion of employees and managers, the prevalence of absenteeism, the proportion of workers in a state of psychosocial well-being or discomfort, and the proportion of workers who left their company during the study year, were reported as the number of individuals and the corresponding percentage. In a first survival models approach, involving single-event models based on first worker absences only and multiple-event models based on all worker absences, we separately evaluated the survival time to short- and long-term work absenteeism using Kaplan-Meier method [22]. We then fitted Cox proportional hazards models to assess the influence of psychosocial feelings, as a quantitative or qualitative variable, on survival time [23]. All Cox proportional hazards models were adjusted twice, first univariately, then adjusted for age, seniority, gender, and job position. In a second multi-state model approach, we evaluated the transition probabilities between the three possible states in which workers could evolve. The three states were: “at work” (state 1), “absent for a short period” ( $\leq 7$  days, state 2), or “absent for a long period” ( $> 7$  days, state 3). For each model, transition probability matrix and average sojourn times (mean time spent by an individual in a state) were calculated with

95CI. As with the survival models, we then evaluated the influence on psychosocial feelings, as a quantitative or qualitative variable, on the probability of moving from one state to another, in a univariate then adjusted model. Since data on psychosocial factors were missing completely at random [24], hierarchical multiple imputation was conducted using the “mice” command from the R package of the same name. Individual and company were treated as random effects and age, seniority, gender, and job position as predictive factors to mitigate bias resulting from complete-cases analysis [25,26]. All models’ estimates were turned into hazard ratios (HR), either crude for univariate models or adjusted for multivariate models, with their 95% confidence interval (95CI). All statistical analyses were carried out using R (version 4.3.1) [27]. Unless specified, a p-value <0.05 was considered as statistically significant.

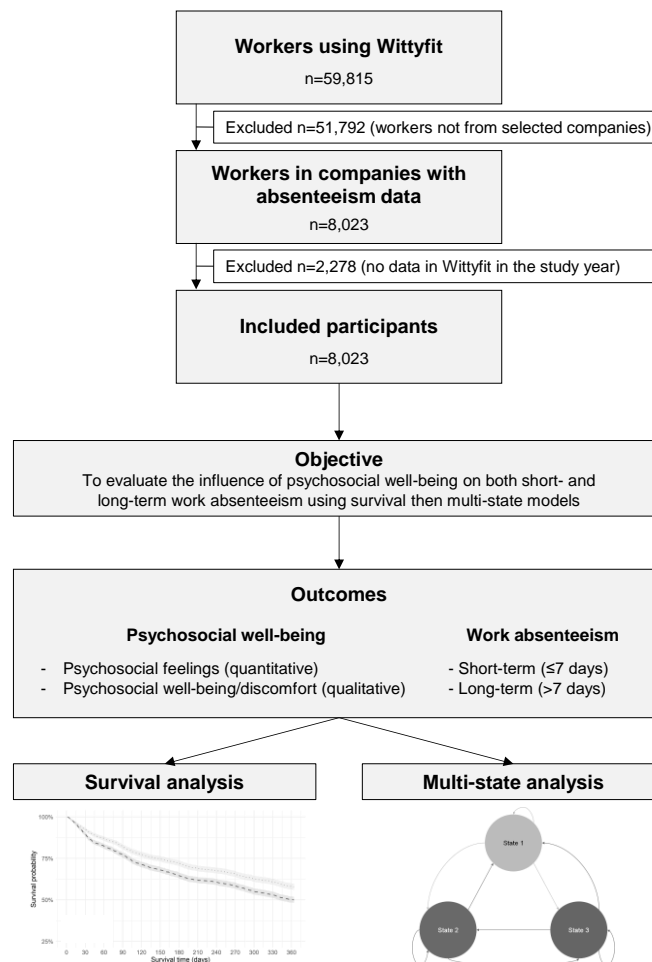
## **Results**

### **Participants**

Data were collected from 5,745 workers from the three companies from the sector of activity of the sale of goods and services, of sizes 1,578, 1,598, and 2,569, with an average age of  $40.5 \pm 10.6$  years old, an average seniority of  $9.0 \pm 8.9$  years. The sample includes 3,104 women (54.0% of the population) and 2,641 men (46.0%), whom 5,290 are employees (92.1%) and 455 are managers (7.9%). Workers’ psychosocial feelings was  $0.63 \pm 0.21$  on average, with a median of 0.65. During the study year, 2,922 workers (50.9%) had at least one period of absenteeism, whether short or long. More precisely, 2,618 workers (45.6%) had at least one period of short-term absenteeism, 1,123 (19.5%) a period of long-term absenteeism, and 819 (14.3%) had both. In the population, there were a total of 6,206 short-term absences for an average number of absences per individual of 1.08, and a total of 2,385 long-term absences for an average number of 0.42 absences per individual. Every day, on average, 92.2% of the population was at work, 5.6% was absent for a short period and 2.2% for a long period. During



the year, 271 workers (4.7%) left their company (Figure 1). More information on worker absenteeism can be found in supplementary materials (Figure S1, Figure S2).



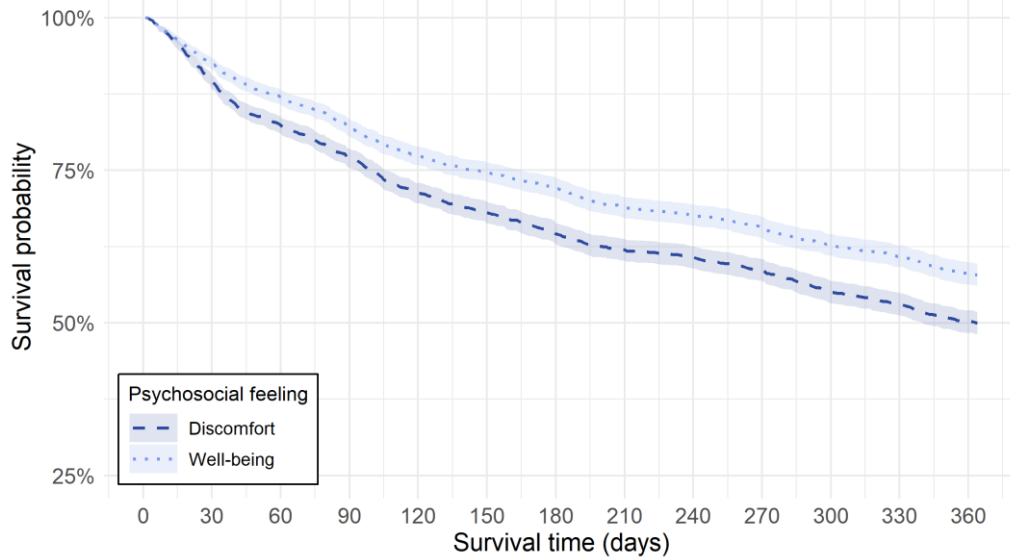
**Figure 1.** Flowchart of the study.

Legend: ‘n’ sample size.

### Survival time to work absenteeism

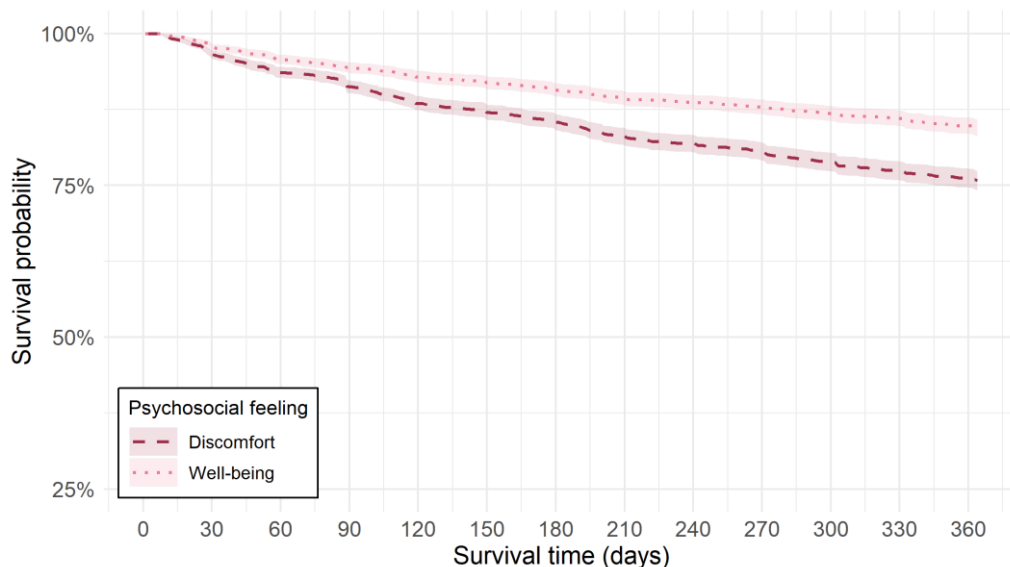
According to single-event survival models, at the end of the year, 53.9% (95CI 25.7 to 55.3%) of workers survived an event of short-term absenteeism and 80.1 % (79.1 to 81.2%) a long-term absenteeism event. Overall, survival probability of short-term absenteeism decreased by  $0.13 \pm 0.11\%$  each day, and that of long-term absenteeism by  $0.06 \pm 0.06\%$ . High psychosocial feelings were associated with lower odds of absenteeism, whether short- (crude HR=0.45, 95CI 0.37 to 0.54,  $p < 0.001$ ; adjusted HR=0.42, 0.35 to 0.51,  $p < 0.001$ ) or even more long-term (0.26, 0.19 to 0.34,  $p < 0.001$ ; 0.24, 0.18 to 0.32,  $p < 0.001$ ). Workers expressing a state of psychosocial

well-being were less likely to be absent, both for a short duration (0.78, 0.72 to 0.84,  $p < 0.001$ ; 0.76, 0.70 to 0.82,  $p < 0.001$ ) and more for a long duration (0.26, 0.19 to 0.34,  $p < 0.001$ ; 0.24, 0.18 to 0.32,  $p < 0.001$ ) (Figures 2 and 3).



**Figure 2.** Evolution of the probability of surviving short-term absenteeism (simple-event survival model).

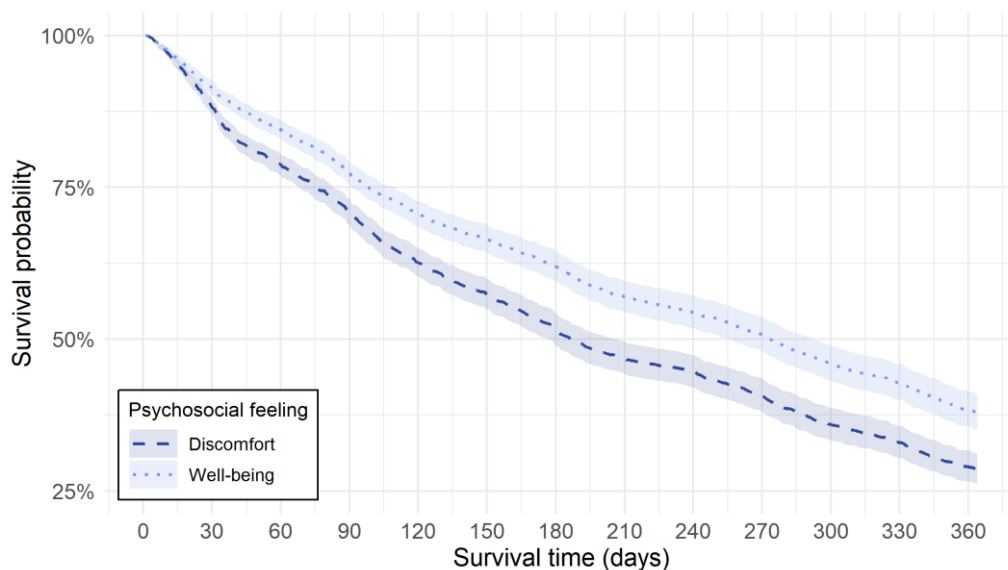
Legend: Kaplan-Meier survival curves illustrate the probability of workers surviving short-term absenteeism event in a state of psychosocial well-being (dotted line) or discomfort (dashed line).



**Figure 3.** Evolution of the probability of surviving long-term absenteeism (simple-event survival model).

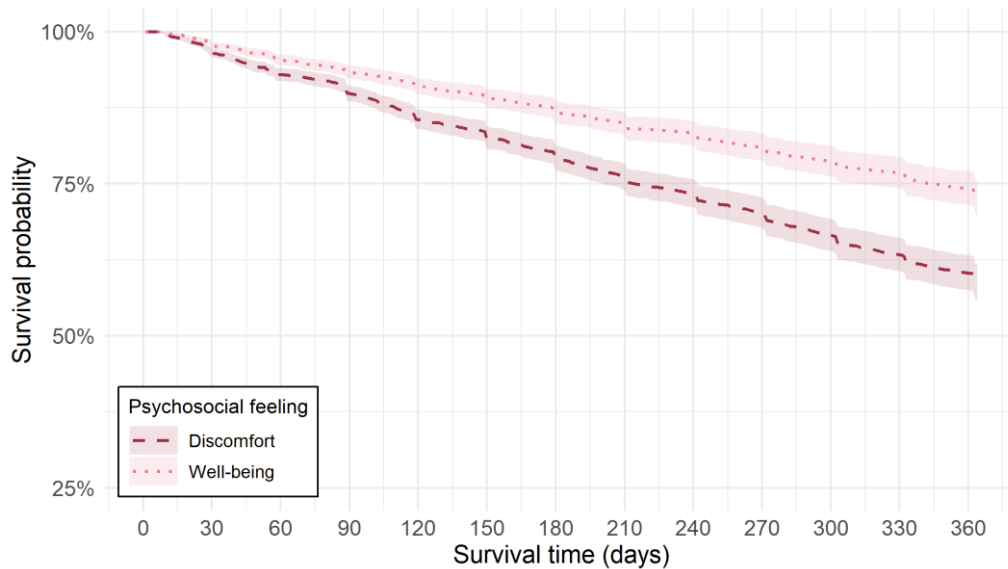
Legend: Kaplan-Meier survival curves illustrate the probability of workers surviving long-term absenteeism event in a state of psychosocial well-being (dotted line) or discomfort (dashed line).

Considering all worker absences combined, the probability of surviving short-term absenteeism event given by the multi-event survival model was reduced to 32.9% (95CI 31.0 to 34.9%), and that of long-term absenteeism to 65.2% (63.1 to 67.3%). Compared to the first model, the average daily survival probability of absenteeism decreased more rapidly, falling to  $0.19 \pm 0.13\%$  for short-term absenteeism, and to  $0.10 \pm 0.15\%$  for long-term absenteeism. As shown with single event survival models, psychosocial feelings of workers were found to reduce the risk of short (crude HR=0.45, 95CI 0.40 to 0.50,  $p < 0.001$ ; adjusted HR=0.42, 0.37 to 0.47,  $p < 0.001$ ) and even more long period absence (0.23, 0.19 to 0.28,  $p < 0.001$ ; 0.21, 0.18 to 0.26,  $p < 0.001$ ). Again, workers who expressed a state of psychosocial well-being proved to be less exposed to the risk of short-term absenteeism (0.77, 0.74, 0.81,  $p < 0.001$ ; 0.75, 0.71 to 0.79,  $p < 0.001$ ) and even less exposed to the risk of long-term absenteeism (0.23, 0.19 to 0.28,  $p < 0.001$ ; 0.21, 0.18 to 0.26,  $p < 0.001$ ) (Figures 4 and 5).



**Figure 4.** Evolution of the probability of surviving short-term absenteeism (multiple-event survival model).

Legend: Kaplan-Meier survival curves illustrate the probability of workers surviving short-term absenteeism event in a state of psychosocial well-being (dotted line) or discomfort (dashed line).

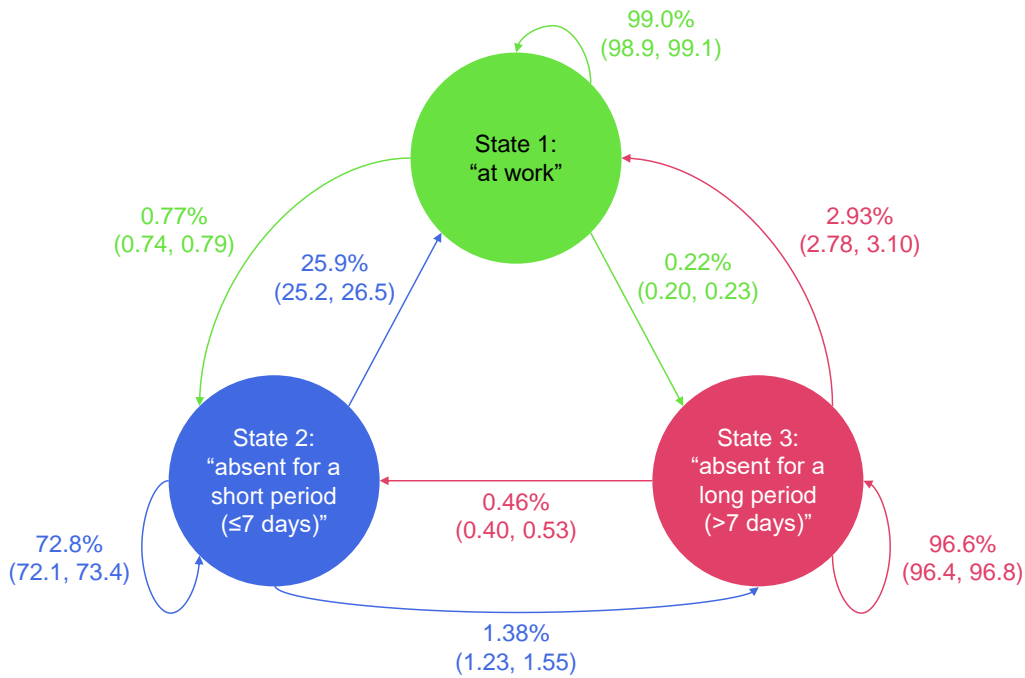


**Figure 5.** Evolution of the probability of surviving long-term absenteeism (multiple-event survival model).

Legend: Kaplan-Meier survival curves illustrate the probability of workers surviving long-term absenteeism event in a state of psychosocial well-being (dotted line) or discomfort (dashed line).

### Multi-state analysis of work absenteeism

When they were in state 1 during a day (i.e., at work), workers had a very low risk to move to state 2 (i.e., absent for a short time; 0.77%, 95CI 0.74 to 0.79%), and even lower to move to state 3 (i.e., absent for a long time; 0.22%, 0.20 to 0.23%) the next day. When in state 2, workers had a 25.9% (25.2 to 26.5%) chance to move to state 1, and a very low risk of moving to state 3 (1.38, 1.23 to 1.55%). In state 3, the probability of moving to state 1 was 2.93% (2.78 to 3.10%) compared to 0.46% (0.40 to 0.53%) of moving to state 2. Mean sojourn time was 89.7 (95CI 87.4 to 91.9) days in state 1, 3.13 (3.04 to 3.22) days in state 2, and 28.9 (27.4 to 30.5) days in state 3 (Figure 6).



**Figure 6.** Estimated time-to-time transition probabilities between states.

Legend: Transition probabilities between states are presented with 95% confidence intervals.

It was found that the psychosocial feelings of workers reduced the risk of being absent one day when they were at work the day before. In fact, workers with high psychosocial feelings were less likely to move from state 1 to state 2 (crude HR=0.85, 95CI 0.75 to 0.97, p=0.007; adjusted HR=0.83, 0.73 to 0.95, p=0.002), and almost half as likely to move to state 3 (0.44, 0.34 to 0.58, p<0.001; 0.43, 0.33 to 0.57, p<0.001). Surprisingly, workers who reported a state of psychosocial well-being were no less likely to move from state 1 to state 2 (0.98, 0.93 to 1.04, p=0.3; 0.97, 0.91 to 1.03, p=0.1). However, there were less likely to move from state 1 to state 3 (0.72, 0.64 to 0.81, p<0.001; 0.71, 0.63 to 0.80, p<0.001) (Table 1).

**Table 1.** Estimated covariate effect of psychosocial feelings and psychosocial well-being on transition intensities between states.

Transition between states (from → to)	Effect of psychosocial feelings/well-being on transition intensity (hazard ratio and 95% confidence interval)			
	Psychosocial feelings (quantitative variable)		Psychosocial well-being (qualitative variable)	
	Univariate model (crude)	Multivariate model (adjusted)	Univariate model (crude)	Multivariate model (adjusted)
1 → 2	<b>0.85 (0.75 to 0.97)</b>	<b>0.83 (0.73 to 0.95)</b>	0.98 (0.93 to 1.04)	0.97 (0.91 to 1.03)
1 → 3	<b>0.44 (0.34 to 0.58)</b>	<b>0.43 (0.33 to 0.57)</b>	<b>0.72 (0.64 to 0.81)</b>	<b>0.71 (0.63 to 0.80)</b>
2 → 1	1.05 (0.92 to 1.19)	1.01 (0.89 to 1.15)	1.02 (0.96 to 1.09)	1.01 (0.95 to 1.07)
2 → 3	0.75 (0.44 to 1.29)	0.74 (0.42 to 1.27)	0.99 (0.78 to 1.26)	0.98 (0.77 to 1.25)
3 → 1	1.03 (0.80 to 1.33)	1.05 (0.81 to 1.36)	0.96 (0.85 to 1.09)	0.97 (0.86 to 1.10)
3 → 2	1.49 (0.82 to 2.69)	1.57 (0.86 to 2.83)	1.23 (0.94 to 1.61)	1.24 (0.94 to 1.62)

Legend: A bold result means that the influence of the covariate is significant. States: ‘1’=“at work”, ‘2’=“absent for a short period”, ‘3’=“absent for a long period”.

## Discussion

The main results showed that when they expressed high psychosocial feelings, or in other words when they were in a state of psychosocial well-being, workers were at less risk of work absenteeism, whether for a short or long term. The two approaches considered, namely survival and multi-state analyses, gave essentially the same results.

### The protective effect of psychosocial well-being on work absenteeism

Psychological well-being is an essential characteristic of mental health, largely associated with happiness [28]. When we know that happiness is associated with lower risks of absenteeism [29], it is not surprising to find the same trend with psychosocial well-being. This is all the truer when we know that the need to be absent can be influenced by work-related factors [30]. In this article, we found that high psychosocial feelings tended to reduce the risk of absenteeism. In both approaches, it was found that the higher the psychosocial feelings of workers, the lower

their risks of absenteeism, both short-term and even more long-term. Research has shown that the psychosocial working environment is strongly associated with absenteeism, both for short- [31] and long-term events [32]. Therefore, our results seem to confirm that psychosocial well-being, like many work-related factors [10], may be at the root of both short- and long-term work absenteeism. Yet, when studied through the lens of psychosocial well-being/discomfort, the results held true, except for the association with short-term work absenteeism in the multi-state analysis, which did not turn to be significant. Previous studies examining the links between psychosocial factors and absenteeism have shown that this association is stronger for long-term absenteeism, which is consistent with our results [33,34]. The multi-state analysis revealed that although psychosocial well-being could reduce the risk of absenteeism, it did not increase the chances of returning to work. Psychosocial well-being is mainly work-related, which means that employers can only work on improving that of their employees when they are present in the workplace. From this perspective, it is necessary that employers work constantly to ensure that their employees work in a healthy working environment [35]. We suggest for example that employers focus on prevention, recognized as an effective tool for improving workers well-being [36].

### **A dual modelling of the evolution of work absenteeism over time**

The study of absenteeism from work using survival analysis has become particularly widespread over the last years [15,17,18]. This is particularly evident in studies that are limited to the evaluation of the time elapsed before a first absence, considered as a definitive event, because this gives results equivalent to standard Poisson regression models [16]. The single-event model we first considered was designed in this way, making it possible to model the daily evolution of the probability of surviving a first event of absenteeism, either of short or long duration. Yet, since work absenteeism is not an absorbing event from which workers cannot return, we decided to consider a multiple-event model in our survival approach, extended to all



periods of worker absenteeism rather than just the first period. However, this method has the disadvantage of making it impossible to interpret the survival curve, as it does not consider the possibility of an individual returning to its initial state (at work) after the occurrence of an event (an absenteeism event), so that the probability of survival can only decrease. Still, despite the drawbacks of the multiple-event model, Cox proportional hazards models gave the same results as the single-event model, with essentially the same effect sizes. Survival models can be used to assess the risk of work absenteeism and the putative influence of covariates, but do not consider changes in state other than the transition from the working state to a state of work absenteeism. Nor do they measure the time spent by individuals in a state of absenteeism or evaluate the risk of repeated absence for workers. These constraints limit the ability of survival models to analyze a repeated phenomenon such as absenteeism from work, and the use of the multi-state model becomes relevant [20]. Unlike the first approach, the model made it possible to estimate that workers had approximately 1% of becoming absent each day (0.77% of becoming absent for a short period, 0.22% for a long one), and spent on average 3.13 (respectively 28.9) days convalescing in the event of a short (respectively long) period of absenteeism. This suggests that the underlying reasons for short absences are rarer than for long absences [10]. The model also showed that the risk of repeated absence of workers was not zero, suggesting that an individual's absenteeism episodes may be interrelated, and therefore share common causes, as already shown for sickness absences [37–40]. Overall, the sole application of the multi-state model made it possible to find results close to those of the survival models in addition to providing additional information. It appears that, as expected, the multi-state model allows better modeling of absenteeism. This approach should be more widely considered by researchers wishing to study absenteeism at work.

## **Limitations**

We acknowledge some limitations within the study. A limitation of the current research is related to the sample consisting only French workers. Data only comes from three companies, and French companies can be impacted to a greater or lesser extent by absenteeism. Similarly, the risk of absenteeism may be different depending on the occupation studied. With over 5,000 workers included in the study, we can expect a variety of different occupations to have been represented, reducing the risk of potential biases. Nevertheless, it would be interesting to add other countries to this study, as we would expect to observe different rates of absenteeism in different countries [42]. Using self-reported responses to assess how workers felt about different psychosocial factors may have led to potential bias. As all responses were collected through surveys within the Wittyfit platform, where volunteers were able to express their feelings anonymously and whenever they wished [21], we were able to control for the impact of these two biases. By treating each absence equally, we were unable to measure the effect of psychosocial well-being according to the type of absence (illness, family event, etc.). Moreover, one would expect the reason of an absence and its duration to be related. However, an absence, whatever the reason, can be due to psychosocial problems. This is particularly true of absences due to illness [43,32], but also applies to accidents [44] or unjustified absences [45], for example. Future studies on the subject will focus on specific reasons for absences to have a better understanding of the mechanisms between absenteeism at work and psychosocial factors.

## **Conclusions**

Psychosocial well-being helps protect workers against short- and long-term absenteeism. It is essential for employers to ensure that their employees are in a state of psychosocial well-being. Survival or multi-state approaches have proved to be relevant and effective tools for studying absenteeism at work. However, the characteristics of the survival model, such as the limited number of states and transitions, may prove limiting for the study of a phenomenon such as

absenteeism, where being absent is not necessarily a definitive state from which workers cannot return, and the multi-state model appears to be a perfect alternative.

## **Acknowledgements**

We express our sincere gratitude to all voluntary workers using Wittyfit, who participated in this study.

## **Conflicts of interest**

RC, SD, and TC are part of Cegid. Other authors have declared that no competing interests exist (FD is responsible for the scientific accuracy of Wittyfit but is not paid by Cegid; as previously published, Wittyfit is a public private partnership with the CHU Clermont-Ferrand).

## **Authors' contributions**

Study design: FD, SD.

Study conduct: RC, FD, BP.

Data collection: SD, TC.

Data analysis: RC, BP.

Data interpretation: RC, SC, BP.

Drafting manuscript: RC.

Revising manuscript content: RC, FD, SC, JSB, BP.

Approving final version of manuscript: all authors.

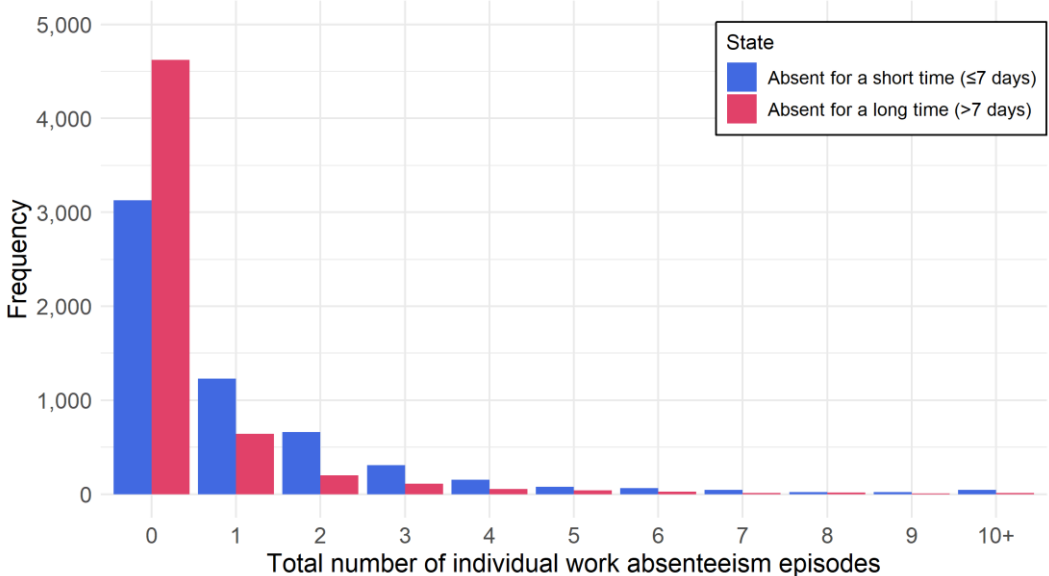
RC, BP, FD take responsibility for the integrity of the data analysis.

## **Funding**

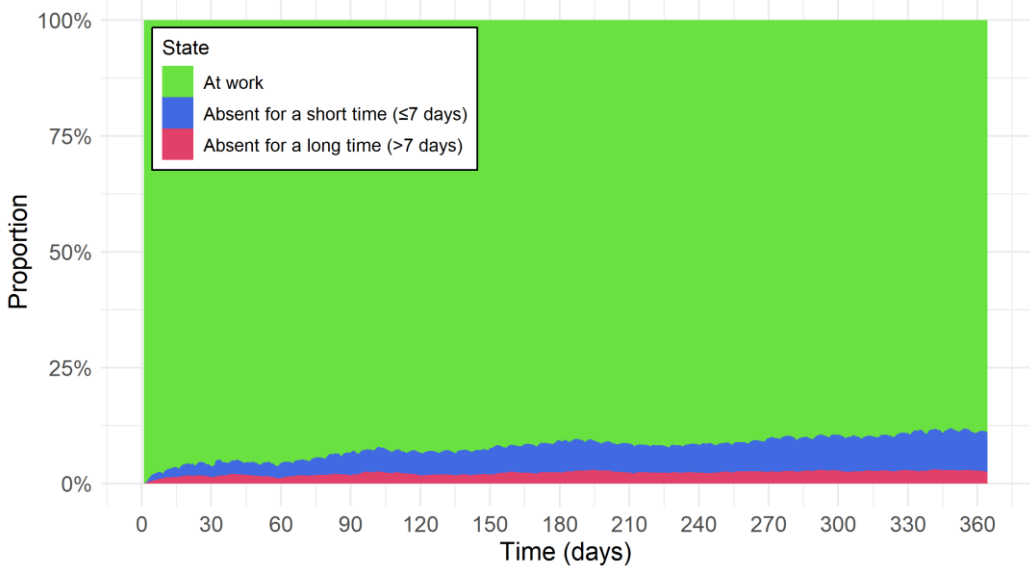
This research received no external funding.

# Supplementary materials

Additional information on the distribution of periods of work absenteeism (Figure S1) and on the evolution of the distribution of states in workers, namely “at work”, “absent for a short time ( $\leq 7$  days)”, and “absent for a long time ( $>7$  days)” (Figure S2), are presented below.



**Figure S1.** Distribution of the total number of work absenteeism episodes per individual. Legend: ‘10+’ more than 10 periods of absenteeism from work.



**Figure S2.** Evolution of the distribution of states. Legend: Colored areas represent the daily proportion of workers in the population of each state at any given time.

## References

1. Patton E. The devil is in the details: Judgments of responsibility and absenteeism from work. *Journal of Occupational and Organizational Psychology*. 2011;84(4):759–79.
2. Nagata T, Mori K, Ohtani M, Nagata M, Kajiki S, Fujino Y, et al. Total Health-Related Costs Due to Absenteeism, Presenteeism, and Medical and Pharmaceutical Expenses in Japanese Employers. *J Occup Environ Med*. 2018;60(5):e273–80.
3. Johnston DA, Harvey SB, Glozier N, Calvo RA, Christensen H, Deady M. The relationship between depression symptoms, absenteeism and presenteeism. *Journal of Affective Disorders*. 2019;256:536–40.
4. Burmeister EA, Kalisch BJ, Xie B, Doumit MAA, Lee E, Ferraresion A, et al. Determinants of nurse absenteeism and intent to leave: An international study. *J Nurs Manag*. 2019;27(1):143–53.
5. Millot C, Pereira B, Miallaret S, Clinchamps M, Vialatte L, Guillin A, et al. Statistical analysis of absenteeism in a university hospital center between 2007 and 2019. *Int J Environ Res Public Health*. 2022;19(19):12966.
6. Kagan I, Hendel T, Savitsky B. Personal initiative and work environment as predictors of job satisfaction among nurses: cross-sectional study. *BMC Nurs*. 2021;20:87.
7. Eiroa-Orosa FJ. Understanding psychosocial wellbeing in the context of complex and multidimensional problems. *Int J Environ Res Public Health*. 2020;17(16):5937.
8. Colin-Chevalier R, Pereira B, Dewavrin S, Cornet T, Baker J, Dutheil F. Psychosocial well-being index and sick leave in the workplace: a structural equation modelling of Wittyfit data (Preprint). 2023;
9. Jeong W, Kim YK, Oh SS, Yoon JH, Park EC. Association between presenteeism/absenteeism and well-being among Korean workers. *J Occup Environ Med*. 2020;62(8):574–80.

10. Blank N, Diderichsen F. Short-term and long-term sick-leave in Sweden: relationships with social circumstances, working conditions and gender. *Scand J Soc Med.* 1995;23(4):265–72.
11. Taimela S, Läärä E, Malmivaara A, Tiekso J, Sintonen H, Justén S, et al. Self-reported health problems and sickness absence in different age groups predominantly engaged in physical work. *Occup Environ Med.* 2007;64(11):739–46.
12. Grigore OM. Factors Contributing to Work-Related Absenteeism during the COVID-19 Pandemic. *Management Dynamics in the Knowledge Economy.* 2020;8(4):401–18.
13. Mastekaasa A, Melsom AM. Occupational segregation and gender differences in sickness absence: evidence from 17 European countries. *European Sociological Review.* 2014;30(5):582–94.
14. Duchemin T, Hocine MN. Modeling sickness absence data: A scoping review. *PLoS One.* 2020;15(9):e0238981.
15. van den Heuvel SG, Boshuizen HC, Hildebrandt VH, Blatter BM, Ariëns GA, Bongers PM. Effect of sporting activity on absenteeism in a working population. *Br J Sports Med.* 2005;39(3):e15.
16. Christensen KB, Andersen PK, Smith-Hansen L, Nielsen ML, Kristensen TS. Analyzing sickness absence with statistical models for survival data. *Scand J Work Environ Health.* 2007;33(3):233–9.
17. Watkins JP, Ellis ED, Girardi DJ, Cragle DL, Richter BS. Illness absences among beryllium sensitized workers. *Am J Public Health.* 2014;104(11):e165-169.
18. Roelen CAM, van Hoffen MFA, Waage S, Schaufeli WB, Twisk JWR, Bjorvatn B, et al. Psychosocial work environment and mental health-related long-term sickness absence among nurses. *Int Arch Occup Environ Health.* 2018;91(2):195–203.

19. Lie SA, Eriksen HR, Ursin H, Hagen EM. A multi-state model for sick-leave data applied to a randomized control trial study of low back pain. *Scand J Public Health*. 2008;36(3):279–83.
20. Gran JM, Lie SA, Øyeflaten I, Borgan Ø, Aalen OO. Causal inference in multi-state models-sickness absence and work for 1145 participants after work rehabilitation. *BMC Public Health*. 2015;15:1082.
21. Dutheil F, Duclos M, Naughton G, Dewavrin S, Cornet T, Huguet P, et al. Wittyfit-live your work differently: study protocol for a workplace-delivered health promotion. *JMIR Research Protocols*. 2017;6(4):e6267.
22. Lesage FX, Berjot S, Deschamps F. Clinical stress assessment using a visual analogue scale. *Occup Med (Lond)*. 2012;62(8):600–5.
23. Dutheil F, Trousselard M, Perrier C, Lac G, Chamoux A, Duclos M, et al. Urinary interleukin-8 is a biomarker of stress in emergency physicians, especially with advancing age-the JOBSTRESS\* randomized trial. *PLoS One*. 2013;8(8):e71658.
24. Dutheil F, Pereira B, Moustafa F, Naughton G, Lesage FX, Lambert C. At-risk and intervention thresholds of occupational stress using a visual analogue scale. *PLoS One*. 2017;12(6):e0178948.
25. Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. Second Edition. New York: John Wiley & Sons; 2002.
26. Li P, Stuart EA, Allison DB. Multiple Imputation: A Flexible Tool for Handling Missing Data. *JAMA*. 2015;314(18):1966–7.
27. Colin-Chevalier R, Dutheil F, Cambier S, Dewavrin S, Cornet T, Baker JS, et al. Methodological issues in analyzing real-world longitudinal occupational health data: a useful guide to approaching the topic. *International Journal of Environmental Research and Public Health*. 2022;19(12):7023.



28. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2021.
29. Tang YY, Tang R, Gross JJ. Promoting psychological well-being through an evidence-based mindfulness training program. *Front Hum Neurosci.* 2019;13:237.
30. Chen C, Okubo R, Hagiwara K, Mizumoto T, Nakagawa S, Tabuchi T. The association of positive emotions with absenteeism and presenteeism in Japanese workers. *J Affect Disord.* 2024;344:319–24.
31. Mehlum IS. How much sick leave is work-related? *Tidsskr Nor Laegeforen.* 2011;131(2):122–5.
32. Mathisen J, Nguyen TL, Jensen JH, Mehta AJ, Rugulies R, Rod NH. Impact of hypothetical improvements in the psychosocial work environment on sickness absence rates: a simulation study. *Eur J Public Health.* 2022;32(5):716–22.
33. Thorsen SV, Flyvholm MA, Pedersen J, Bültmann U, Andersen LL, Bjorner JB. Associations between physical and psychosocial work environment factors and sickness absence incidence depend on the lengths of the sickness absence episodes: a prospective study of 27 678 Danish employees. *Occup Environ Med.* 2021;78(1):46–53.
34. Allebeck P, Mastekaasa A. Chapter 5. Risk factors for sick leave - general studies. *Scand J Public Health.* 2004;32(63\_suppl):49–108.
35. Nielsen ML, Rugulies R, Christensen KB, Smith-Hansen L, Kristensen TS. Psychosocial work environment predictors of short and long spells of registered sickness absence during a 2-year follow up. *J Occup Environ Med.* 2006;48(6):591–8.
36. Schütte S, Chastang JF, Malard L, Parent-Thirion A, Vermeylen G, Niedhammer I. Psychosocial working conditions and psychological well-being among employees in 34 European countries. *Int Arch Occup Environ Health.* 2014;87(8):897–907.

37. de Wijn AN, van der Doef MP. Reducing psychosocial risk factors and improving employee well-being in emergency departments: a realist evaluation. *Front Psychol.* 2022;12:728390.
38. Koopmans PC, Roelen CAM, Groothoff JW. Risk of future sickness absence in frequent and long-term absentees. *Occup Med (Lond).* 2008;58(4):268–74.
39. Koopmans PC, Roelen C a. M, Groothoff JW. Frequent and long-term absence as a risk factor for work disability and job termination among employees in the private sector. *Occup Environ Med.* 2008;65(7):494–9.
40. Laaksonen M, He L, Pitkäniemi J. The durations of past sickness absences predict future absence episodes. *J Occup Environ Med.* 2013;55(1):87–92.
41. Notenbomer A, van Rhenen W, Groothoff JW, Roelen CAM. Predicting long-term sickness absence among employees with frequent sickness absence. *Int Arch Occup Environ Health.* 2019;92(4):501–11.
42. Antczak E, Miszczyńska KM. Causes of sickness absenteeism in Europe—analysis from an intercountry and gender perspective. *Int J Environ Res Public Health.* 2021;18(22):11823.
43. Hystad SW, Eid J, Brevik JI. Effects of psychological hardiness, job demands, and job control on sickness absence: a prospective study. *J Occup Health Psychol.* 2011;16(3):265–78.
44. Alali H, Braeckman L, Van Hecke T, De Clercq B, Janssens H, Wahab MA. Relationship between non-standard work arrangements and work-related accident absence in Belgium. *J Occup Health.* 2017;59(2):177–86.
45. Fischer FM. Retrospective study regarding absenteeism among shiftworkers. *Int Arch Occup Environ Health.* 1986;58(4):301–20.

# Discussion générale

# Discussion générale

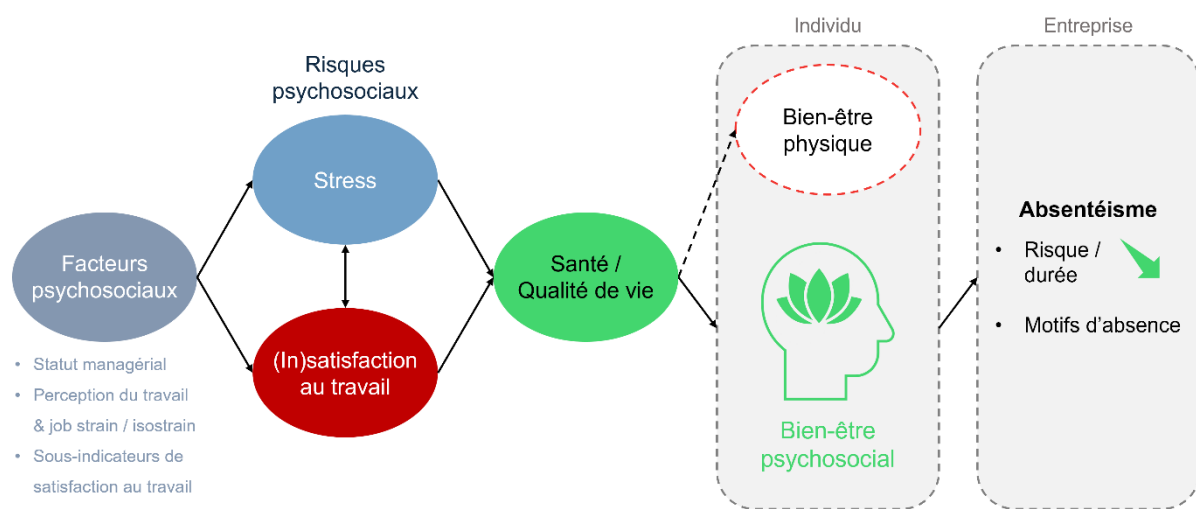
## 1. L'étude des données de santé au travail au service de la lutte contre les risques psychosociaux

Il n'existe pas de solution universelle pour lutter contre les RPS en milieu professionnel. La réalité d'une entreprise peut être très différente de celle d'une autre entreprise, et les travailleurs d'une même organisation peuvent être exposés inégalement à ces risques. Agir efficacement contre les RPS requiert une évaluation préalable des risques encourus par les travailleurs sur leur lieu de travail [123]. Cette évaluation doit tenir compte des spécificités de chaque entreprise comme l'environnement de travail, la nature des activités, ou les caractéristiques des salariés. Elle doit également s'inscrire dans une démarche participative, impliquant à la fois les employeurs et les employés [42]. Dans ce travail de thèse, nous avons analysé les données comportementales des travailleurs qui utilisent la plateforme Wittyfit, plateforme qui permet à ses utilisateurs de contribuer à l'évaluation et à l'amélioration de leur bien-être au travail. Parce qu'elle offre aux entreprises un outil capable d'évaluer les risques encourus par les salariés dans l'exercice de leur activité, et parce qu'elle repose sur la participation active et volontaire de ses utilisateurs, la plateforme permet aux entreprises d'agir contre les RPS.

Pour répondre aux enjeux épidémiologiques et médico-économiques de ce travail de thèse, nous avons mené des études centrées sur les différents indicateurs mesurés par la plateforme. Nos objectifs étaient de démontrer que l'analyse des données de Wittyfit permettait à la fois de quantifier l'exposition aux RPS des travailleurs et utilisateurs de la plateforme, de mesurer l'influence des facteurs psychosociaux engendrant ces risques, et d'étudier leurs liens avec des indicateurs socio-économiques, notamment l'absentéisme. Dans une approche transversale, nous avons mesuré la prévalence d'employés et de managers en situation de job strain ou d'isostrain en utilisant le modèle de Karasek. Ces risques sont connus pour être associés au stress et donc néfastes pour la santé [124–127]. De la même manière, en utilisant des

questionnaires adaptés, nous avons mesuré la prévalence de travailleurs en situation d'insatisfaction au travail selon leur position managériale. Comme le job strain et l'isotrain, l'insatisfaction au travail est un facteur de stress [20,128] susceptible d'avoir des effets délétères pour la santé des travailleurs [20,129]. Quantifier les facteurs de risques au travers du prisme du statut managérial nous a paru pertinent pour deux raisons, à savoir l'importance du rôle de manager en entreprise et le faible nombre d'études menées sur cette population [130,131]. En outre, nous avons constaté que certaines caractéristiques sociodémographiques telles l'âge, l'ancienneté dans la profession, et le sexe (en particulier le fait d'être une femme) pouvaient altérer le ressenti psychosocial des travailleurs. L'analyse longitudinale opérée dans notre quatrième étude a permis de confirmer ces résultats en apportant de plus la preuve de l'association entre stress et satisfaction au travail. Ces résultats suggèrent qu'une intervention dont l'objectif est de protéger les individus contre un potentiel risque de stress peut aussi s'avérer efficace pour lutter contre l'insatisfaction au travail, et inversement. L'intersection de ces deux indicateurs liés au travail appuie l'aspect multidimensionnel du bien-être psychosocial [10]. Ce constat nous a par ailleurs conduit à explorer un grand nombre de dimensions du travail dans notre modèle d'équation structurelle pour construire un indice de bien-être psychosocial. Après avoir mesuré l'exposition des travailleurs aux RPS et évaluer leurs facteurs de risques, nous avons exploré les relations entre les différents facteurs et un indicateur de performance des entreprises, l'absentéisme. Facteurs psychosociaux et absentéisme sur le lieu de travail sont intimement liés [70]. Dans les derniers articles de cette thèse, nous avons examiné les associations entre les facteurs psychosociaux que nous avons précédemment étudiés et l'absentéisme. Dans la première étude transversale, nous avons constaté qu'un bien-être psychosocial élevé était associé à un risque et à une durée d'absence pour maladie plus faibles. L'absentéisme pour cause de maladie est un indicateur majeur du bien être des travailleurs en entreprise [132]. Nos résultats corroborent l'idée selon laquelle l'amélioration de l'environnement psychosocial au travail contribue à réduire le risque d'absentéisme, peu

importe la durée de l'arrêt maladie [133]. Pour aller plus loin, dans la seconde étude placée cette fois-ci dans un cadre longitudinal, nous avons quantifié l'effet protecteur du bien-être psychosocial sur le risque d'absentéisme. La littérature a déjà pu mettre en évidence l'effet protecteur du bien-être psychosocial sur l'absentéisme [134]. L'emploi de modèles de survie et multi-états dans notre étude, classés parmi les modèles longitudinaux, renforce le niveau de preuve de ces résultats. Finalement, notre travail de thèse a permis d'accroître les connaissances autour des liens entre facteurs psychosociaux, RPS, qualité de vie et bien-être, et absentéisme (Figure 5).



**Figure 5.** Contribution de la thèse à l'approfondissement des connaissances sur les relations entre les facteurs psychosociaux, la qualité de vie au travail, le bien-être et l'absentéisme

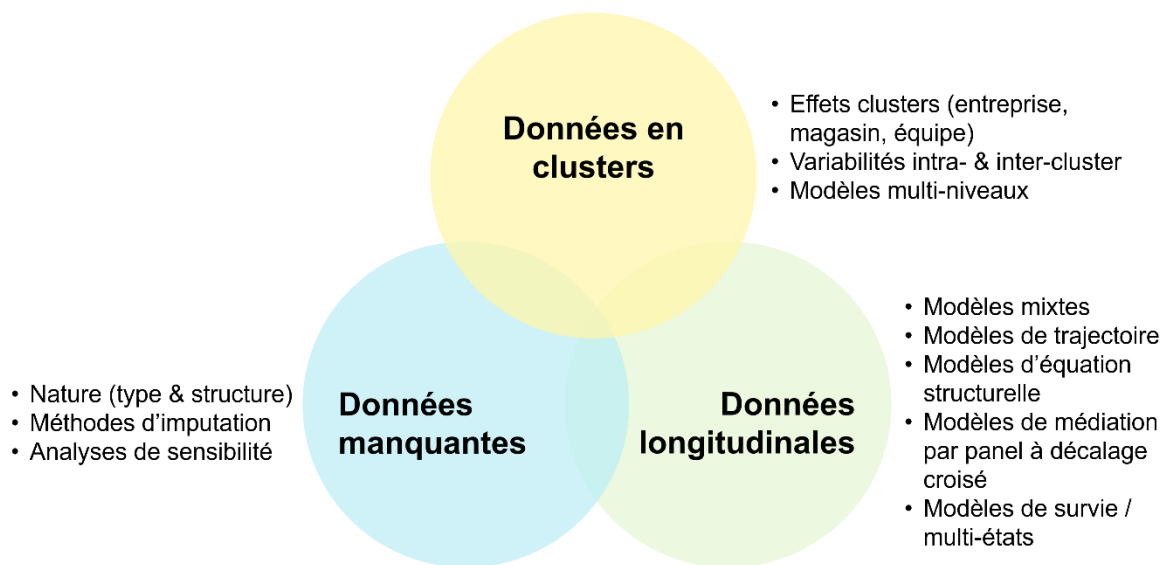
Lutter contre les RPS sur le lieu de travail permet non seulement d'améliorer le bien-être psychosocial et la santé des travailleurs, mais assure aussi le bien-être et la prospérité des entreprises [135]. Nous avons montré que cela pouvait se traduire, par exemple, par une réduction de l'absentéisme, favorisant la productivité et la performance globale des organisations [74]. Ce travail de thèse a démontré que l'analyse des données de santé au travail rendait possible l'étude des facteurs psychosociaux pour les travailleurs et les conséquences économiques néfastes pour les entreprises. En utilisant des modèles capables de mesurer l'exposition des travailleurs aux RPS, de prédire l'évolution de leur bien-être, en tenant compte de leurs caractéristiques et de celles des entreprises, le tout en analysant des données collectées

grâce à l'engagement volontaire des travailleurs et à leur capacité à évaluer leur propre état de bien-être, nous avons démontré que la médecine 4P était un outil efficace pour améliorer la santé des travailleurs et la performance des entreprises.

## **2. Les défis de l'analyse de données de santé au travail : enjeux et réponses méthodologiques**

Les bases de données du monde réel offre une source intarissable d'informations utile à la recherche en santé. Ces bases peuvent contenir une grande quantité de données collectées au sein d'échantillons représentatifs d'individus dans la population et permettent de suivre leur évolution au cours du temps, à moindre coût. Les données du monde réel sont une source inestimable pour les chercheurs, mais l'analyse de ces données comporte des défis méthodologiques qu'il convient de comprendre et de prendre en compte. La qualité des données, la gestion des possibles biais, et la nécessité d'appliquer des méthodes et modèles statistiques adéquats pour faire face aux nombreuses problématiques constituent des aspects cruciaux que les chercheurs se doivent de prendre en considération dans leurs analyses [136]. Dans le cadre de cette thèse, nous avons exploré les défis méthodologiques associés à l'analyse de données du monde réel, longitudinales, liées à la santé au travail. Notre attention s'est principalement portée sur trois défis majeurs : la gestion des données en cluster, la problématique des données manquantes et les spécificités des données longitudinales. S'il existe des situations dans lesquelles il est possible de répondre à ces enjeux indépendamment les uns des autres, une majeure partie des difficultés de l'analyse de données de santé au travail se situe à la frontière entre ces différents concepts. Notre travail a permis de mettre en évidence les diverses méthodes et solutions existantes dans la littérature scientifique qui permettent de répondre à ces enjeux, tant de manière isolée que simultanée (Figure 6).





**Figure 6.** Solutions et méthodes pour répondre aux enjeux méthodologiques de l'analyse de données de santé au travail

Pour aider les chercheurs à surmonter ces obstacles méthodologiques, nous avons élaboré une méthodologie basée sur des approches et des modèles statistiques éprouvés, tirés de la littérature spécialisée. Aborder les principales questions liées à l'analyse des données longitudinales peut considérablement améliorer la qualité des analyses. Traiter les données manquantes dans un jeu de données en les estimant ou en les imputant peut, par exemple, pallier au risque de biais et de perte de puissance statistique des analyses [137]. Plus largement, répondre aux défis de l'analyse de santé au travail peut contribuer à garantir que les résultats d'une étude soient généralisables [138]. Bien que la qualité des analyses soit intrinsèquement liée à la résolution des enjeux méthodologiques, cela peut parfois se révéler être une tâche complexe, voire dépendre en partie des décisions prises par les chercheurs. Par exemple, toujours dans le cas de données manquantes, nous avons vu qu'il est souvent difficile, voire impossible en pratique, de vérifier les hypothèses qui guident le choix d'un type de données manquantes par rapport à un autre [119]. Les chercheurs sont ainsi confrontés à un dilemme méthodologique délicat qui nécessite une approche réfléchie et des choix avisés pour atteindre des résultats de haute qualité. L'approche que nous avons exposée n'a pas pour ambition d'être une solution infaillible garantissant la fiabilité des résultats de n'importe quelle étude. Elle se veut plutôt être un cadre

général permettant aux chercheurs de détecter les éventuels biais dans leurs études, de les reconnaître et de travailler à les minimiser. Nous avons pris ce cadre comme base pour nos propres études tout au long de cette thèse, démontrant ainsi sa pertinence pour relever les défis méthodologiques qui nous ont été imposés. L'un de ces défis majeurs auquel nous avons été confrontés était lié à la nature de nos données. Les utilisateurs de la plateforme Wittyfit provenant de diverses entreprises, chacune avec une identité et des caractéristiques propres, nos données présentaient une structure hiérarchique. Cette structure nous a amené à considérer l'entreprise comme un cluster, et donc à imputer des observations manquantes entreprise par entreprise ou à utiliser des modèles mixtes en considérant un « effet entreprise » comme effet aléatoire. Une autre difficulté résidait dans la gestion des données manquantes, en particulier dans la définition de leur nature. Étant donné qu'il est impossible de suivre les utilisateurs au-delà des données recueillies dans la plateforme, nous ne pouvions pas déterminer avec certitude leur type, ce qui nous a obligé à les considérer comme étant MNAR. À l'inverse des études avec des données manquantes MCAR, il nous était impossible de lever pleinement et avec certitude certains biais comme le biais de sélection, d'information ou de confusion [139]. L'analyse de sensibilité devenait indispensable lorsque le taux de données manquantes était trop élevé afin d'évaluer la robustesse de nos résultats [140]. Comme de nombreuses bases de données du monde réel, la base Wittyfit est sujette à des problèmes de qualité des données, notamment en raison de données incomplètes et incohérentes [141]. Bien que Wittyfit puisse recueillir de manière répétée le ressenti de ses utilisateurs, elle ne peut garantir ni la régularité, ni la continuité de ces données. Dans nos études longitudinales, cette limitation s'est traduite par une attrition importante, nous obligeant à réduire la durée de la période d'étude et à opter pour des intervalles de temps de temps très larges, ce qui a restreint la généralisabilité de nos résultats. Malgré les nombreux enjeux et défis que nous avons dû surmonter au cours de nos études, nous avons démontré qu'en suivant rigoureusement la méthodologie que nous avons définie, il était possible d'obtenir des résultats pertinents et en accord avec la littérature. Ces

réalisations témoignent de l'efficacité de notre approche méthodologique et renforcent la crédibilité de notre travail de recherche, soulignant ainsi les avancées significatives que nous avons accomplies dans le domaine de la santé au travail.

# Conclusions et perspectives

## Conclusions et perspectives

La lutte contre les RPS en milieu professionnel est complexe car chaque entreprise et chaque travailleur présentent des spécificités qui demandent une évaluation plus personnalisée des risques. Cette thèse a exploré la possibilité d'utiliser l'analyse de données de santé au travail, en particulier des données issues de la plateforme Wittyfit, pour quantifier l'exposition des travailleurs aux RPS, identifier et évaluer les facteurs psychosociaux liés à ces risques, et mesurer l'impact de ces facteurs sur des indicateurs clés de performance des entreprises comme l'absentéisme. Nous avons évalué l'exposition des travailleurs aux RPS en étudiant des facteurs tels que le stress, l'insatisfaction au travail, le job strain, l'isostrain, ou encore le mal-être au travail. Parallèlement, nous avons examiné l'impact des caractéristiques sociodémographiques des individus et de leur environnement et conditions de travail sur ces risques. Nous avons enfin étudié et mis en évidence l'association entre les facteurs psychosociaux et l'absentéisme sur le lieu de travail, soulignant l'importance de garantir la santé des travailleurs pour les entreprises. Cette thèse a contribué à démontrer la pertinence de l'approche de la médecine 4P en milieu professionnel. Par l'analyse des données de santé au travail, nous avons démontré qu'il était de prédire l'exposition des travailleurs aux RPS et les risques pour leur santé. En identifiant les travailleurs à risque et en mesurant les conséquences sanitaires et économiques des facteurs psychosociaux, nous avons proposé des pistes d'amélioration pouvant être intégrées dans les stratégies de prévention contre les RPS. Notre travail a également mis en évidence les différences pouvant exister entre les entreprises, en particulier en ce qui concerne dans la mesure des RPS. Nous avons également observé des disparités entre les travailleurs, par exemple en comparant les situations de travail et les niveaux de satisfaction des managers et des employés. Ces constats soulignent la nécessité de mettre en œuvre des approches plus personnalisées pour améliorer la QVT sur le lieu de travail. Enfin, notre thèse a montré que la participation active des travailleurs à l'évaluation de leur bien-être psychosocial est un outil pertinent pour lutter contre les RPS.

D'un point de vue statistique et méthodologique, cette thèse a permis de faire l'état de l'art des enjeux de l'analyse de données de santé au travail et de présenter les modèles et méthodes pour y répondre. Elle a mis en évidence l'importance de traiter les données en cluster, de gérer les données manquantes dans les analyses et de tenir compte des spécificités des données longitudinales. Au-delà de la simple reconnaissance de ces enjeux, nous avons élaboré dans cette thèse une méthodologie robuste visant à surmonter ces obstacles et à élargir notre compréhension des défis liés à l'analyse des données du monde réel en santé au travail. Cette approche méthodologique offre un cadre solide pour relever les difficultés courantes et améliorer la qualité des futures analyses dans ce domaine en constante évolution. Appliquée à nos données, notre méthodologie a permis de fournir des résultats pertinents, robustes, alignés sur la littérature.

Pour conclure, cette thèse renforce la valeur de l'analyse des données de santé au travail, en mettant en évidence l'utilité des données du monde réel, en particulier celles provenant de la base Wittyfit, dans la lutte contre les RPS. Nous avons démontré qu'il était possible, à partir de ces données, d'étudier l'évolution de l'état de santé des travailleurs au fil des années ou d'établir des corrélations entre facteurs psychosociaux et répercussions économiques. Cependant, en raison la récente création de la base de données Wittyfit et des difficultés dans la collecte des données socio-économiques des entreprises qui utilisent la plateforme, nous n'avons pas pu examiner ces mêmes corrélations au cours du temps. Par exemple, il aurait été pertinent d'étudier comment une détérioration de la qualité de vie des travailleurs pouvait être prédictive d'une dégradation de leur état de santé ou d'un risque d'absentéisme. Pour des raisons similaires, nous n'avons pas été en mesure d'utiliser l'analyse de données de santé au travail sur de longues périodes d'étude pour anticiper les conséquences sanitaires et économiques, à court, moyen et long terme d'une intervention de prévention sur le lieu de travail. Par conséquent, à l'avenir, il serait judicieux de répliquer nos études sur les relations entre facteurs psychosociaux et absentéisme en étudiant comment l'évolution du bien-être psychosocial des

travailleurs influence le risque d'absence au cours du temps. Par leur caractère longitudinal, ces études pourraient en outre permettre d'appuyer ou de renforcer notre approche méthodologique. Ce travail de thèse a démontré la pertinence de l'utilisation des bases de données de santé au travail pour évaluer les RPS en milieu professionnel et contribuer à améliorer la qualité de vie des travailleurs. La digitalisation rapide et progressive de la médecine du travail offre des perspectives intéressantes, en particulier pour améliorer l'accessibilité et l'efficacité des soins, le suivi en temps réel de la santé des travailleurs, et la possibilité de réaliser des analyses de données plus poussées pour une meilleure prévention et gestion des RPS. Avec la révolution de la digitalisation et de l'accumulation continue des données de santé au travail, on peut envisager un futur où les travailleurs, les employeurs et les médecins du travail, dans un rapport gagnant-gagnant-gagnant, contribueraient tous ensemble à garantir la QVT pour tous.

# Références



## Références

1. International Health Conference. Constitution of the World Health Organization. 1946. *Bulletin of the World Health Organization*. 2002;80(12):983–4.
2. Larson JS. The World Health Organization's definition of health: social versus spiritual health. *Soc Indic Res*. 1996;38(2):181–92.
3. Eiroa-Orosa FJ. Understanding psychosocial wellbeing in the context of complex and multidimensional problems. *Int J Environ Res Public Health*. 2020;17(16):5937.
4. Guyatt GH, Feeny DH, Patrick DL. Measuring Health-Related Quality of Life. *Ann Intern Med*. 1993;118(8):622–9.
5. Parra-Giordano D, Quijada Sánchez D, Grau Mascayano P, Pinto-Galleguillos D. Quality of work life and work process of assistance nurses. *Int J Environ Res Public Health*. 2022;19(11):6415.
6. Brooks BA, Anderson MA. Defining quality of nursing work life. *Nurs Econ*. 2005;23(6):319–26, 279.
7. Khajehnasiri F, Foroushani AR, Kashani BF, Kassiri N. Evaluation of the quality of working life and its effective factors in employed nurses of Tehran University of Medical Sciences Hospitals. *J Educ Health Promot*. 2021;10:112.
8. Hackman JR, Oldham GR. The job diagnostic survey: an instrument for the diagnosis of jobs and the evaluation of job redesign projects. 1974.
9. Cooper CL, Mumford E. The quality of working life in Western and Eastern Europe. Greenwood Press; 1979.
10. Warr P, Cook J, Wall T. Scales for the measurement of some work attitudes and aspects of psychological well-being. *Journal of Occupational Psychology*. 1979;52(2):129–48.
11. Mirvis PH, Lawler III EE. Accounting for the quality of work life. *Journal of Organizational Behavior*. 1984;5(3):197–212.
12. Baba VV, Jamal M. Routinization of job context and job content as related to employees' quality of working life: A study of Canadian nurses. *Journal of Organizational Behavior*. 1991;12(5):379–86.
13. Nanjundeswaraswamy TS, Swamy DR. Quality of worklife of employees in private technical institutions. *International Journal for Quality Research*, ISSN 1800-6450. 2013;7:431–42.
14. Gollac M, Askenazy P, Baudelot C, Brochard P, Al E. Mesurer les facteurs psychosociaux de risque au travail pour les maîtriser. In 2011.
15. Ryff CD. Beyond Ponce de Leon and life satisfaction: new directions in quest of successful ageing. *International Journal of Behavioral Development*. 1989;12(1):35–55.

16. Kristensen TS, Hannerz H, Høgh A, Borg V. The Copenhagen Psychosocial Questionnaire-a tool for the assessment and improvement of the psychosocial work environment. *Scandinavian Journal of Work, Environment & Health*. 2005;31(6):438–49.
17. Cohen S, Kamarck T, Mermelstein R. Perceived stress scale. *Measuring stress: A guide for health and social scientists*. 1994;10(2):1–2.
18. Rick J, Briner RB. Psychosocial risk assessment: problems and prospects. *Occup Med (Lond)*. 2000;50(5):310–4.
19. Ilić IM, Arandjelović MŽ, Jovanović JM, Nešić MM. Relationships of work-related psychosocial risks, stress, individual factors and burnout – Questionnaire survey among emergency physicians and nurses. *Med Pr Work Health Saf*. 2017;68(2):167–78.
20. Faragher EB, Cass M, Cooper CL. The relationship between job satisfaction and health: a meta-analysis. *Occup Environ Med*. 2005;62(2):105–12.
21. Leka S, Griffiths A, Cox T, Organization WH. *Work organisation and stress: systematic problem approaches for employers, managers and trade union representatives*. World Health Organization; 2003.
22. Backé EM, Seidler A, Latza U, Rossnagel K, Schumann B. The role of psychosocial stress at work for the development of cardiovascular diseases: a systematic review. *Int Arch Occup Environ Health*. 2012;85(1):67–79.
23. de Wijn AN, van der Doef MP. Reducing psychosocial risk factors and improving employee well-being in emergency departments: a realist evaluation. *Front Psychol*. 2022;12:728390.
24. Steptoe A, Kivimäki M. Stress and cardiovascular disease. *Nat Rev Cardiol*. 2012;9(6):360–70.
25. Santosa A, Rosengren A, Ramasundarahettige C, Rangarajan S, Gulec S, Chifamba J, et al. Psychosocial risk factors and cardiovascular disease and death in a population-based cohort from 21 low-, middle-, and high-income countries. *JAMA Netw Open*. 2021;4(12):e2138920.
26. Letvak S, Buck R. Factors influencing work productivity and intent to stay in nursing. *Nurs Econ*. 2008;26(3):159–65.
27. Schmidt B, Schneider M, Seeger P, van Vianen A, Loerbroks A, Herr RM. A comparison of job stress models: associations with employee well-being, absenteeism, presenteeism, and resulting costs. *J Occup Environ Med*. 2019;61(7):535–44.
28. Nei D, Snyder LA, Litwiller BJ. Promoting retention of nurses: a meta-analytic examination of causes of nurse turnover. *Health Care Manage Rev*. 2015;40(3):237–53.
29. Wu Y, Zhou L, Zhang X, Yang X, Niedermann G, Xue J. Psychological distress and eustress in cancer and cancer treatment: Advances and perspectives. *Sci Adv*. 2022;8(47):eabq7982.
30. Kung CSJ, Chan CKY. Differential roles of positive and negative perfectionism in predicting occupational eustress and distress. *Personality and Individual Differences*. 2014;58:76–81.

31. Wood S, Ghezzi V, Barbaranelli C, Di Tecco C, Fida R, Farnese ML, et al. Assessing the risk of stress in organizations: getting the measure of organizational-level stressors. *Front Psychol* [Internet]. 2019 [cited 2024 Apr 8];10. Available from: <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2019.02776/full>
32. Lesage FX, Berjot S, Deschamps F. Clinical stress assessment using a visual analogue scale. *Occup Med (Lond)*. 2012;62(8):600–5.
33. Giménez-Espert M del C, Prado-Gascó V, Soto-Rubio A. Psychosocial risks, work engagement, and job satisfaction of nurses during COVID-19 pandemic. *Frontiers in Public Health*. 2020;8.
34. Locke EA. The nature and causes of job satisfaction. *Handbook of Industrial and Organizational Psychology*. 1976;1:1297–343.
35. Satuf C, Monteiro S, Pereira H, Esgalhado G, Marina Afonso R, Loureiro M. The protective effect of job satisfaction in health, happiness, well-being and self-esteem. *Int J Occup Saf Ergon*. 2018;24(2):181–9.
36. Hou J, He Y, Zhao X, Thai J, Fan M, Feng Y, et al. The effects of job satisfaction and psychological resilience on job performance among residents of the standardized residency training: a nationwide study in China. *Psychol Health Med*. 2020;25(9):1106–18.
37. Bagheri Hosseinabadi M, Etemadinezhad S, Khanjani N, Ahmadi O, Gholinia H, Galeshi M, et al. Evaluating the relationship between job stress and job satisfaction among female hospital nurses in Babol: An application of structural equation modeling. *Health Promot Perspect*. 2018;8(2):102–8.
38. Aaviksoo E, Baburin A, Kiiwet RA. Risk factors for sickness absence among Estonian employees. *Occup Med (Lond)*. 2013;63(2):156–9.
39. Kresal F, Suklan J, Roblek V, Jerman A, Meško M. Psychosocial Risk Factors for Low Back Pain and Absenteeism among Slovenian Professional Drivers. *Cent Eur J Public Health*. 2017;25(2):135–40.
40. Han GH, Jekel M. The mediating role of job satisfaction between leader-member exchange and turnover intentions. *J Nurs Manag*. 2011;19(1):41–9.
41. Weiss DJ, Dawis RV, England GW. Manual for the Minnesota Satisfaction Questionnaire. *Minnesota Studies in Vocational Rehabilitation*. 1967;22:120–120.
42. Martin A, Karanika-Murray M, Biron C, Sanderson K. The psychosocial work environment, employee mental health and organizational interventions: improving research and practice by taking a multilevel approach. *Stress and Health*. 2016;32(3):201–15.
43. Karasek RA. Job Demands, Job Decision Latitude, and Mental Strain: Implications for Job Redesign. *Administrative Science Quarterly*. 1979;24(2):285–308.
44. Karasek R, Baker D, Marxer F, Ahlbom A, Theorell T. Job decision latitude, job demands, and cardiovascular disease: a prospective study of Swedish men. *Am J Public Health*. 1981;71(7):694–705.

45. Theorell T, Karasek RA, Eneroth P. Job strain variations in relation to plasma testosterone fluctuations in working men - a longitudinal study. *Journal of Internal Medicine*. 1990;227(1):31–6.
46. Siegrist J. Adverse health effects of high-effort/low-reward conditions. *J Occup Health Psychol*. 1996;1(1):27–41.
47. de Jonge J, Bosma H, Peter R, Siegrist J. Job strain, effort-reward imbalance and employee well-being: a large-scale cross-sectional study. *Social Science & Medicine*. 2000;50(9):1317–27.
48. Siegrist J, Starke D, Chandola T, Godin I, Marmot M, Niedhammer I, et al. The measurement of effort–reward imbalance at work: European comparisons. *Social Science & Medicine*. 2004;58(8):1483–99.
49. Niedhammer I, Chastang JF, Gendrey L, David S, Degioanni S. [Psychometric properties of the French version of Karasek’s “Job Content Questionnaire” and its scales measuring psychological pressures, decisional latitude and social support: the results of the SUMER]. *Sante Publique*. 2006;18(3):413–27.
50. Leka S, Cox T, Zwetsloot G. The European framework for psychosocial risk management (PRIMA-EF). *The European Framework for Psychosocial Risk Management: PRIMA-EF*, Leka S and Cox T (Eds). 2008;1–16.
51. Leka S, Jain A, Organization WH. Health impact of psychosocial hazards at work: an overview. 2010;
52. Fernandes C, Pereira A. Exposure to psychosocial risk factors in the context of work: a systematic review. *Rev Saude Publica*. 2016;50:24.
53. Roussos PL. The psychosocial risks and impacts in the workplace assessment tool: construction and psychometric evaluation. *Behav Sci (Basel)*. 2023;13(2):104.
54. Milczarek M, Brun E, Houtman I, Goudswaard A, Evers M, Bovenkamp M, et al. Expert forecast on emerging psychosocial risks related to occupational safety and health. 2007;
55. Schaufeli WB, Greenglass ER. Introduction to special issue on burnout and health. *Psychology & health*. 2001;16(5):501–10.
56. Niedhammer I, Bèque M, Chastang JF, Bertrais S. Psychosocial work exposures and suicide ideation: a study of multiple exposures using the French national working conditions survey. *BMC Public Health*. 2020;20(1):895.
57. van der Molen HF, Nieuwenhuijsen K, Frings-Dresen MHW, de Groene G. Work-related psychosocial risk factors for stress-related mental disorders: an updated systematic review and meta-analysis. *BMJ Open*. 2020;10(7):e034849.
58. Shah SMA, Mohammad D, Qureshi MFH, Abbas MZ, Aleem S. Prevalence, psychological responses and associated correlates of depression, anxiety and stress in a global population, during the coronavirus disease (covid-19) pandemic. *Community Ment Health J*. 2021;57(1):101–10.

59. Davis MT, Holmes SE, Pietrzak RH, Esterlis I. Neurobiology of Chronic Stress-Related Psychiatric Disorders: Evidence from Molecular Imaging Studies. *Chronic Stress (Thousand Oaks)*. 2017;1:2470547017710916.
60. Seidler A, Thinschmidt M, Deckert S, Then F, Hegewald J, Nieuwenhuijsen K, et al. The role of psychosocial working conditions on burnout and its core component emotional exhaustion – a systematic review. *J Occup Med Toxicol*. 2014;9:10.
61. Mehta RK, Parijat P. Associations between psychosocial risk factors and musculoskeletal disorders: application to the IT profession in India. *Work*. 2012;41 Suppl 1:2438–44.
62. van der Molen HF, Foresti C, Daams JG, Frings-Dresen MHW, Kuijer PPFM. Work-related risk factors for specific shoulder disorders: a systematic review and meta-analysis. *Occup Environ Med*. 2017;74(10):745–55.
63. van den Berge M, van der Beek AJ, Türkeli R, van Kalken M, Hulsege G. Work-related physical and psychosocial risk factors cluster with obesity, smoking and physical inactivity. *Int Arch Occup Environ Health*. 2021;94(4):741–50.
64. Kivimäki M, Nyberg ST, Batty GD, Fransson EI, Heikkilä K, Alfredsson L, et al. Job strain as a risk factor for coronary heart disease: a collaborative meta-analysis of individual participant data. *Lancet*. 2012;380(9852):1491–7.
65. Schneiderman N, McIntosh RC, Antoni MH. Psychosocial risk and management of physical diseases. *J Behav Med*. 2019;42(1):16–33.
66. Cuevas AG, Chen R, Thurber KA, Slopen N, Williams DR. Psychosocial stress and overweight and obesity: findings from the Chicago community adult health study. *Ann Behav Med*. 2019;53(11):NP.
67. Amiri S, Behnezhad S. Job strain and mortality ratio: a systematic review and meta-analysis of cohort studies. *Public Health*. 2020;181:24–33.
68. Liao W, Chen Y, Gao Q, Gan R, Li M, Liu Z, et al. Impact of stress hyperglycemia ratio, derived from glycated albumin or hemoglobin A1c, on mortality among ST-segment elevation myocardial infarction patients. *Cardiovasc Diabetol*. 2023;22(1):334.
69. Patton E. The devil is in the details: Judgments of responsibility and absenteeism from work. *Journal of Occupational and Organizational Psychology*. 2011;84(4):759–79.
70. Niedhammer I, Chastang JF, Sultan-Taïeb H, Vermeylen G, Parent-Thirion A. Psychosocial work factors and sickness absence in 31 countries in Europe. *Eur J Public Health*. 2013;23(4):622–9.
71. Halbesleben JRB, Whitman MV, Crawford WS. A dialectical theory of the decision to go to work: bringing together absenteeism and presenteeism. *Human Resource Management Review*. 2014;24(2):177–92.
72. Martocchio JJ, Harrison DA. To be there or not to be there?: Questions, theories, and methods in absenteeism research. *Research in personnel and human resources management : a research annual*. 1993;11.

73. Asay GRB, Roy K, Lang JE, Payne RL, Howard DH. Absenteeism and Employer Costs Associated With Chronic Diseases and Health Risk Factors in the US Workforce. *Prev Chronic Dis.* 2016;13:E141.
74. Nagata T, Mori K, Ohtani M, Nagata M, Kajiki S, Fujino Y, et al. Total Health-Related Costs Due to Absenteeism, Presenteeism, and Medical and Pharmaceutical Expenses in Japanese Employers. *J Occup Environ Med.* 2018;60(5):e273–80.
75. Ivanova JI, Birnbaum HG, Samuels S, Davis M, Phillips AL, Meletiche D. The cost of disability and medically related absenteeism among employees with multiple sclerosis in the US. *Pharmacoeconomics.* 2009;27(8):681–91.
76. Chirico F, Heponiemi T, Pavlova M, Zaffina S, Magnavita N. Psychosocial risk prevention in a global occupational health perspective. A descriptive analysis. *Int J Environ Res Public Health.* 2019;16(14):2470.
77. Houtman I, van Zwieten M, Leka S, Jain A, de Vroome E. Social dialogue and psychosocial risk management: added value of manager and employee representative agreement in risk perception and awareness. *Int J Environ Res Public Health.* 2020;17(10):3672.
78. Hood L, Heath JR, Phelps ME, Lin B. Systems biology and new technologies enable predictive and preventative medicine. *Science.* 2004;306(5696):640–3.
79. Hood L, Balling R, Auffray C. Revolutionizing medicine in the 21st century through systems approaches. *Biotechnol J.* 2012;7(8):992–1001.
80. Hood L, Flores M. A personal view on systems medicine and the emergence of proactive P4 medicine: predictive, preventive, personalized and participatory. *New Biotechnology.* 2012;29(6):613–24.
81. Flores M, Glusman G, Brogaard K, Price ND, Hood L. P4 medicine: how systems medicine will transform the healthcare sector and society. *Personalized Medicine.* 2013;10(6):565–76.
82. Alonso SG, de la Torre Díez I, Zapiraín BG. Predictive, personalized, preventive and participatory (4P) medicine applied to telemedicine and eHealth in the literature. *J Med Syst.* 2019;43(5):140.
83. Boffetta P, Collatuzzo G. Application of P4 (predictive, preventive, personalized, participatory) approach to occupational medicine. *Med Lav.* 2022;113(1):e2022009.
84. Dutheil F, Duclos M, Naughton G, Dewavrin S, Cornet T, Huguet P, et al. Wittyfit-live your work differently: study protocol for a workplace-delivered health promotion. *JMIR Research Protocols.* 2017;6(4):e6267.
85. Delgado DA, Lambert BS, Boutris N, McCulloch PC, Robbins AB, Moreno MR, et al. Validation of digital visual analog scale pain scoring with a traditional paper-based visual analog scale in adults. *J Am Acad Orthop Surg Glob Res Rev.* 2018;2(3):e088.
86. de Korte EM, Wiezer N, Janssen JH, Vink P, Kraaij W. Evaluating an mHealth App for Health and Well-Being at Work: Mixed-Method Qualitative Study. *JMIR Mhealth Uhealth.* 2018;6(3):e72.

87. Paganin G, Simbula S. New technologies in the workplace: can personal and organizational variables affect the employees' intention to use a work-stress management app? *Int J Environ Res Public Health*. 2021;18(17):9366.
88. Ryan C, Bergin M, Chalder T, Wells JS. Web-based interventions for the management of stress in the workplace: focus, form, and efficacy. *J Occup Health*. 2017;59(3):215–36.
89. Makady A, de Boer A, Hillege H, Klungel O, Goettsch W. What is real-world data? A review of definitions based on literature and stakeholder interviews. *Value in Health*. 2017;20(7):858–65.
90. Baro E, Degoul S, Beuscart R, Chazard E. Toward a literature-driven definition of Big Data in healthcare. *Biomed Res Int*. 2015;2015:639021.
91. Deering S, Pratap A, Suver C, Borelli AJ, Amdur A, Headapohl W, et al. Real-world longitudinal data collected from the SleepHealth mobile app study. *Sci Data*. 2020;7(1):418.
92. Maissenhaelter BE, Woolmore AL, Schlag PM. Real-world evidence research based on big data. *Onkologe (Berl)*. 2018;24(Suppl 2):91–8.
93. Corrigan-Curay J, Sacks L, Woodcock J. Real-world evidence and real-world data for evaluating drug safety and effectiveness. *JAMA*. 2018;320(9):867–8.
94. Clinchamps M, Auclair C, Prunet D, Pfabigan D, Lesage FX, Baker JS, et al. Burnout among hospital non-healthcare staff: influence of job demand-control-support, and effort-reward imbalance. *J Occup Environ Med*. 2021;63(1):e13–20.
95. Dogbla L, Gouvenelle C, Thorin F, Lesage FX, Zak M, Ugbolue UC, et al. Occupational risk factors by sectors: an observational study of 20,000 workers. *Int J Environ Res Public Health*. 2023;20(4):3632.
96. Costa da Silva R de C, Pires de Novais MA, Zucchi P. Notifications of work-related injuries and diseases: an observational study on a mining disaster. *BMC Health Serv Res*. 2023;23(1):936.
97. Becker A, Angerer P, Weber J, Müller A. The prevention of musculoskeletal complaints: long-term effect of a work-related psychosocial coaching intervention compared to physiotherapy alone—a randomized controlled trial. *Int Arch Occup Environ Health*. 2020;93(7):877–89.
98. Uchiyama Y, Sasanuma N, Nanto T, Fujita K, Takahashi M, Iwasa S, et al. COVID-19 Patient Returned to Work after Long Hospitalization and Follow-up: A Case Report. *Prog Rehabil Med*. 2021;6:20210025.
99. Laird NM. Missing data in longitudinal studies. *Statistics in Medicine*. 1988;7(1–2):305–15.
100. Graham JW. Missing Data Analysis: Making It Work in the Real World. *Annu Rev Psychol*. 2008;60(1):549–76.
101. Garcia TP, Marder K. Statistical approaches to longitudinal data analysis in neurodegenerative diseases: huntington's disease as a model. *Curr Neurol Neurosci Rep*. 2017;17(2):14.

102. Holditch-Davis D, Levy J. Potential pitfalls in collecting and analyzing longitudinal data from chronically ill populations. *Newborn Infant Nurs Rev.* 2010;10(1):10–8.
103. Diggle P, Heagerty P, Liang KY, Zeger S. *Analysis of longitudinal data. SECOND EDITION.* OUP Oxford; 2002. 396 p.
104. Rotnitzky A, Jewell NP. Hypothesis testing of regression parameters in semiparametric generalized linear models for cluster correlated data. *Biometrika.* 1990;77(3):485–97.
105. Holodinsky JK, Austin PC, Williamson TS. An introduction to clustered data and multilevel analyses. *Family Practice.* 2020;37(5):719–22.
106. Gibbons RD, Hedeker D, DuToit S. Advances in analysis of longitudinal data. *Annu Rev Clin Psychol.* 2010;6:79–107.
107. Begg MD, Parides MK. Separation of individual-level and cluster-level covariate effects in regression analysis of correlated data. *Statistics in Medicine.* 2003;22(16):2591–602.
108. Little TD, Lang KM, Wu W, Rhemtulla M. Missing Data. In: *Developmental Psychopathology.* John Wiley & Sons, Ltd; 2016. p. 1–37.
109. Little RJA. Modeling the Drop-Out Mechanism in Repeated-Measures Studies. *Journal of the American Statistical Association.* 1995;90(431):1112–21.
110. Kang H. The prevention and handling of the missing data. *Korean J Anesthesiol.* 2013;64(5):402–6.
111. Desai M, Esserman DA, Gammon MD, Terry MB. The use of complete-case and multiple imputation-based analyses in molecular epidemiology studies that assess interaction effects. *Epidemiol Perspect Innov.* 2011;8:5.
112. He Y. Missing data analysis using multiple imputation: getting to the heart of the matter. *Circ Cardiovasc Qual Outcomes.* 2010;3(1):98.
113. Weed L, Lok R, Chawra D, Zeitzer J. The impact of missing data and imputation methods on the analysis of 24-hour activity patterns. *Clocks Sleep.* 2022;4(4):497–507.
114. Schafer JL, Olsen MK. Multiple imputation for multivariate missing-data problems: a data analyst's perspective. *Multivariate Behavioral Research.* 1998;33(4):545–71.
115. Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res.* 1999;8(1):3–15.
116. Allison PD. *Missing Data.* SAGE Publications; 2001. 100 p. (Quantitative Applications in the Social Sciences; vol. 136).
117. Little RJA, Rubin DB. *Statistical Analysis with Missing Data. Second Edition.* New York: John Wiley & Sons; 2002.
118. Gelman A, Hill J. *Data analysis using regression and multilevel/hierarchical models.* Cambridge University Press; 2006. 651 p.
119. Fitzmaurice G. Missing data: implications for analysis. *Nutrition.* 2008;24(2):200–2.



120. Thabane L, Mbuagbaw L, Zhang S, Samaan Z, Marcucci M, Ye C, et al. A tutorial on sensitivity analyses in clinical trials: the what, why, when and how. *BMC Med Res Methodol.* 2013;13:92.
121. de Souza RJ, Eisen RB, Perera S, Bantoto B, Bawor M, Dennis BB, et al. Best (but oft-forgotten) practices: sensitivity analyses in randomized controlled trials. *Am J Clin Nutr.* 2016;103(1):5–17.
122. Mowbray FI, Manlongat D, Shukla M. Sensitivity analysis: a method to promote certainty and transparency in nursing and health research. *Can J Nurs Res.* 2022;54(4):371–6.
123. Di Tecco C, Persechino B, Iavicoli S. Psychosocial risks in the changing world of work: moving from the risk assessment culture to the management of opportunities. *Med Lav.* 2023;114(2):e2023013.
124. Rugulies R, Krause N. Job strain, iso-strain, and the incidence of low back and neck injuries. A 7.5-year prospective study of San Francisco transit operators. *Social Science & Medicine.* 2005;61(1):27–39.
125. Madsen IEH, Nyberg ST, Magnusson Hanson LL, Ferrie JE, Ahola K, Alfredsson L, et al. Job strain as a risk factor for clinical depression: systematic review and meta-analysis with additional individual participant data. *Psychol Med.* 2017;47(8):1342–56.
126. Mutambudzi M, Javed Z. Job strain as a risk factor for incident diabetes mellitus in middle and older age U.S. workers. *The Journals of Gerontology: Series B.* 2016;71(6):1089–96.
127. Ferrario MM, Veronesi G, Bertù L, Grassi G, Cesana G. Job strain and the incidence of coronary heart diseases: does the association differ among occupational classes? A contribution from a pooled analysis of Northern Italian cohorts. *BMJ Open.* 2017;7(1):e014119.
128. Rahman M, Sen AK. Effect of job satisfaction on stress, performance and health in self-paced repetitive work. *Int Arch Occup Environ Health.* 1987;59(2):115–21.
129. Bae SH. The association between health status and job satisfaction among female workers: A nationwide cross-sectional study. *Nurs Health Sci.* 2021;23(4):908–15.
130. Vanroelen C, Levecque K, Louckx F. Differential exposure and differential vulnerability as counteracting forces linking the psychosocial work environment to socioeconomic health differences. *J Epidemiol Community Health.* 2010;64(10):866–73.
131. Penconek T, Tate K, Bernardes A, Lee S, Micaroni SPM, Balsanelli AP, et al. Determinants of nurse manager job satisfaction: A systematic review. *Int J Nurs Stud.* 2021;118:103906.
132. Gimeno D, Benavides F, Amick B, Benach J, Martinez JM. Psychosocial factors and work related sickness absence among permanent and non-permanent employees. *J Epidemiol Community Health.* 2004;58(10):870–6.
133. Mathisen J, Nguyen TL, Jensen JH, Mehta AJ, Rugulies R, Rod NH. Impact of hypothetical improvements in the psychosocial work environment on sickness absence rates: a simulation study. *Eur J Public Health.* 2022;32(5):716–22.

134. Jeong W, Kim YK, Oh SS, Yoon JH, Park EC. Association between presenteeism/absenteeism and well-being among Korean workers. *J Occup Environ Med.* 2020;62(8):574–80.
135. Kim CY. Psychological well-being, knowledge management behavior and performance: the moderating role of leader-member exchange. *Front Psychol.* 2021;12:566516.
136. Rogers JR, Lee J, Zhou Z, Cheung YK, Hripcsak G, Weng C. Contemporary use of real-world data for clinical trial conduct in the United States: a scoping review. *J Am Med Inform Assoc.* 2021;28(1):144–54.
137. Palmer RF, Royall DR. Missing Data? Plan on It! *Journal of the American Geriatrics Society.* 2010;58(s2):S343–8.
138. Malik HB, Norman JB. Best practices and methodological strategies for addressing generalizability in neuropsychological assessment. *J Pediatr Neuropsychol.* 2023;9(2):47–63.
139. Perkins NJ, Cole SR, Harel O, Tchetgen Tchetgen EJ, Sun B, Mitchell EM, et al. Principled approaches to missing data in epidemiologic studies. *Am J Epidemiol.* 2018;187(3):568–75.
140. Mavridis D, White IR. Dealing with missing outcome data in meta-analysis. *Res Synth Methods.* 2020;11(1):2–13.
141. Grimberg F, Asprion PM, Schneider B, Miho E, Babrak L, Habbabeh A. The real-world data challenges radar: a review on the challenges and risks regarding the use of real-world data. *Digit Biomark.* 2021;5(2):148–57.

# Annexes

# Annexes

## Annexe A : Captures d'écran de la plateforme Wittyfit

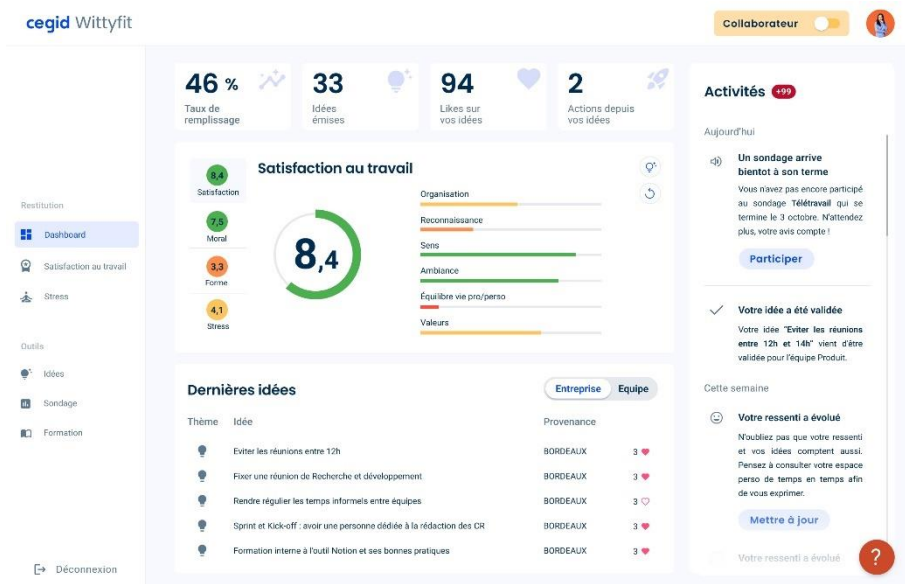
Les captures d'écran ci-dessous sont issues de la plateforme Wittyfit et présentent les deux parties de la plateforme : la partie dédiée aux collaborateurs (**Figures A1, A2 et A3**), et la partie dédiée aux managers (**Figures A4 et A5**). Les utilisateurs disposant d'un statut de managers peuvent basculer d'une partie à l'autre de la plateforme en utilisant le bouton prévu à cet effet, leur permettant d'agir à la fois comme un collaborateur de l'entreprise pouvant évaluer son bien-être au travail ou comme un manager souhaitant consulter l'état de bien-être de son équipe ou de son entreprise.



The screenshot shows a questionnaire interface with a progress bar at the top: Général (checked), Questionnaire 2/8, Vos idées, and Validation. The main heading is "Indiquez votre ressenti" with a placeholder text below it. There are three sliders: "Satisfaction au travail", "Forme", and "Stress". The "Stress" slider has a warning icon and a note: "Cet indicateur peut vous paraître inversé par rapport aux autres". A "Button" is located at the bottom right.

**Figure A1.** Capture d'écran de la plateforme Wittyfit : questionnaires

Légende : Quand ils se connectent à la plateforme, les utilisateurs sont invités à exprimer/mettre à jour leur ressenti dans les différentes thématiques de la plateforme. Le remplissage des questionnaires s'effectue en cascade et permet d'évaluer, dans l'ordre, le ressenti des collaborateurs dans (1) les indicateurs principaux de la plateforme (par exemple dans cette illustration, la satisfaction au travail, la forme et le stress), (2) les sous-indicateurs associés aux indicateurs principaux (par exemple, la charge de travail et l'autonomie sont définies comme des sous-indicateurs du stress), et (3) les thématiques associées aux sous-indicateurs (par exemple, les questions relatives à la quantité de travail ou à l'interruption de tâches sont définies comme des thématiques liées à la charge de travail). Les utilisateurs peuvent, quand ils le souhaitent, mettre à jour leurs réponses aux questionnaires.



**Figure A2.** Capture d'écran de la plateforme Wittyfit : dashboard collaborateur

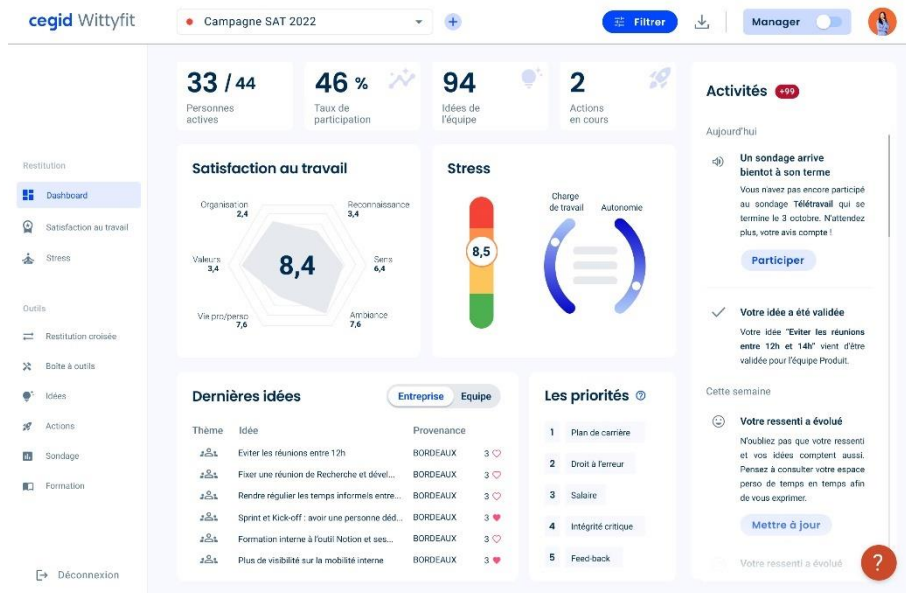
Légende : Le dashboard collaborateur est l'interface centrale depuis laquelle un utilisateur peut consulter son état de bien-être dans les différents indicateurs principaux, consulter les idées émises par les utilisateurs de la plateforme de son entreprise, son équipe, ou lui-même et les mettre en avant, ou participer à des enquêtes et sondages.



**Figure A3.** Capture d'écran de la plateforme Wittyfit : suggestions d'idées

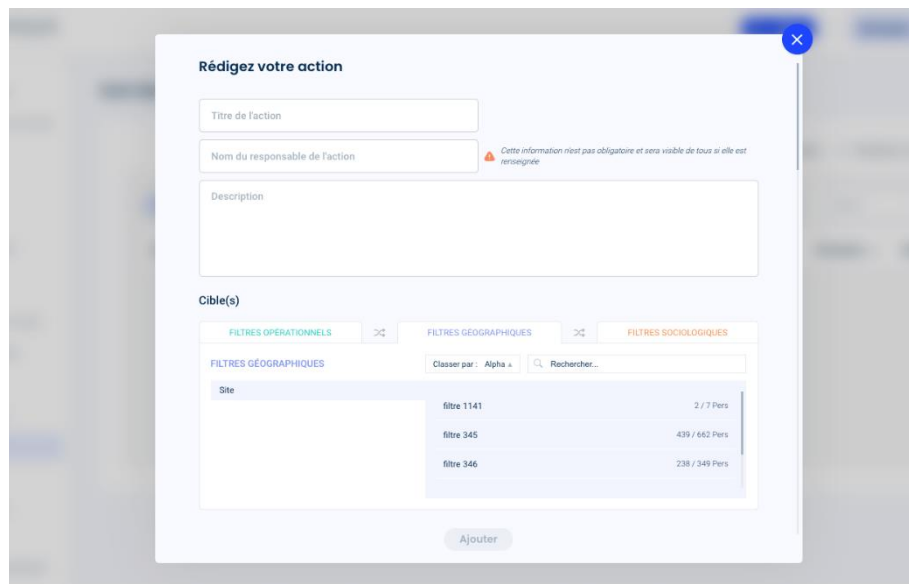
Légende : Après avoir complété les différents questionnaires ou depuis son dashboard, un utilisateur peut émettre des idées visant à améliorer le bien-être global sur le lieu de travail. Après avoir sélectionné une thématique spécifique, l'utilisateur peut développer librement son idée. Une fois validée, l'idée est transmise à son entreprise,

tant aux managers qui pourront s'en saisir pour mener des actions qui vont dans le sens de l'idée de l'émetteur qu'aux autres collaborateurs qui pourront la pousser.



**Figure A4.** Capture d'écran de la plateforme Wittyfit : dashboard manager

Légende : Le dashboard manager est l'interface centrale depuis laquelle un utilisateur disposant du statut de manager peut consulter l'état de bien-être de son équipe, consulter les idées émises par les utilisateurs de la plateforme de son entreprise, ses équipes, ou lui-même et les mettre en avant, consulter les thématiques prioritaires de ses équipes, mettre en place des plans d'action pour améliorer la QVT, émettre des sondages, ou suivre des formations. Le menu déroulant en haut de son dashboard lui permet de naviguer entre le mode live qui se base sur les dernières données à jour des utilisateurs, et les campagnes menées dans son entreprise. Une campagne correspond à une période durant laquelle les utilisateurs ont été invités par leur entreprise à mettre à jour leur ressenti, permettant à l'entreprise d'établir un état des lieux à un temps précis.



**Figure A5.** Capture d'écran de la plateforme Wittyfit : mise en place d'une action

Légende : Dans une optique d'amélioration de la QVT, un manager peut mettre en place une action spécifique à une thématique cible, visant une partie ou l'intégralité des équipes sous sa responsabilité. Après avoir nommé et décrit l'action qu'il souhaite mener, le manager peut sélectionner sa population cible à partir des différents filtres opérationnels, géographiques et sociologiques à sa disposition. Il doit enfin planifier son action en lui donnant un début et en fin. Les travailleurs ciblés par l'action seront invités à évaluer la thématique ciblée avant et après la mise en place de l'action, permettant au manager de mesurer l'efficacité de son action.

## **Annexe B : Présentation orale au Small Meeting d’hiver du LAPSCO (2021)**

### **Facteurs psychosociaux et économiques en milieu professionnel à l’heure de la médecine 4P : évolution longitudinale et prédictive du comportement des travailleurs**

**Rémi COLIN**

*[encadrement : Frédéric Dutheil, Bruno Pereira, Samuel Dewavrin]*

Les facteurs psychosociaux renvoient à des aspects du travail susceptibles d’agir sur la santé physique et mentale des travailleurs.<sup>1</sup> Ces facteurs comprennent, entre autres, les exigences de l’emploi, la satisfaction professionnelle ou encore les relations sociales. Les risques psychosociaux (RPS) engendrés, à savoir le stress et les violences internes et externes, représentent un enjeu majeur en milieu professionnel tant leurs effets sur la santé des travailleurs et sur la performance des entreprises peuvent être néfastes.

La médecine 4P (prédictive, préventive, prédictive et personnalisée) repose sur une approche systémique de la maladie, sur les technologies émergentes et sur les outils d’analyse pour promouvoir le bien-être des individus.<sup>2</sup> À l’instar de la médecine curative conventionnelle, elle a pour objectifs l’anticipation des symptômes et des maladies, leur prévention et leur traitement de manière spécifique.<sup>3</sup> Une approche possible contre les RPS en milieu professionnel et fidèle au principe de la médecine 4P consiste alors en l’identification des situations à risques, la mise en débat entre tous les salariés d’une évolution dans le fonctionnement organisationnel de l’entreprise et l’accompagnement de cette démarche.

Wittyfit est une plateforme visant à promouvoir et améliorer le bien-être des travailleurs en entreprise. Au travers de sondages et de questionnaires anonymes, les volontaires peuvent faire part de leur ressenti sur différentes thématiques liées à la santé. En plus de mesurer l’évolution de l’expérience collaborateur, les résultats permettent d’évaluer les RPS et la qualité de vie au travail des salariés.<sup>4</sup> En retour, les travailleurs reçoivent un feedback personnel et individualisé de leur état de santé. De même, les managers peuvent consulter l’état de santé global de leur équipe et ainsi mener des actions spécifiques de promotion du bien-être au travail puis les évaluer. Cette démarche s’inscrit parfaitement dans le cadre de la médecine 4P.

Notre projet consiste à étudier les données de la base Wittyfit. Les enjeux de santé publique et de médico-économie sont d’une part l’identification des facteurs psychosociaux et des facteurs économiques responsables de la baisse de performance des entreprises et d’autre part l’estimation des coûts engendrés par les RPS pour les travailleurs et pour les entreprises. Les enjeux statistiques sont quant à eux le développement de modèles spécifiques à l’analyse de données longitudinales de santé au travail et à la prise en compte des problématiques qui leurs sont liées, notamment les concepts de données en clusters, manquantes et longitudinales en elles-mêmes.<sup>5-8</sup>

Une première étude longitudinale s’intéressait à l’évolution sur une année de la satisfaction des travailleurs et au risque d’absence pour maladie en milieu professionnel. Cette



étude a permis de déceler quatre trajectoires spécifiques d'évolution et de mesurer le risque d'absentéisme associé à chaque trajectoire. Il a ainsi été constaté que le risque était plus élevé pour les individus dont la satisfaction était restée faible ou avait diminué au cours du temps, suggérant que la satisfaction au travail joue un rôle protecteur sur le risque d'absence pour maladie. Enfin, elle a permis de mettre en évidence l'influence de la position du travailleur sur l'appartenance à une trajectoire.

## Références

1. Thomas, K., Nilsson, E., Festin, K., Henriksson, P., Lowén, M., Löf, M., & Kristenson, M. (2020). Associations of Psychosocial Factors with Multiple Health Behaviors : A Population-Based Study of Middle-Aged Men and Women. *International Journal of Environmental Research and Public Health*, 17(4), 1239.
2. Hood, L., & Friend, S. H. (2011). Predictive, personalized, preventive, participatory (P4) cancer medicine. *Nature Reviews. Clinical Oncology*, 8(3), 184-187.
3. Slim, K., Selvy, M., & Veziat, J. (2021). Conceptual innovation : 4P Medicine and 4P surgery. *Journal of Visceral Surgery*, 158(3, Supplement), S12-S17.
4. Dutheil F, Duclos M, Naughton G, Dewavrin S, Cornet T, Huguet P, et al. WittyFit-Live Your Work Differently: Study Protocol for a Workplace-Delivered Health Promotion. *JMIR Res Protoc*. 2017;6(4).
5. Little, R. J. A., & Rubin, D. B. (2002). *Statistical Analysis with Missing Data* (Second Edition). John Wiley & Sons.
6. Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, 38(4), 963-974.
7. Verbeke, G. (1997). Linear Mixed Models for Longitudinal Data. In G. Verbeke & G. Molenberghs (Éds.), *Linear Mixed Models in Practice : A SAS-Oriented Approach* (p. 63-153). Springer.
8. Edwards, L. J. (2000). Modern statistical techniques for the analysis of longitudinal data in biomedical research. *Pediatric Pulmonology*, 30(4), 330-344.

## Annexe C : Présentation orale au Small Meeting d'hiver du LAPSCO (2023)

### NOM Prénom

COLIN Rémi

### Titre de la présentation

Stress et satisfaction au travail au cours du temps, l'influence du statut managérial : une modélisation bivariée et longitudinale des données de Wittyfit

### Durée de présentation estimée (15 ou 30 minutes avec questions)

20 minutes

### Résumé de la présentation (500 mots maxi)

Le stress désigne l'incapacité d'un organisme à répondre à une menace, qu'elle soit physique ou émotionnelle.<sup>1</sup> La satisfaction au travail, elle, désigne un état d'équilibre entre les attentes d'un travailleur vis-à-vis de son emploi et les valeurs professionnelles que cet emploi lui permet d'atteindre.<sup>2</sup> Connus pour s'influencer négativement,<sup>3</sup> et parce qu'ils permettent de mieux comprendre l'état de bien-être des travailleurs, ces deux indicateurs représentent aujourd'hui deux enjeux majeurs pour les entreprises.

Diverses études ont montré que le statut professionnel influence le stress<sup>4</sup> et la satisfaction au travail,<sup>5</sup> en faisant un facteur commun. Des facteurs sociodémographiques comme l'âge, l'ancienneté et le sexe présentent également cette particularité.<sup>5-6</sup>

A notre connaissance, il n'existe aucune évaluation de l'influence du statut professionnel, notamment managérial, sur les deux dimensions du bien-être au travail représentées par le stress et la satisfaction au travail, à court comme à long terme. Le rôle de manager étant associé avec des tâches et des responsabilités plus importantes, mais aussi à un risque accru pour la santé,<sup>7</sup> il semble pertinent de comprendre si les managers peuvent être affectés différemment des employés, et comment cela peut persister ou non dans le temps.

Ainsi, l'objectif principal de notre étude était d'évaluer l'influence bivariée du statut professionnel, en particulier du statut managérial, sur le stress et la satisfaction des travailleurs. Les objectifs secondaires étaient (i) d'explorer l'évolution simultanée du stress et de la satisfaction au travail, (ii) d'évaluer leurs influences réciproques et (iii) de mesurer l'influence d'autres facteurs sociodémographiques, à savoir l'âge, l'ancienneté et le sexe, sur le stress et la satisfaction des travailleurs, au cours du temps et selon leur statut professionnel, à savoir manager ou employé.

### Références principales (7 maxi)

1. Chrousos, G. P. (2009). Stress and disorders of the stress system. *Nature reviews endocrinology*, 5(7), 374-381.

2. Locke, E. A. (1976). The nature and causes of job satisfaction. *Handbook of industrial and organizational psychology*.

3. Pecino, V., Mañas, M. A., Díaz-Fúnez, P. A., Aguilar-Parra, J. M., Padilla-Góngora, D., & López-Liria, R. (2019). Organisational climate, role stress, and public employees' job satisfaction. *International journal of environmental research and public health*, 16(10), 1792
4. Kim, K. W., & Cho, Y. H. (2020). The moderating effect of managerial roles on job stress and satisfaction by employees' employment type. *International Journal of Environmental Research and Public Health*, 17(21), 8259.
5. Tarcan, M., Hikmet, N., Schooley, B., Top, M., & Tarcan, G. Y. (2017). An analysis of the relationship between burnout, socio-demographic and workplace factors and job satisfaction among emergency department health professionals. *Applied nursing research*, 34, 40-47.
6. Gazioglu, S., & Tansel, A. (2006). Job satisfaction in Britain: individual and job related factors. *Applied economics*, 38(10), 1163-1171.
7. Ikesu, R., Miyawaki, A., Svensson, A. K., Svensson, T., Kobayashi, Y., & Chung, U. I. (2021). Association of managerial position with cardiovascular risk factors: A fixed-effects analysis for Japanese employees. *Scandinavian journal of work, environment & health*, 47(6), 425.

## **Annexe D : Présentation orale au Small Meeting d’hiver du LAPSCO (2024)**

### **Indice de bien-être psychosocial et arrêt maladie sur le lieu de travail : modélisation par équation structurelle des données Wittyfit**

**Rémi COLIN**

« La santé est un état de complet bien-être physique, mental et social, et ne consiste pas seulement en une absence de maladie ou d’infirmité ». Cette définition, inscrite dans le Préambule de la Constitution par l’OMS pour la toute première fois en 1946 [1], et toujours présente à ce jour, met en exergue l’aspect multidimensionnel de la santé. Parmi les trois axes qui définissent la santé humaine, le bien-être social d’un individu se caractérise par un état de satisfaction, de prospérité, et de qualité de vie au sein d’une société. En fait, le bien-être social, ou plus largement le bien-être psychosocial [2], se rapproche de la notion de qualité de vie [3]. Cette seconde notion est étroitement liée au travail, si bien qu’on associe souvent qualité de vie et conditions de travail, formant ensemble le concept de qualité de vie au travail. L’amélioration de la qualité de vie au travail est devenue un enjeu majeur pour les entreprises désireuses de fidéliser leur personnel, d’attirer de nouveaux travailleurs et de maintenir un niveau élevé d’engagement et de bien-être chez leurs employés.

La qualité de vie au travail englobe de nombreux facteurs susceptibles d’affecter le bien-être psychologique et social des travailleurs, autrement dit leur bien-être psychosocial. Ces facteurs, qu’on appelle facteurs psychosociaux, sont connus, par exemple, pour exercer une influence sur le risque d’absentéisme au travail, en particulier sur le risque d’arrêt maladie [4]. En plus d’avoir un impact significatif sur la santé financière et la productivité des organisations, les arrêts maladie représentent un risque considérable pour la santé des travailleurs [5].

Le bien-être psychosocial est un concept large et complexe, et il semble donc pertinent d’étudier son effet potentiel sur l’absentéisme pour maladie en milieu professionnel. Ainsi, l’objectif principal de cette étude était de construire un modèle d’équation structurelle d’un indice de bien-être psychosocial englobant dix facteurs psychosociaux : la satisfaction au travail, l’ambiance de travail, la reconnaissance, l’équilibre entre vie personnelle et vie professionnelle, le sens, l’organisation du travail, les valeurs professionnelles, la charge de travail, l’autonomie et le stress. L’objectif secondaire était de mesurer le bien-être psychosocial des travailleurs à l’aide de l’indice ainsi créé et d’étudier son influence à la fois sur le risque et sur la durée des arrêts maladie.

## REFERENCES

1. International Health Conference. Constitution of the World Health Organization. 1946. Bulletin of the World Health Organization. 2002;80(12):983–4.
2. Larson JS. The World Health Organization's definition of health: social versus spiritual health. Soc Indic Res. 1996;38(2):181–92.
3. Eiroa-Orosa FJ. Understanding psychosocial wellbeing in the context of complex and multidimensional problems. Int J Environ Res Public Health. 2020;17(16):5937.
4. Niedhammer I, Chastang JF, Sultan-Taïeb H, Vermeylen G, Parent-Thirion A. Psychosocial work factors and sickness absence in 31 countries in Europe. Eur J Public Health. 2013;23(4):622–9.
5. Johnston DA, Harvey SB, Glozier N, Calvo RA, Christensen H, Deady M. The relationship between depression symptoms, absenteeism and presenteeism. Journal of Affective Disorders. 2019;256:536–40.

## Annexe E : Présentation orale aux Journées de l'Ecole Doctorale SVSAE

(2023)



**Family name:** COLIN

**First name:** Rémi

**Year of PhD:** 3<sup>rd</sup>

**Research team and Laboratory:** Équipe 3 : Stress  
Physiologique et Psychosocial, Bien-Être, Laboratoire de  
Psychologie Sociale et Cognitive (LaPSCo UMR 6024 UCA-  
CNRS)

**PhD supervisor(s):** Frédéric DUTHEIL, Bruno PEREIRA

**Email:** Remi.COLIN@doctorant.uca.fr

### Psychosocial well-being index and sick leave in the workplace: a structural equation modelling of Wittyfit data.

Rémi Colin-Chevalier<sup>1\*</sup>, Bruno Pereira<sup>2</sup>, Samuel Dewavrin<sup>3</sup>, Thomas Cornet<sup>3</sup>, Julien S Baker<sup>4</sup>, & Frédéric Dutheil<sup>1</sup>.

<sup>1</sup> CNRS, LaPSCo, Physiological and Psychosocial Stress, Université Clermont Auvergne, CHU Clermont-Ferrand, Occupational and Environmental Medicine, F-63000 Clermont-Ferrand, France

<sup>2</sup> Biostatistics Unit, DRCl, CHU Clermont-Ferrand, F-63000 Clermont-Ferrand, France

<sup>3</sup> Cegid, F-69009 Lyon, France

<sup>4</sup> Center for Health and Exercise Science Research, Hong Kong Baptist University, Kowloon Tong, Hong Kong 999077, China

**Background:** Psychosocial well-being, which assesses emotional, psychological, social, and collective well-being, could help measure risk and duration of sick leave in workers.

**Objective:** This study aims to build a structural equation model of a psychosocial well-being index (PWI) based on 10 psychosocial factors and investigate its association with sick leave.

**Methods:** Data of workers using Wittyfit was collected in 2018. Psychosocial factors (job satisfaction, atmosphere, recognition, work-life balance, meaning, work organization, values, workload, autonomy, and stress) were self-assessed using health-related surveys, while sick leave records were provided by volunteer companies.

**Results:** A total of 1,399 workers were included in the study (mean age: 39.4±9.4, mean seniority: 9.2±7.7, 49.8% of women, 12.0% managers). The prevalence of absenteeism was 34.5%, with an average of 8.48±28.7 days of sick leave per worker. Structural equation modelling facilitated computation of workers' PWI (AIC: 123,016.2, BIC: 123,231.2, RMSEA: 0.03). All factors, except workload ( $p=0.9$ ), were influential, with meaning ( $\beta=0.72$ , 95CI 0.69 to 0.74), values (0.69, 0.67 to 0.70) and job satisfaction (0.64, 0.61 to 0.66) being the main drivers ( $p<0.001$ ). Overall, psychosocial well-being was found to be a protective factor for sick leave, with a 2% decreased risk (OR=0.98, 95CI 0.98 to 0.99,  $p<0.001$ ) and duration (IRR=0.98, 95CI 0.97 to 0.99,  $p<0.001$ ) per PWI point.

**Conclusions:** The PWI provides a measure of psychosocial well-being and helps predict sick leave in the workplace. This new indicator could be used to analyze the association between psychosocial well-being and other health outcomes.

**Trial registration:** Clinicaltrials.gov: NCT02596737.

**Keywords:** psychosocial well-being; psychosocial factors; sick leave; structural equation modelling



## Rémi Colin

# Facteurs psychosociaux et économiques en milieu professionnel à l'heure de la médecine 4P : évolution longitudinale et prédictive du comportement des travailleurs

Thèse d'université – Clermont-Ferrand – 2024

---

### Résumé

La santé englobe le bien-être physique, mental et social, transcendant l'absence de maladie. Cette vision multidimensionnelle de la santé souligne l'importance du suivi à long terme de la santé des individus. Le bien-être social, ou plus largement psychosocial, associé à la qualité de vie au travail (QVT), est essentiel pour garantir la santé des travailleurs. Les risques psychosociaux (RPS) désignent des facteurs liés au travail pouvant influencer la santé et le bien-être des individus. Les RPS peuvent avoir des conséquences graves sur la santé des travailleurs, entraînant des problèmes tels que la dépression ou le burnout, mais aussi sur la productivité des entreprises, entraînant des pertes économiques conséquentes. Les employeurs peuvent lutter contre ces risques en évaluant les RPS sur le lieu de travail et en mettant en place des actions de prévention avec l'implication des salariés, tel que recommandé par la médecine 4P (prédictive, préventive, personnalisée et participative). Wittyfit, la plateforme dédiée à l'amélioration de la QVT, permet aux employeurs et aux salariés de faire face à ces défis. L'émergence des nouvelles technologies a révolutionné la santé au travail, permettant des soins de santé plus personnalisés grâce à la collecte et à l'analyse de données du monde réel. Ces données incluent une variété d'informations sur la santé des individus qui sont essentielles pour évaluer le bien-être des travailleurs, identifier les risques auxquels ils sont exposés, et suivre l'évolution de leur état de santé. Cependant, l'analyse des données du monde réel de santé au travail est complexe en raison de leur nature longitudinale, de la présence de clusters naturels de données, et des données manquantes. Cette thèse se concentre sur l'analyse des données de santé au travail pour améliorer la QVT et le bien-être des travailleurs. L'hypothèse principale est que l'étude de ces données permet d'identifier les facteurs psychosociaux et économiques en milieu professionnel. Les enjeux médico-économiques portent sur l'évaluation des RPS, l'identification des facteurs de risques, et l'étude de leurs conséquences sur la performance des entreprises avec un focus sur l'absentéisme. Les enjeux statistiques portent sur le développement et l'application de modèles statistiques complexes répondant aux défis méthodologiques de l'analyse de données de santé au travail.

**Mots-clés :** qualité de vie au travail, risques psychosociaux, productivité des entreprises, médecine 4P, données de santé au travail

### Abstract

Health encompasses physical, mental, and social well-being, transcending the absence of disease. This multidimensional vision of health underlines the importance of long-term monitoring of individual health. Social well-being, or more broadly psychosocial well-being, associated with quality of life at work (QWL), is essential to guarantee workers' health. Psychosocial risks (PSR) refer to work-related factors that can influence the health and well-being of individuals. PSR can have serious consequences for workers' health, leading to problems such as depression or burnout, but also for business productivity, resulting in substantial economic losses. Employers can fight against these risks by assessing RPS in the workplace and implementing preventive actions with employee involvement, as recommended by P4 medicine (predictive, preventive, personalized and participative). Wittyfit, the platform dedicated to improving QWL, enables employers and employees to meet these challenges. The emergence of new technologies has revolutionized occupational health, enabling more personalized healthcare through the collection and analysis of real-world data. This data includes a variety of information on individual health that is essential for assessing workers' well-being, identifying the risks to which they are exposed, and tracking changes in their health status. However, the analysis of real-world occupational health data is complex due to its longitudinal nature, the presence of natural clusters of data, and missing data. This thesis focuses on the analysis of occupational health data to improve QWL and worker well-being. The main hypothesis is that the study of these data enables the identification of psychosocial and economic factors in the workplace. The medico-economic issues concern the evaluation of PSR, the identification of risk factors, and the study of their consequences on company performance, with a focus on absenteeism. Statistical issues concern the development and application of complex statistical models to meet the methodological challenges of occupational health data analysis.

**Keywords:** quality of life at work, psychosocial risks, business productivity, P4 medicine, occupational health data