Comités de Retour d'Expérience et culture de sécurité des soins

Bastien Boussat

▶ **To cite this version:**

Bastien Boussat. Comités de Retour d'Expérience et culture de sécurité des soins. Médecine humaine et pathologie. Université Grenoble Alpes, 2018. Français. NNT : 2018GREAS008 . tel-02382764

HAL Id: tel-02382764
https://theses.hal.science/tel-02382764

Submitted on 27 Nov 2019

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.

Communauté
UNIVERSITÉ Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTE UNIVERSITE GRENOBLE ALPES

Spécialité : **Modèles, méthodes et algorithmes en biologie**

Arrêté ministériel : 25 mai 2016

Présentée par

Bastien BOUSSAT

Thèse dirigée par **Patrice FRANÇOIS**

préparée au sein du **Laboratoire TIMC-IMAG, équipe BCM**
dans **l'École Doctorale Ingénierie de la Santé, la Cognition et l'Environnement**

Comités de Retour d'Expérience et culture de sécurité des soins

Thèse soutenue publiquement le **21 Mars 2018**,
devant le jury composé de :

M. Pascal AUQUIER
PUPH, Assistance Publique-Hôpitaux de Marseille, Président
M. Antoine DUCLOS
PUPH, Hospices Civils de Lyon, Rapporteur
M. Bruno FALISSARD
PUPH, Assistance Publique- Hôpitaux de Paris, Rapporteur
M. Olivier FRANÇOIS
PU, Grenoble INP TIMC-IMAG, Examinateur
M. Patrice FRANÇOIS
PUPH, CHU de Grenoble Alpes, Directeur de thèse
M. Thomas PERNEGER
PUPH, Hôpitaux Universitaires de Genève, Examinateur

SOMMAIRE

Résumé : Le Comité de Retour d'Expérience (CREX) est un dispositif de gestion de la sécurité des patients en équipe, engageant les professionnels de santé dans la pratique collective de l'analyse des causes racines des événements indésirables associés aux soins. Ce dispositif, développé en France en 2005, est désormais implanté dans la majorité des établissements de santé français. Curieusement, après 10 ans de développement dans des milliers d'équipes médicales, aucun travail scientifique publié ne s'était penché sur leur fonctionnement et leurs effets sur la sécurité des patients. Mon travail de thèse a permis d'évaluer le fonctionnement des CREX, ainsi que leur association aux dimensions de la culture de sécurité des professionnels de santé. Nous avons montré que les CREX sont appropriés pour fonctionner dans les diverses spécialités d'un établissement de santé. De plus, ce dispositif permet aux soignants de réaliser des analyses des causes racines, et de décider collectivement d'actions visant à corriger les défaillances profondes des organisations. Afin d'étudier les relations entre la participation à un CREX et les dimensions de la culture de sécurité des soignants, nous avons effectué une enquête auprès de 3888 professionnels de santé, en utilisant le questionnaire Hospital Survey On Patient Safety Culture (HSOPS), dont nous avons précisé les propriétés métrologiques (psychométrie, variabilité des méthodes d'agrégation des scores, méthodes d'imputation des données manquantes). Nous avons mis en évidence des relations favorables entre la participation à un CREX et la culture de sécurité des soignants. Les CREX permettent de favoriser le travail en équipe, la multidisciplinarité, ainsi que de modifier la culture de l'erreur. Malgré ses limites, notre travail contribue à mieux connaitre les CREX et leur place dans les dispositifs de gestion de la sécurité des patients. Cette thèse plaide pour le développement d'une recherche centrée sur l'implication des professionnels de santé dans la gestion de la sécurité des patients.

Mots-clés : Comité de Retour d'Expérience ; Sécurité du patient ; Analyse des causes racines ; Culture de sécurité ; Gestion des risques ; Organisation des services de santé

Abstract: The Experience Feedback Committee (EFC) is a tool designed to involve medical teams in patient safety management, through root cause analysis of adverse events within the team. This program was created in 2005, and was implemented in the vast majority of French hospitals. Despite its wide implementation in thousands of French medical teams, the EFC still lacks scientific evaluation. To our knowledge, the EFC framework and the association between EFC and patient safety culture have never been assessed. In this context, our work aimed to determine whether patient safety culture, as measured by the Hospital Survey On Patient Safety Culture (HSOPS), differed regarding care provider involvement in EFC activities. Using the original data from a cross-sectional survey of 3,888 employees at a single university hospital in France, we analyzed the differences in HSOPS dimension scores according involvement in EFC activities. We also specified the metrological properties of the transcultural adaptation of the HSOPS into French (psychometrics properties, variability of scoring strategies and missing data imputation methods). Our findings suggest that EFC participation may improve patient safety culture, teamwork and non-blame oriented processes. Despite several limitations, our study contributes to a better understanding of EFC and its position amongst patient safety systems. This thesis advocates research developments centered on health care givers involvement in patient safety management.

Keywords: Experience Feedback Committee; Patient safety; Root cause analysis; Patient safety culture; Risk management; Health services administration

1. INTRODUCTION

Depuis le début des années 2000 s'est effectuée une véritable prise de conscience de la communauté médicale et scientifique, quant à la nécessité d'améliorer la sécurité des soins dans les établissements de santé (BMJ 2000). Aussi de nombreuses initiatives ont été mises en place par les organismes en charge de la santé, sous l'impulsion notamment de l'Alliance mondiale pour la sécurité des patients, ou encore de l'Agency for Healthcare Research and Quality aux Etats-Unis (William "Bill" Frist 1999, OMS 2004). Malgré les efforts déployés pour améliorer la sécurité des patients, une étude publiée en 2016 par une équipe de l'Université John Hopkins classait les erreurs médicales au 3^{e} rang des causes de décès hospitaliers aux Etats-Unis, devant les maladies respiratoires ou les accidents vasculaires cérébraux (Makary and Daniel 2016).

En France, les enquêtes nationales sur les événements indésirables graves associés aux soins (ENEIS) ne montrent aucune amélioration des densités d'incidences de ces événements entre 2004 et 2009 (Michel, Quenon et al. 2007, Michel, Minodier et al. 2010). Un des principaux axes de bataille pour la sécurité des soins se situe au niveau des services de santé, dans les équipes médicales (Darzi 2009, Pronovost and Marsteller 2011). En effet, de nombreux obstacles sont liés à la sociologie des organisations de soins et il apparait essentiel que les professionnels de santé partagent un ensemble de valeurs communes plaçant la sécurité de leurs patients au centre de leurs pratiques (Flin, Burns et al. 2006, Zohar, Livne et al. 2007). Ces valeurs, qui influencent les comportements et les attitudes des cliniciens et des équipes soignantes en fixant des règles et repères communs, doivent être orientées vers un même objectif d'amélioration de la sécurité des patients. Cet ensemble de valeurs fonde le concept de culture de sécurité des patients, dont la promotion a été placée par l'Organisation Mondiale pour la Santé au 3^{e} rang des 20 objectifs prioritaires pour la recherche en santé des pays développés (Bates, Larizgoitia et al. 2009).

Dès lors, diverses initiatives visant à développer cette culture chez les professionnels de santé ont vu le jour. C'est notamment le cas des TeamSTEPPS (Ward, Zhu et al. 2015, Duclos, Peix et al. 2016) et des Comprehensive Unit-based Safety Program (CUSP) aux Etats-Unis (Weaver, Lofthus et al. 2015), qui sont des dispositifs situés au niveau des équipes médicales, dont l'objectif est d'améliorer le travail en équipe, la communication, le leadership, les comportements des équipes faces aux événements indésirables associés aux soins (EIAS). Un programme similaire, intitulé Comité de Retour d'Expérience (CREX), a été créé en France en 2005, avec l'idée d'engager les professionnels de terrain dans la gestion de la sécurité des soins (Francois, Sellier et al. 2013). Cette méthode, inspirée de l'aviation civile, permet aux équipes médicales d'analyser collectivement les EIAS auxquels ils sont confrontés en routine, et de mettre en place des actions visant à corriger les défaillances de leur organisation. Fondés sur le retour d'expérience et l'analyse des causes racines, les CREX ont connu un important déploiement sous l'impulsion de la Haute Autorité de Santé (HAS), et sont désormais présents dans la grande majorité des établissements de santé français. Toutefois, les CREX n'ont fait l'objet que de peu d'évaluation scientifique. Fort de ce constat, le projet ACREX (Apport des CREX) a été retenu et financé par la Direction Générale de l'Offre de Soins (DGOS), dans le cadre de l'appel d'offre national PREPS (Programme de Recherche sur la Performance du Système de Soins). Ce projet visait à étudier l'association entre la participation à CREX et la culture de sécurité, au travers d'une enquête menée auprès de l'ensemble des professionnels de santé du Centre Hospitalier Universitaire de Grenoble.

La présente thèse d'université s'appuie sur les données de ce projet de recherche ACREX, et a pour objectif principal d'explorer la relation entre la participation à un CREX et les différentes dimensions de la culture de sécurité du patient. Afin d'évaluer cette relation, notre stratégie de recherche s'est articulée autour de deux principaux axes de travail, que sont l'étude du

fonctionnement des CREX implantés au CHU de Grenoble, et une enquête de culture de sécurité réalisée auprès des professionnels de santé de l'établissement.

Dans un premier chapitre seront présentés le concept de CREX ainsi que les résultats de l'étude du fonctionnement des CREX. Dans un second chapitre seront présentés les travaux relatifs aux propriétés du questionnaire servant à mesurer la culture de sécurité. Enfin, le troisième chapitre présentera les résultats concernant l'association entre la participation à un CREX et les dimensions de la culture de sécurité.

2. CHAPITRE I : Etude du fonctionnement des Comités de Retour d'Expérience

2.1. Introduction : l'origine des CREX

Un comité de retour d'expérience est une méthode de gestion de la sécurité des soins destinée à une équipe médicale. Le CREX est un comité pluri professionnel qui prend en charge les signalements des EIAS survenus dans le service. Le CREX étudie les événements, définit des priorités, organise l'analyse méthodique des événements et propose des actions visant à améliorer la sécurité des soins.

Cette méthode a été adaptée aux structures de soins à partir des systèmes de sécurité de l'aviation civile. Cette transposition a été réalisée par la société Air France consulting, dans le cadre des interventions de la Mission nationale d'expertise et d'audits hospitaliers (Meah) (MeaH 2008). Le projet a été expérimenté entre 2005 et 2007 dans trois services de radiothérapie (Woynar, Debouck et al. 2007, Lartigau, Coche-Dequeant et al. 2008). Cette expérience a été un succès qui a conduit à généraliser les CREX dans tous les services de radiothérapie et à envisager des mécanismes de mutualisation des problèmes et des solutions rapportées par les différentes équipes (Lartigau, Vitoux et al. 2009, Debouck, Petit et al. 2010). Dans un rapport de 2011, l'Institut de radioprotection et de sûreté nucléaire (IRSN) soulignait l'importance des CREX dans les progrès constatés dans le domaine de la sécurité en radiothérapie (Thellier 2011). Dans le même temps, la Meah, a proposé d'étendre l'implantation dans d'autres services médicaux ou médicotechniques (MeaH 2008). Dès 2007, un consultant de la société Air France consulting, mandaté par la Meah, a conduit la mise en place de CREX dans quatre services du CHU de Grenoble : pharmacie, réanimation médicale, maladie infectieuses et cardiologie (Francois, Sellier et al. 2013). Devant le succès de cette implantation, les professionnels ayant adhéré aux CREX et les groupes s'étant pérennisés au-delà de l'intervention de la Meah, le dispositif a été proposé à d'autres services de l'établissement. En

2014, on dénombrait 20 CREX, dans diverses spécialités comme la cardiologie, la psychiatrie, la chirurgie digestive ou encore la pédiatrie. Depuis 2010, la Haute Autorité de Santé promeut l'implantation des CREX dans tous les établissements de santé en France, qui est désormais évalué lors de la procédure de certification des établissements de santé (Bally and Chevalier 2014).

2.2. Fonctionnement du CREX

2.2.1. Le comité

Le CREX est un groupe de six à 12 professionnels représentant les différents métiers du service : médecins, cadres de santé, personnels paramédicaux, secrétaires, etc. Il est coordonné par un responsable et il désigne en son sein un secrétaire de séance chargé de la prise de notes et de la rédaction des comptes rendus. Le comité se réunit régulièrement, généralement une fois par mois, selon un jour et un horaire fixe, afin que l'activité s'intègre dans la routine du service.

2.2.2. Les réunions du comité

Les réunions durent entre une heure et une heure et demie et sont standardisées selon le déroulement suivant :

1) examen des signalements du mois : les signalements d'EIAS ou d'événements porteurs de risques reçus dans le mois précédent la réunion, sont énoncés et examinés brièvement ;

2) choix d'un événement à traiter : le groupe doit choisir parmi les événements examinés, celui qui sera analysé en profondeur. Ce choix s'appuie implicitement sur les critères de criticité, qui est le produit de la fréquence par la gravité de l'événement. Ce choix s'effectue par consensus du groupe, ou par un vote simple ou pondéré ;

3) désignation du « pilote » de l'analyse : le comité désigne la personne chargée d'investiguer l'événement choisi et d'élaborer des propositions d'action. Ce pilote doit utiliser la méthode Orion, méthode d'analyse systémique des événements indésirables, développée lors de la

création des CREX par Air France consulting. Il est préférable que le pilote ne soit pas impliqué directement dans l'événement à analyser.

4) Ecoute du rapport d'analyse du mois précédent : le pilote désigné lors de la précédente réunion pour investiguer l'événement prioritaire fait le rapport de son analyse et propose des actions d'amélioration ;

5) élaboration d'un plan d'actions : après débat, le comité choisit et planifie la ou les actions à mettre en œuvre, c'est-à-dire inscrire chaque action dans un calendrier, fixer les objectifs, désigner un responsable, assurer un suivi et prévoir l'évaluation des effets ;

6) suivi des actions en cours : le comité parcourt la liste des actions qui ont été planifiées lors des CREXs antérieurs et s'informe de l'état d'avancement de chaque action et des éventuelles difficultés rencontrées. Une action est supprimée de la liste dès lors qu'elle est considérée comme terminée.

2.2.3. Analyse des causes racines des EIAS : méthode Orion

Depuis le début des années 2000, l'analyse des causes racines s'est imposée comme la méthodologie de référence pour investiguer les EIAS et tenter de les prévenir. Basée sur l'approche systémique du modèle de James Reason, l'analyse des causes racines vise à détecter les facteurs organisationnels complexes favorisant les erreurs humaines, considérant celle-ci comme des conséquences plutôt que des causes (Reason 2000). La méthode d'analyse des causes racines la plus célèbre est la méthode ALARM développée en 1999 par un groupe de recherche de l'Imperial College of London (Vincent, Taylor-Adams et al. 2000). Lors de la création des CREX, la société Air France consulting a développé une méthode d'analyse des causes racines, proche de la méthode ALARM, intitulée méthode Orion, qui impose de procéder par étapes (Debouck, Rieger et al. 2012) :

1) collecter les données. Le pilote doit rassembler les informations sur l'événement, son contexte de survenue et la façon dont il a été géré. Ces informations sont recherchées dans des documents (dossier médical, dossier de soins, formulaires, etc.) et auprès des acteurs et témoins de l'événement. Le pilote peut procéder par entretiens individuels ou provoquer une réunion de « débriefing » dont le compte-rendu doit être validé par les participants ;
2) décrire la chronologie de l'événement. Le pilote doit recenser et trier les faits pour les ordonner dans le temps et reconstituer le « film » de l'événement (avant, pendant, après). A partir de cette chronologie, il faut repérer les défaillances, les actions inappropriés, les écarts par rapports à des standards ;
3) identifier les causes des erreurs. Une cause est ici un élément du contexte de l'action qui a favorisé la survenue de l'erreur. Il s'agit de comprendre pourquoi s'est produit l'écart en explorant systématiquement différents domaines : l'organisation, les conditions de travail, le fonctionnement de l'équipe, les comportements individuels, les procédures, le patient, la politique de l'équipe ;
4) identifier les facteurs d'influence. Plus on s'éloigne de l'événement, plus la relation entre causes et effet devient incertaine. On recherche alors les facteurs d'influence qui sont les fragilités du système qui ont contribué aux causes de l'événement. Ces facteurs sont recherchés dans les même domaines que les causes ;
5) proposer des actions correctives. Le pilote de l'analyse doit élaborer, avec les acteurs concernés, des actions visant à réduire le risque des défaillances et à mettre en place des défenses ou barrières de sécurité. Chaque action doit être décrite avec une proposition de l'entité ou la personne la mieux placée pour la mettre en œuvre et une estimation de la faisabilité ;
6) rédiger le rapport d'analyse. Les résultats de l'investigation et les propositions sont rapportées dans un document dont le format standardisé correspond aux étapes de l'analyse.

2.3. Evaluation des CREX

Comme évoqué précédemment, les CREXs ont été initialement développés et mis en place dans des services de radiothérapie, puis diffusés dans de nombreux hôpitaux et dans de nombreux services d'autres spécialités médicales et chirurgicales. Quelques travaux ont été publiés, essentiellement descriptifs, lors de l'implantation des CREX en radiothérapie, alors que le déploiement des CREX dans d'autres spécialités n'avait jamais été évalué (Francois, Sellier et al. 2013). Les questions posées sur le fonctionnement et l'utilité des CREX a conduit notre équipe à étudier les CREX implantés dans certaines spécialités (services d'urgences, d'anesthésie réanimation, de pharmacie hospitalière, et de neuropsychiatrie) (Lecoanet, Sellier et al. 2013, Caporossi, Brudieu et al. 2014, Boussat, Bougerol et al. 2015). Nous illustrons cette approche par la présentation de deux articles. Le premier rapporte le fonctionnement d'un CREX implanté dans un service de psychiatrie, le second rapporte le fonctionnement global des CREX de notre établissement.

2.3.1. Monographie du CREX de Neuropsychiatrie : article publié – Experience Feedback Committee : a management tool to improve patient safety in mental health

Compte-tenu de l'implantation des comités dans diverses spécialités, l'une des questions était de savoir si les comités étaient adaptés pour traiter les problématiques de sécurité des patients dépassant le cadre spécifique de la radiothérapie. En effet, si les EIAS sont souvent communs aux différentes spécialités médicales et chirurgicales, certains problèmes de sécurité peuvent être spécifiques. C'est notamment le cas pour les services de psychiatrie, qui font face à des événements liés à des problèmes comportementaux comme les violences, les fugues ou les tentatives de suicides (Nath and Marcus 2006).

Objectif : L'objectif de cette étude était d'analyser le fonctionnement d'un CREX implanté dans un service hospitalier de psychiatrie, mais aussi d'explorer sa contribution aux problématiques spécifiques rencontrées dans le domaine de la santé mentale.

Méthode : Nous avons conduit une étude descriptive basée sur tous les documents produits par le CREX du pôle de Neuropsychiatrie du CHU de Grenoble, entre mars 2010 et janvier 2013. Nous avons analysé tous les signalements d'EIAS, les comptes-rendus des réunions et les rapports Orion produits par le comité. Les EIAS étaient classés selon une classification internationale élaborée par un groupe de l'Alliance Mondiale pour la Sécurité des Patients (World Alliance For Patient Safety Drafting, Sherman et al. 2009), et les autres documents étaient analysés à partir de formulaires standardisés incluant les étapes formelles des CREX établies par la société Air France consulting. Les analyses ont été conduites indépendamment par deux investigateurs (PF et BB), et les différences de classement étaient discutées jusqu'à l'obtention d'un consensus.

Résultats : Sur les 38 mois de fonctionnement analysés, 30 réunions ont été organisées, auxquelles ont participé 22 professionnels de santé représentant les différentes professions du pôle. 475 EIAS ont été examinés en équipe, dont la majorité (92%) était sans conséquence médicale directe pour les patients. 11 EIAS ont fait l'objet d'une analyse des causes racines effectuée selon la méthode Orion, conduisant à la décision de 21 actions correctrices, dont 8 s'attaquaient à des problématiques spécifiques de la santé mentale (exemple : formations pour réagir face aux situations de violence ou encore protocole pour guider la prise en charge des tentatives des suicide).

Conclusion : Cette étude nous a permis d'établir que le CREX était adapté pour traiter des EIAS spécifiques aux secteurs de psychiatrie. De plus, nous avons reporté une bonne adhésion des participants au cours des 3 années de suivi, avec la mise en place de nombreuses actions d'amélioration de la sécurité des patients. Cependant, nous avons noté une diminution des rapports d'analyse ainsi qu'une résurgence de certains types EIAS au cours du temps, mettant en évidence certaines faiblesses de la méthode CREX.

Boussat *et al. Ann Gen Psychiatry (2015) 14:23*
DOI 10.1186/s12991-015-0062-2

ANNALS OF GENERAL PSYCHIATRY

PRIMARY RESEARCH Open Access

Experience Feedback Committee: a management tool to improve patient safety in mental health

Bastien Boussat[1,2*], Thierry Bougerol[3†], Olivier Detante[4†], Arnaud Seigneurin[1,2†] and Patrice François[1,2†]

Abstract

Background: A management tool, called the Experience Feedback Committee, has been applied for patient safety and successfully used in medical departments. The purpose of this study was to analyse the functioning of an Experience Feedback Committee in a psychiatric department and to explore its contribution to the particular issues of patient safety in mental health.

Methods: We conducted a descriptive study based on all the written documents produced by the Experience Feedback Committee between March 2010 and January 2013. The study was conducted in Grenoble University Hospital in France. We analysed all reported incidents, reports of meetings and event analysis reports. Adverse events were classified according to the Conceptual Framework for the International Classification for Patient Safety.

Results: A total of 30 meetings were attended by 22 professionals including seven physicians and 12 paramedical practitioners. We identified 475 incidents reported to the Experience Feedback Committee. Most of them (92 %) had no medical consequence for the patient. Eleven incidents were investigated with an analysis method inspired by civil aviation security systems. Twenty-one corrective actions were set up, including eight responses to the specific problems of a mental health unit, such as training to respond to situations of violence or management of suicide attempts.

Conclusions: The Experience Feedback Committee makes it possible to involve mental healthcare professionals directly in safety management. This tool seems appropriate to manage specific patient safety issues in mental health.

Keywords: Risk management, Mental health, Neuropsychiatry, Quality improvement, Interdisciplinary communication

Background

Patient safety has become a public health priority in the past 15 years since the publication of major epidemiological studies on healthcare-related adverse events [1–5]. The last Canadian national study estimated that 7.5 adverse events occurred for 100 hospital admissions, including a high proportion of preventable events and adverse events leading to death [6].

Even if many adverse events are similar in all medical units, there are specific patient safety issues in mental health. Indeed, neuropsychiatric units face events caused by behavioural problems such as violence, absconding, self-harm and suicide attempts [7–11]. Mental health adverse events result from multiple factors, mixing human behaviour risks and healthcare organizational weaknesses. Considering the lack of readily available information to guide patient safety systems in mental health, the improvement of management tools is essential to promote a patient safety culture among healthcare professionals [11–13].

Since the 1970s, civil aviation has developed operating experience feedback to improve passenger safety. Air transport safety systems require that any incident, even minor, must be treated by a systemic analysis within the air crew. Inspired by those security systems,

*Correspondence: BBoussat@chu-grenoble.fr
†Thierry Bougerol, Olivier Detante, Arnaud Seigneurin and Patrice François contributed equally to this work
[2] Laboratory TIMC, UMR 5525, CNRS, Joseph Fourier University, Grenoble, France
Full list of author information is available at the end of the article

BioMed Central

© 2015 Boussat et al. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

a specific mechanism, called the Experience Feedback Committee (EFC), was created in 2005 to analyse adverse events within a medical team. In France, the method was adapted to healthcare facilities with the help of Air France Consulting and was successfully implemented in an emergency department and a radiotherapy unit [14–19]. The EFC is a team composed of professionals representing the diversity of the functions encountered in the medical unit. The EFC members meet monthly to examine adverse events reported to their medical unit. An event requiring a thorough analysis is chosen at each meeting according to its severity and frequency and corrective actions are suggested based on the results of the analysis. The main principles of the method are managing patient safety within a medical team and setting up corrective actions concerning latent factors that contributed to the occurrence of events or near-miss events.

The aim of this study was to describe the functioning of the EFC in a neuropsychiatric department and to discuss its contribution to the management of patient safety.

Methods

Study design

We conducted a descriptive study based on the written reports of the neuropsychiatric department EFC from its beginning in March 2010 until January 2013.

Setting

The study was conducted in a 1347-bed acute-care university hospital. The neuropsychiatry department has an annual patient volume of 4800 stays.

Adverse events and near-misses affecting patients (such as nosocomial infections, technical complications, negligence, diagnostic mishaps, therapeutic incidents, etc.) are reported by healthcare professionals, through a voluntary internal reporting system, to the hospital's central safety unit using a standardized report form. This unit is composed of a medical doctor, a pharmacist and an engineer specializing in quality management. The reports of events classified by severity and risk areas are presented during a weekly meeting involving representatives of the administration and professionals in charge of specific risk areas such as the risks associated with drugs (pharmacovigilance), nosocomial infections (infection vigilance), healthcare materials and devices (medical device vigilance), transfusion (haemovigilance), etc. The central safety unit directly investigates the most serious events and those involving several hospital departments. Other events are transmitted to the appropriate operator and to executives of relevant departments. For departments where an EFC has been implemented, the central safety unit addresses the reports of events to the EFC leader every month.

Neuropsychiatric department EFC

The neuropsychiatric department EFC was set up in March 2010 and works through a written procedure in accordance with the method proposed by Air France Consulting [17, 18, 20]. The Committee is composed of volunteer representatives of the various professions within the neuropsychiatric division. A few days before the committee meeting, the EFC leader receives a file with event reports concerning the neuropsychiatric division. Committee meetings are conducted according to a standardized framework: (1) reading the list of reported events, (2) choosing a priority event to investigate by consensus according to the criticality of each incident, (3) choosing the professional responsible for the investigation, (4) reviewing the analysis made of the event chosen the previous month, (5) choosing corrective actions and (6) monitoring on-going actions. The investigation is carried out during the month following the EFC by a designated person using a method, called ORION©, developed from methods of systemic analysis used in civil aviation and adapted to the healthcare domain by Air France Consulting [14–17]. Previously trained investigators must follow the main steps of the ORION© method to fill out a standardized report (Additional file 1: Appendix): collecting data and existing recommendations, describing the chronological facts that occurred before, during and after the event, describing the failures, looking for causes of errors and latent factors that could have contributed to the failures, setting up corrective actions and writing a report of the analysis. Causes and latent factors are sought in different areas such as political, organizational, working conditions, team functioning, procedures, actors and the patient.

Data collection

All written documents from the EFC of the neuropsychiatric division were analysed. The events reported were classified according to the source of the report, the type of event and the consequence for the patient using the International Classification for Patient Safety [21]. Written reports from meetings were analysed using a standardized form that included the theoretical steps of an EFC meeting and the contents of the ORION© analysis (as described above). All documents were analysed by two independent investigators. Differences in rating were discussed until a consensus was reached.

Statistical analysis

We reported the characteristics of the EFC's main functioning (meetings and participants), the adverse events reported and the analysis reports as medians and interquartile ranges (IQR; i.e., 25th and 75th percentiles) for continuous variables and number and percentages for

categorical variables. The analysis was performed using R version 3.0.1.

Results

The committee set up 30 meetings during the study period. A total of 22 professionals participated in the EFC (Table 1), including seven physicians, four head nurses, four auxiliary nurses, three nurses, one secretary, one physical therapist, one cleaning staff member and one quality engineer. The median number of attendants was eight (IQR 6–9) per meeting. A report was written for each meeting. Priority events were chosen in half of the meetings, analysis reports were presented and corrective actions were decided in more than one-third of the meetings (Table 1). The previous corrective actions were monitored in 19 meetings (63.3 %).

A total of 475 reported incidents were transmitted to the EFC (Table 2). A median number of 12 incidents (IQR 7–20) were discussed per meeting. Incidents were mainly (97.1 %) reported by a professional of the department and 93.3 % of them occurred inside the department. Reported incidents concerned mainly clinical administration (29.3 %) (including incidents in patient identification, patient transfer, admission, discharge), behaviour (24 %) (concerning patient or staff) and patient accidents (12 %). The majority of incidents had no clinical consequence for the patient (91.8 %) or the care process (70.1 %). In 29 cases (6.1 %), the reported event had a mild or moderate consequence for the patient (Table 2). Among the 20 events involving a mild consequence for the patient, ten events were related to a patient fall, four events concerned violence against the staff or another patient, three events were related to inadequate equipment and three events were suicide attempts. Among the nine events involving a moderate consequence, eight falls led to orthopaedic fracture or head trauma, and one admission error concerned a patient hospitalized in a corridor.

Fifteen priority incidents were chosen for investigation, including six incidents related to clinical administration, four incidents related to behaviour problems, four incidents related to infrastructure and fixtures and one to the keeping of archives. Four incidents chosen were not investigated. Among the 11 analyses carried out, nine reports were written, whereas two reports were only oral (Table 3). Three reports involved a problem of coordination with other hospital departments. First, an error of emergency transfer was reported for a patient hospitalized in a corridor who was transported to the intensive care unit for optimal surveillance because of loss of consciousness. Second, stretchers for an imaging examination emergency were recurrently unavailable. Third, organizational problems of the psychiatric consultation were reported. Three reports analysed adverse events associated with behavioural disorders: the inability of a professional to properly manage a patient with a suicide

Table 1 Main functioning characteristics of the Experience Feedback Committee of the neuropsychiatry department

	***N* = 22**	**%**
Participants		
Physicians	7	31.8
Head nurses	4	18.2
Auxiliary nurses	4	18.2
Nurses	3	13.6
Secretary	1	4.5
Quality engineer	1	4.5
Physical therapist	1	4.5
Cleaning staff member	1	4.5
Median number of participations per participant (IQR 25–75)	9	(5–15)
Median number of participants per meeting (IQR 25–75)	8	(6–9)
	***N* = 30**	**%**
Meetings		
Writing of minutes	30	100.0
Listening to the events reported during the previous month	30	100.0
Choosing a priority event to analyse during the following month	15	50.0
Listening to the analysis report from the event chosen the previous month	11	36.7
Deciding corrective actions	11	36.7
Following up the previous corrective actions	19	63.3

Table 2 Characteristics of the events reported during the Experience Feedback Committee meetings

Characteristics	*N* = 475	%
Incident type		
Clinical administration	139	29.3
Behaviour	114	24.0
Patient accidents	57	12.0
Infrastructure/building/fixtures	41	8.6
Medical device/equipment	30	6.3
Resources/organizational management	30	6.3
Nutrition	20	4.2
Clinical process/procedure	16	3.4
Medication/IV fluids	14	2.9
Documentation	11	2.3
Healthcare-associated infection	3	0.6
Blood/blood products	0	0.0
Oxygen/gas/vapour	0	0.0
Degree of Harm		
None, without care modification	343	72.2
None, with care modification	103	21.7
Mild	20	4.2
Moderate	9	1.9
Severe	0	0.0
Death	0	0.0
Report provider		
Staff from the neuropsychiatric department	461	97.1
Staff from another department	16	3.4
Place of the event		
In the neuropsychiatric department	443	93.3
In another department	32	6.7

risk in child psychiatry, an attempted suicide by strangulation with a phone cord and finally the investigation of repeated fugues in adult psychiatry. Three ORION© analyses were carried out to investigate a technical equipment failure including an ECG machine, the computer network and a power failure. The last reports concerned a loss of patient records, a delivery mistake for special meals and the fall of an elderly patient.

The expertise of the written reports showed that the ORION© method was frequently followed. The chronology of the facts and the identification of contributing or latent factors were described in 80 % of the cases. However, existing recommendations were only described in one-half of the cases. Twenty-six corrective actions were proposed by the professionals who performed the analyses. The committee decided to implement 21 actions. Written guidelines ($n = 11$) were the most common type of action (see box). Other actions included staff training, improvement of the availability of material resources and a deeper analysis of an event in the context of a medical thesis. For example, the ORION© report related to the fall of an elderly patient showed several factors contributing to the event: some factors related to the patient (inappropriate behaviour of a patient with a depressive syndrome, decreased alertness and reflexes due to anxiolytic treatment) and organizational factors (lack of assessment of the risk for falls at admission, lack of supervision due to a high level of department activity). The corrective actions selected were structured around these two main types of factors. A systematic screening of risk factors for falling at admission was established. Secondary prevention actions were also decided: pharmaceutical adaptation to reduce iatrogenic events and implementation of monitoring for high-risk patients by a team composed of a physiotherapist, an occupational therapist and a movement therapist. Finally, adjustments were made at the facility level with adjustable-height beds, night lights in the rooms and more convenient showers.

Box Corrective actions set up

- Guideline writing
  - Inpatient transfer from ED
  - Patient medical record management
  - Suicide risk assessment at admission
  - Management of patient with behaviour problems
  - Security guards system
  - Job profile of secretary
  - Failing risk assessment at admission
  - Protocol to prevent failing risk
  - User manual for ECG advise
  - Emergency consultation procedure
  - Electrical failure procedure
- Training
  - Respond to situations of violence
  - Management of elderly patients
  - Respond to suicide attempt
  - Proper management of medical records
- Material resources
  - Establishment of an isolation room
  - Change of ECG device
  - Loan of two ECG device
  - Implementation of a planning software for inpatient transport
- Research
  - Completion of a medical thesis on inpatient transfers in the hospital

Discussion

This study highlighted that the EFC implemented in the mental health department functions routinely with patient safety incidents analysed and corrective actions set up. The EFC method, which was successfully implemented in medical units, is also relevant to managing patient safety in mental health. Nath and Marcus

Table 3 Characteristics of the analysis reports and of the corrective actions

	N = 11	%
Analysis reports		
Written reports	9	81.8
Oral reports	2	18.2
Description of the data collection method	9	81.8
Individual interviews	9	81.8
Collective debriefing	1	9.1
Files	5	45.5
Area visits	6	54.5
Description of the chronology of facts	9	81.8
Description of existing recommendations	6	54.5
Error identification	6	54.5
Identification of contributing or latent factors	10	90.9
Management	5	45.5
Organization and procedures	7	63.6
Working environment	6	54.5
Teamwork	3	27.3
Technical processes	5	45.5
Professionals	4	36.4
Patients	2	18.2
Corrective actions		
Proposed actions	*N* = 26	
Staff Training	4	15.4
Writing procedures	12	46.1
Organizational changes	4	15.4
Increasing material resources	5	19.2
Decided actions	*N* = 21	
With a professional in charge	14	66.7
From the department	12	57.1
From another department	2	9.5
With a defined deadline	9	42.9

demonstrated that some patient safety incidents and contributing factors are specific to mental health [11]. Although the EFC examined events that might have been found in other medical units, events specific to mental health were also reported and analysed. For example, suicide attempts and missing person incidents were investigated and the analyses highlighted several organizational flaws in suicide risk assessment, monitoring patients at risk and securing the department's premises. Consequently, actions concerning the security of the building and the assessment of suicide risk in child psychiatry (specific staff training to manage patients with violent behaviour and improved guidelines) were set up.

Considering patient safety incidents in mental health as a result of a complex set of contributing factors, the EFC provided a structured framework to analyse them within the department's routine [8, 13, 22, 23]. The principle is to choose only one event per meeting to perform a thorough analysis using the ORION© method. As advocated by the Reason model, this method aims to identify factors related to the design of the system's organization or the workplace environment rather than individual error [24]. The method implies the main steps of the Association of Litigation and Risk Management (ALARM) protocol but seems easier to use for healthcare professionals who are not specialists in risk management [25]. The essential contribution of an EFC is to provide a formal framework to correct the latent failures in the department's organization.

Over the 3 years studied, the theoretical framework for conducting an EFC was not always followed. Indeed, the analysis of events did not always include all the steps defined in the ORION© method and the search for contributing factors was often superficial. These deviations can be partly explained by the staff's lack of time and availability. Carrying out the investigations to determine the causes of events as well as writing the report is time-consuming. Inconsistent monitoring of the corrective actions previously set up resulted in the resurgence of certain patient safety incidents that had previously been investigated, highlighting the importance of monitoring corrective actions by the EFC. A low level of expertise in conducting the analyses may also explain these deviations: only professionals involved in the EFC at its beginning were given formal training, and additional training was not proposed afterwards. Consequently, regular training seems necessary to ensure the quality of meetings and event analysis.

To function properly, the EFC requires reports of adverse events experienced by professionals. Several studies have shown that healthcare professionals, particularly physicians, agree with the importance of incident reporting and the concept of learning from errors [26, 27]. Nevertheless, in practice, many incidents are not reported [28, 29]. Self-report of patient safety incidents is hindered by several barriers such as time constraints, complex forms, fear of punishment, shame as well as lack of education and feedback [28, 30–33]. In the present study, we were not able to estimate the proportion of unreported incidents. However, the committee had enough incidents to discuss every month. Mental healthcare professionals are probably informed more easily of the corrective actions set up and can observe their effects considering that reported incidents are analysed by professionals in the department. Consequently, the existence of an EFC in a mental health department may improve incident reporting.

Psychiatric units provide global and coordinated care for patients through the involvement of many professionals. However, Priest and Borella [34] showed a higher risk of incidents due to the increasing distribution of patient

care over multiple practitioners. One of the strengths of the EFC is to gather all categories of professionals working in the unit and reinforcing interprofessional collaboration and promoting teamwork. This multidisciplinary approach also contributes to identifying system vulnerabilities more easily [35].

This study had several limitations. First, the functioning of an EFC depends on the professionals involved and the study was conducted in only one department. Second, the reporting system based on self-reporting by healthcare professionals did not provide the proportion of unreported incidents and did not take into account the patient's complaints that were treated by the hospital's legal department. Consequently, we were not able to measure the impact of the EFC on the completeness of incident reports and more generally on the prevalence of adverse events. Third, the impact of the EFC on patient safety was not assessed using clinical outcomes. However, we assumed that corrective actions against identified vulnerabilities resulted in an improvement of patient safety, including an improvement in the patient safety culture among professionals attending the committee.

Conclusions

The EFC is a tool allowing the direct involvement of mental health professionals to manage patient safety. This innovative management tool is adapted to the specific adverse events encountered in mental health. The theoretical framework for conducting an EFC was not always followed, suggesting the need for simplifying the method for professionals subject to tight time constraints.

Additional file

Additional file 1: Appendix. ORION standardized report

Author's contributions

BB: data management, analysis and interpretation; manuscript drafting; final approval of the version to be published; agreement to be accountable for all aspects of the work. TB: contribution to the design of the work; revising for important intellectual content; final approval of the version to be published; agreement to be accountable for all aspects of the work. OD: contribution to the design of the work; revising for important intellectual content; final approval of the version to be published; agreement to be accountable for all aspects of the work. AS: data analysis and interpretation; revising for important intellectual content; final approval of the version to be published; agreement to be accountable for all aspects of the work. PF: contribution to the design of the work; data analysis and interpretation; revising for important intellectual content; final approval of the version to be published; agreement to be accountable for all aspects of the work.

Author details

[1] Public Health Department, Grenoble University Hospital, Grenoble, France. [2] Laboratory TIMC, UMR 5525, CNRS, Joseph Fourier University, Grenoble, France. [3] Psychiatry Department, Grenoble University Hospital, Grenoble, France. [4] Neurology Department, Grenoble University Hospital, Grenoble, France.

Acknowledgements

We acknowledge Linda Northrup for her assistance in editing the manuscript. We acknowledge all professionals who participated in the EFC: C. Besançon, M. Bonnet, I. Bouillaud, C. Boutte, D. Chergui, F. Dejoie, O. Detante, M. Dubuc, C. Girard, S. Im, F. Imburchia, E. Lagrange, J.C. Peyrin, M. Poullet, C. Rivière, C. Robart, E. Ruhaud, E. Sellier, S. Sevenier, B. Teppa, C. Vidy, G. Weil.

Funding

This work was supported by the French Ministry of Health, project PREPS 2012.

Compliance with ethical guidelines

Competing interests

The authors declare that they have no competing interests

Received: 27 January 2015 Accepted: 20 August 2015
Published online: 03 September 2015

References

1. Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG et al (2004) Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical Practice Study I. Qual Health Care 13(2):145–151. doi:10.1136/qshc.2002.003822
2. Pietro DA, Shyavitz LJ, Smith RA, Auerbach BS (2000) Detecting and reporting medical errors: why the dilemma? BMJ 320(7237):794–796
3. Reason J (2000) Human error: models and management. BMJ 320(7237):768–770
4. Vincent C, Neale G, Woloshynowych M (2001) Adverse events in British hospitals: preliminary retrospective record review. BMJ 322(7285):517–519
5. Weingart SN, Wilson RM, Gibberd RW, Harrison B (2000) Epidemiology of medical error. BMJ 320(7237):774–777
6. Baker GR, Norton PG, Flintoft V, Blais R, Brown A, Cox J et al (2004) The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. CMAJ Can Med Assoc J 170(11):1678–1686
7. Bowers L, Simpson A, Alexander J (2003) Patient-staff conflict: results of a survey on acute psychiatric wards. Soc Psychiatry Psychiatr Epidemiol 38(7):402–408. doi:10.1007/s00127-003-0648-x
8. Desai RA, Dausey DJ, Rosenheck RA (2005) Mental health service delivery and suicide risk: the role of individual patient and facility factors. Am J Psychiatry 162(2):311–318. doi:10.1176/appi.ajp.162.2.311
9. Druss BG (2007) Improving medical care for persons with serious mental illness: challenges and solutions. J Clin Psychiatry 68(Suppl 4):40–44
10. Laitanantpong D (2006) The patient risk in psychiatric service at King Chulalongkorn Memorial Hospital, Thai Red Cross Society. J Med Assoc Thail Chotmaihet thangphaet 89(Suppl 3):S174–S179
11. Nath SB, Marcus SC (2006) Medical errors in psychiatry. Harv Rev Psychiatry 14(4):204–211. doi:10.1080/10673220600889272
12. Hairon N (2008) Guidelines focus on improving patient safety in mental health. Nurs Times 104(48):19–20
13. Brickell TA, McLean C (2011) Emerging issues and challenges for improving patient safety in mental health: a qualitative analysis of expert perspectives. J Patient Saf 7(1):39–44. doi:10.1097/PTS.0b013e31820cd78e
14. Woynar S, Debouck F, Cellier P, Bourhis J, Cauterman M, Lartigau E (2007) Towards global security in radiation oncology. Using event notification feed-back. Cancer Radiother J Soc Fr De Radiother Oncol 11(6–7):320–328. doi:10.1016/j.canrad.2007.09.003
15. Lartigau E, Vitoux A, Debouck F (2009) Crex and Orion analysis in radiation oncology: towards a mutualisation of corrective actions. Cancer Radiother J Soc Fr Radiother Oncol 13(6–7):458–460. doi:10.1016/j.canrad.2009.08.143
16. Debouck F, Petit HB, Lartigau E (2010) From Crex mutualisation to clinical audit. Cancer Radiother J Soc Fr Radiother Oncol 14(6–7):571–575. doi:10.1016/j.canrad.2010.07.018
17. Debouck F, Rieger E, Petit H, Noel G, Ravinet L (2012) ORION(R): a simple and effective method for systemic analysis of clinical events and precursors occurring in hospital practice. Cancer Radiother J Soc Fr Radiother Oncol 16(3):201–208. doi:10.1016/j.canrad.2011.12.002

18. Francois P, Sellier E, Imburchia F, Mallaret MR (2013) Experience feedback committee: a method for patient safety improvement. Rev Epidemiol Sante Publique 61(2):155–161. doi:10.1016/j.respe.2012.09.001
19. Lecoanet A, Sellier E, Carpentier F, Maignan M, Seigneurin A, Francois P (2013) Experience feedback committee in emergency medicine: a tool for security management. Emer Med J EMJ. doi:10.1136/emermed-2013-202767
20. Lartigau E, Coche-Dequeant B, Dumortier V, Giscard S, Lacornerie T, Lasue A et al (2008) Experience feed back committee in radiotherapy (CREx): a compulsory tool for security management evaluation at two years. Cancer Radiother J Soc Fr Radiother Oncol 12(6–7):610–613. doi:10.1016/j.canrad.2008.07.007
21. World Alliance For Patient Safety Drafting Group, Sherman H, Castro G, Fletcher M, World Alliance for Patient Study, Hatlie M et al (2009) Towards an International Classification for Patient Safety: the conceptual framework. Int J Qual Health Care J Int Soc Qual Health Care/ISQua. 21(1):2–8. doi:10.1093/intqhc/mzn054
22. Grasso BC, Bates DW (2003) Medication errors in psychiatry: are patients being harmed? Psychiatric Serv 54(5):599
23. Measham TJ (1995) The acute management of aggressive behaviour in hospitalized children and adolescents. Can J Psychiatry Rev Can Psychiatr 40(6):330–336
24. Reason JT, Carthey J, de Leval MR (2001) Diagnosing "vulnerable system syndrome": an essential prerequisite to effective risk management. Qual Health Care QHC 10(Suppl 2):ii21–ii25
25. Vincent C, Taylor-Adams S, Chapman EJ, Hewett D, Prior S, Strange P et al (2000) How to investigate and analyse clinical incidents: clinical risk unit and association of litigation and risk management protocol. BMJ 320(7237):777–781
26. Firth-Cozens J (2003) Evaluating the culture of safety. Qual Saf Health Care 12(6):401
27. Sarvadikar A, Prescott G, Williams D (2010) Attitudes to reporting medication error among differing healthcare professionals. Eur J Clin Pharmacol 66(8):843–853. doi:10.1007/s00228-010-0838-x
28. Blendon RJ, DesRoches CM, Brodie M, Benson JM, Rosen AB, Schneider E et al (2002) Views of practicing physicians and the public on medical errors. New Engl J Med 347(24):1933–1940. doi:10.1056/NEJMsa022151
29. Haller G, Courvoisier DS, Anderson H, Myles PS (2011) Clinical factors associated with the non-utilization of an anaesthesia incident reporting system. Br J Anaesth 107(2):171–179. doi:10.1093/bja/aer148
30. Hobgood C, Weiner B, Tamayo-Sarver JH (2006) Medical error identification, disclosure, and reporting: do emergency medicine provider groups differ? Acad Emer Med Off J Soc Acad Emer Med 13(4):443–451. doi:10.1197/j.aem.2005.11.005
31. Leape LL, Berwick DM (2005) Five years after to err is human: what have we learned? JAMA 293(19):2384–2390
32. Leape LL (2002) Reporting of adverse events. New Engl J Med 347(20):1633–1638. doi:10.1056/NEJMNEJMhpr011493
33. Vincent C, Stanhope N, Crowley-Murphy M (1999) Reasons for not reporting adverse incidents: an empirical study. J Eval Clin Pract 5(1):13–21
34. Priest SL, Borella AB (1998) A crash course in systems analysis. J Healthcare Info Manag JHIM 12(3):17–23
35. Schmutz J, Manser T (2013) Do team processes really have an effect on clinical performance? A systematic literature review. Br J Anaesth 110(4):529–544. doi:10.1093/bja/aes513

2.3.2. Fonctionnement de l'ensemble des CREX de l'établissement : article soumis - Experience feedback committees: implementation and functioning in hospital medical departments.

Dans cet article soumis le 22 décembre 2017, nous avons étudié le fonctionnement des 20 CREX du CHU de Grenoble.

Objectif : L'objectif de ce travail était d'analyser le fonctionnement global des CREX au sein d'un établissement de santé, et de rechercher les facteurs associés à leur productivité, définie à partir du nombre d'actions correctrices décidées par les comités.

Méthode : Nous avons analysé l'ensemble des documents produits par les CREX du CHU de Grenoble sur une durée de 12 mois. Les documents étaient analysés de façon indépendante par deux investigateurs (PF et BB), et classées en reprenant la méthodologie choisie lors de l'analyse du CREX de Neuropsychiatrie.

Résultats : Sur les 12 mois d'analyse, nous avons identifié 164 réunions, 1707 signalements d'événements indésirables, 91 analyses Orion, menant à la décision de 206 actions correctrices. La médiane de la série d'actions correctrices décidées par les comités était de 5 (étendue 0-62). Les CREX les plus productifs, définis comme ayant décidé au moins 5 actions correctrices par an, examinaient significativement plus d'événements indésirables (109 versus 45, $p<0,001$) et conduisaient significativement plus d'analyses des causes racines (7 rapports Orion versus 1, $p<0,001$). Les rapports Orion produits par ces comités identifiaient plus de causes racines, et les actions correctrices décidées étaient mieux planifiées (pilote désigné, calendrier de mise en place effective).

Conclusion : Cette étude a donc permis d'avoir une vision d'ensemble du fonctionnement des CREX de l'établissement, mis en places dans la diversité des spécialités (médicales, chirurgicales et médicotechniques). Globalement, nous avons mis en évidence une bonne adhésion des professionnels de santé à la méthode CREX, un nombre important d'EIAS

analysés et d'actions correctrices décidées. Nous avons toutefois noté le rôle important des leaders médicaux des CREX, ainsi que les difficultés de certains CREX pour mener des analyses des causes racines des EIAS. En effet, les rapports Orion ne suivaient pas toujours les étapes nécessaires à l'identification des causes racines, aboutissant à une recherche souvent superficielle des facteurs favorisants les EIAS. Nous avons également relevé une relative inefficacité des CREX pour analyser et corriger les événements indésirables dépassant leurs périmètres (exemple : EIAS au cours d'un transfert de patients ou défaut de transmissions entre services). Ces conclusions sont en cohérence avec la littérature, qui émet des critiques quant à l'efficacité des méthodes d'analyse des causes racines conduites par les professionnels de santé (Peerally, Carr et al. 2017). Nous discuterons de ces limites dans le chapitre IV de la thèse.

Experience feedback committees: implementation and functioning in hospital medical departments.

Patrice François [1, 2]MD, PhD, André Lecoanet [1]MD, Alban Caporossi [1]MD, Anne-Marie Dols [1]MD, Arnaud Seigneurin [1,2]MD, PhD, Bastien Boussat[1,2]MD.

ABSTRACT

Background: The experience feedback committee (EFC) is a tool designed to involve medical teams in patient safety management, through root cause analysis (RCA) within the team.

Objectives: To investigate the functioning of EFCs in the departments of a single university hospital in France and to consider their contribution to the management of patient safety.

Methods: Cross-sectional, observational study, based on an analysis of the documents produced by the EFCs for 1 year. Data were collected independently by two investigators in meeting minutes, adverse event reports, and event analysis reports.

Results: The study included all 20 EFCs operating in the hospital's medical departments. During the study year, committees held 164 meetings, reviewed 1,707 adverse events, conducted 91 event analyses and decided on 206 corrective actions. The median number of corrective actions adopted by each EFC was five actions (range, 0–62). A root cause analysis (RCA) was present in 76% of the analysis reports, but these analyses were often incomplete. There was also a lack of planning corrective actions: an implementation deadline was only defined in 26% of the actions.

Conclusions: Healthcare professionals adhered to the system-based approach to patient safety, but we observed difficulties in holding regular meetings and deviations from the theoretical framework. These findings confirm the difficulties of practicing RCA in the healthcare setting. Nevertheless, EFCs can be vectors of safety culture and teamwork.

Keywords: Patient Safety; Root Cause Analysis; Risk Management; Safety Culture, Hospital Medicine

INTRODUCTION

Despite the efforts made in all healthcare systems for nearly 20 years to improve safety, progress remains slow and the results are modest.[1] The main obstacles to the development of effective safety management systems relate to the sociology of healthcare organisations and the culture of healthcare professionals.[2] Thus, one of the main challenges for the promotion of patient safety is to directly involve health professionals, especially doctors and clinical teams, in healthcare safety management.[3] Particular attention is paid to the ability of healthcare professionals to work as a team, namely their ability to communicate and collaborate. A wide variety of interventions has been proposed to improve teamwork and adopt patient safety-oriented operating modes. These interventions include programs inspired by the "crew resource management" developed in civil aviation or the programs labeled “TeamSTEPPS”.[4-6]

In France, a specific program, called the Experience Feedback Committee (EFC), has been created to involve the staff of a medical department in patient safety management.[7] Originating from civil aviation security systems, the program is based on root cause analysis (RCA) and prevention of adverse events. It has been adapted to healthcare facilities and was successfully implemented in radiotherapy units by the company Air France Consulting in 2005.[7, 8] Thereafter healthcare authorities promoted the method, which was widely implemented in French hospitals in various departments.[9-11] An EFC is a team including a variety of professionals who represent the diversity of the functions encountered in the medical department. The EFC members usually meet monthly to examine reported incidents that occurred in their department. They choose priority incidents that need to be analysed and propose corrective actions. The main principles of this method are as follows: patient safety must be managed within a medical team, the team must focus on near-miss events and the corrective actions must involve latent factors that contributed to the occurrence of the near-miss event.[8, 12]

In 2007 the first EFCs were implemented by its designer in four departments in our hospital. Based on this experience, the hospital managers wished to extend the method to other volunteer departments.

The objective of this study was to investigate the functioning of the EFC in the departments of a university hospital and to consider its contribution to the management of patient safety.

METHODS

Study design

We conducted a cross-sectional, observational study of EFCs established in the Grenoble University Hospital (France), based on the analysis of

documents produced by each EFC over 1 year of activity. The study was approved by the Institutional research ethics committee of south-eastern France (IRB 6705).

Setting

The study was conducted in a 1347-bed acute-care university hospital including 42 clinical and medical-technical departments.

Adverse and near-miss events are reported to a central safety unit using a voluntary internal reporting system based on a standardized reporting form. This unit is made up of one medical doctor, one pharmacist and one quality engineer. The events reported were classified by severity and risk area. The central safety unit directly investigates the most serious events and those involving several hospital departments. Other events are transmitted to the appropriate operator and to the executives of the relevant departments. For departments where an EFC is implemented, the central safety unit addresses the reports of events to the EFC leader every month.

Experience Feedback Committee Framework

The functioning of the EFC is defined in local guidelines in accordance with the framework proposed by Air France Consulting.[8] The departments can obtain methodological assistance from the quality-assurance team.

The Committee is composed of volunteer representatives of the department's various professions. A few days before the committee meeting, the EFC leader receives a file including the department's events reports from the central safety unit. The Committee can also set up a specific reporting system for the EFC.

The committee meets regularly, usually once a month, according to a fixed schedule. Meetings last between 1 and 2 h. Committee meetings are conducted according to a standardised framework: 1) reading the list of reported events, 2) choosing a priority event to investigate, 3) choosing the professional responsible for the investigation, 4) reviewing the analysis carried out for the event chosen the previous month, 5) deciding on corrective actions and 6) monitoring on-going actions.

The investigation is carried out during the month following the EFC meeting by a designated person using the Orion method developed by aviation safety experts.[8] The main steps of the method are as follows: collecting data and existing guidelines describing the chronological facts that occurred before, during and after the event; describing the failures; looking for causes of errors and latent factors that could have contributed to the failures; setting up corrective actions; and writing a report of the analysis. Causes and latent factors are found in different areas such as political, organisational, working conditions, team functioning, procedures, actors and the patient.

Study sample

All EFCs established in the hospital departments more than 1 year before were eligible. The purpose of the study was presented to the EFC leaders and their consent was required for participation in the study.

Data collection

All written documents produced by the EFC during a 1-year period before the inclusion were analysed. These documents included meeting minutes, event reports, event analysis reports, and all documents related to corrective actions decided by the EFC. Reported events were classified according to the source of the report, the type of event and the consequence for the patient, using the International Classification for Patient Safety.[13] Written reports from meetings were analysed using a standardised form that included the theoretical steps of an EFC meeting. The event analysis reports were analysed using a standardised form in accordance with the Orion method. The corrective actions were classified by type and planning elements (i.e., designating a person in charge and setting a deadline for implementation). No direct or indirect identification of patients or healthcare professionals was possible in the data collected. The data were collected independently by two investigators (PF & BB). Differences in recording were discussed until a consensus was reached.

Statistical analysis

EFC baseline characteristics were reported as numbers and percentages for categorical variables, and median and interquartile range (IQR, 25^{th} and 75^{th} percentiles) for continuous variables. We generated another variable to identify the productive EFCs (the median number of corrective actions per year was chosen to dichotomize this variable) Secondarily, we compared the EFC characteristics across subgroups of EFCs defined by this computed variable using the chi-square test or Fischer's exact test, when appropriate, for categorical variables, and the Kruskal-Wallis test for continuous variables. *P*-values less than 0.05 were considered statistically significant. Analyses were performed using Stata 14.0 (Stata Corp, College Station, TX, USA).

RESULTS

From 2007 to 2014, EFCs were implemented in 20 departments, including seven medical departments (gastroenterology, infectious diseases, cardiology, pediatrics, internal medicine, vascular medicine, neurology), six medical-technical departments (nuclear medicine, pharmacy, sterilisation, biology, radiotherapy), five emergency or intensive care

departments and two surgical departments. Five EFCs had periods from 6 to 26 months of inactivity. During the 1-year period studied, the EFCs held 164 meetings (2–12 per EFC). They examined 1,707 reported events, conducted in-depth analysis of 91 events and decided on 206 corrective actions.

EFC meetings and participants.

Meeting minutes were found for 160 (98%) of the 164 committee meetings. These reports mentioned 351 participants including 99 physicians, 48 head nurses and 76 nurses or other paramedics (Table 1). The review of the month's event reports was noted in almost all meeting minutes. The presentation of an event analysis report, a list of decided actions and the follow-up of previous actions were present in 58% of all meeting minutes.

Table 1: Information included in the meeting minutes of the 20 EFCs (*n*=**160** reports)

		All meetings		Per EFC	
Presence of information		*n*	%	Median	[IQR]
–	Review of event reports	158	99	8	[7; 10]
–	Choice of an event to investigate	98	61	5	[1; 8]
–	Presentation of analysis report	93	58	5	[2; 7]
–	List of actions decided	92	58	4	[2; 7]
–	Follow-up of previous actions	93	58	5	[2; 8]
Meeting attendees		351	100	15.5	[14; 21]
–	Physicians	99	28	4	[3; 6]
–	Head nurses	48	14	2	[1; 4]
–	Nurses and other paramedics	76	22	3	[1; 6]
–	Students	51	15	1	[0;5]
–	Others (secretary, technicians, etc.)	77	21	2	[0;6]

IQR = Interquartile range

Events reported

Of the 20 EFCs, 11 used event reporting from the central unit for risk management, six used only the reports collected in the EFC department and three used the two sources of reports (Table 2). Most events occurred in the EFC department (83%) and were reported by the professionals of the department (86%). These events related mainly to medication issues (21%), organisation of care (20%) or medical devices (17%). Most of the reported events (91%) did not have harmful consequences for the patients. However, three patients died and 11 underwent severe harm.

Table 2. Characteristics of event reports reviewed by EFCs (n = **1707**)

	All EFCs		Per EFC	
	n	%	Median	[IQR]
Reporting route				
Central unit of risk management	1195	67	34	[0; 84]
EFC department	585	33	0	[0; 41]
Individual who reports				
EFC department staff	1475	86	50	[31; 106]
Other staff	232	14	3	[0; 11]
Location of occurrence				
EFC department	1422	83	49	[31; 103]
Other department	285	17	13	[2; 18]
Topics of event				
Medications	365	21	3	[0; 15]
Care organisation	337	20	10	[5; 25]
Medical device and equipment	298	17	8	[1; 10]
Care process and practices	253	15	3	[1; 7]
Patients and relatives	126	7	1	[0; 4]
Environment: premises, hygiene	125	7	3	[0; 6]
Staff	96	5	1	[0; 1]
Patient records	43	3	1	[0; 2]
Other†	64	5	1	[0; 2]
Severity of reported events				
Event without harm	1555	91	44	[24; 99]
Minor harm	96	6	3	[1; 7]
Moderated harm	42	2	1	[0; 2]
Severe harm	11	1	0	[0; 0]
Death	3	0	0	[0; 0]

Abbreviation: IQR, interquartile range (i.e., 25th and 75th percentiles)

†Transfusion, nosocomial infections, food, medical gases, bed availability, etc.

Root cause analysis of events

Among all the events reviewed, 98 events were selected for in-depth analysis. An analysis report presented at a committee meeting was produced for 91 event investigations, and 72 of them (79%) were presented according to the format of the Orion method (Table 3). The data collection procedures were specified by only 38% of the reports. The chronology of the facts was described in 88% of cases, and the search for the causes was present in 84%. However, this cause analysis was often incomplete, not exploring all categories of causes. Ninety-five per cent of the reports included at least one proposal for corrective action.

Table 3. Characteristics of analysis reports reported to committee (*n*=**91**)

	All EFCs		Per EFC	
	n	%	Median	[IQR]
Presentation of the analysis report	91	100	5	[1; 7]
Presentation format				
Orion format	72	79	5	[3; 6]
Oral	10	11	1	[1; 2]
Oral with visual support	8	9	4	[1; 7]
Methods of collecting data	35	38	1	[0; 4]
Individual interviews	32	35	4	[3; 4]
Debriefing	8	9	0	[0; 1]
Patient record	14	15	2	[0; 2]
Site visit	22	24	1	[1; 3]
Search for documents	13	14	1	[0; 1]
Chronology of the facts	80	88	5	[3; 7]
Description of the chronology	75	82	5	[2; 6]
Identifying errors	52	57	3	[1; 5]
Investigation of causes and contributing factors	76	84	6	[3; 6]
Organisation	64	70	5	[3; 6]
Working conditions	51	56	4	[2; 4]
Team functioning	50	55	3	[2; 5]
Policy	49	54	2	[1; 6]
Staff	49	54	3	[2; 5]
Guidelines, procedures	48	53	2	[1; 6]
Patients	30	33	2	[1; 3]
Other	9	10	0	[0; 1]
Proposals for corrective actions	86	95	6	[4; 7]

Abbreviation: IQR, interquartile range (i.e., 25th and 75th percentiles)

Corrective actions

The annual number of corrective actions decided by each EFC ranged from 0 to 62, with a median of 5 (Table 4). The most frequent actions were to change an organisational point (34%), write or modify a procedure (30%) or organise staff training (22%). The majority of actions had a designated person responsible (77%), who generally worked in the same department (70%). Only 26% of all corrective actions included a deadline for implementation.

Table 4. Characteristics of the actions adopted by the Committee (*n*=**206**)

	All EFCs		Per EFC	
	n	%	Median	[IQR]
Actions adopted by the EFC	**206**	100	5	[0; 13]
Action type				
Organisation improvement	70	34	1	[0; 5]
Write or revise a procedure	61	30	1.5	[0; 3]
Train staff	45	22	1	[0; 2.5]
Improve a device	16	8	0	[0; 0.5]
Other	13	6	0	[0; 0]
Person in charge of the action:				
Member of department	145	70	1	[0; 8]
Other	15	7	0	[0; 1]
Undesignated	46	22	0.5	[0; 3]
Defined deadline	53	26	0	[0; 3.5]

Abbreviation: IQR, interquartile range (i.e., 25th and 75th percentiles)

Factors related to the EFC productiveness

The most productive committees, defined as EFCs that have decided at least 5 actions in the year, reviewed more events (109 versus 45, p=0.02) and achieved more investigations (7 analysis reports versus 1, p<0.001) (Table 5).

The analysis reports more often provided root causes of the event (p<0.001). The most active committees more often designated a person in charge of the corrective action (p<0.001) and their actions more often had a deadline for implementation (p=0.01).

Table 5: Comparison of characteristics of EFCs according to the number of actions decided in the year (less than 5 versus greater than or equal to 5).

	< 5 actions	≥ 5 actions	p
Department specialty; *n* (%)			0.07
Clinical department	9 (90)	5 (50)	
Medical-technical department	1 (10)	5 (50)	
EFC seniority, median [IQR], y	2 [1; 2]	1 [1; 2]	0.08
Number of attendees, median [IQR]	15 [9; 19]	20 [14; 22]	0.29
Number of events reported, median [IQR]	45 [33; 60]	109 [76; 142]	0.02
Number of analysis reports, median [IQR]	1 [0; 4]	7 [6; 8]	<0.001
Mode of presentation; *n* (%)			
Orion format	5 (50)	9 (90)	0.07
Other format	5 (50)	1 (10)	
Search for causes, median [IQR]	0 [0; 2]	6 [5; 7]	<0.001
Designated person in charge of action, median [IQR]	0 [0; 1]	8 [5; 18]	<0.001
Defined deadline, median [IQR]	0 [0; 0]	4 [0; 6]	0.01
Follow-up of previous actions, median [IQR]	3 [0; 7]	6 [2; 8]	0.14

Abbreviation: IQR, interquartile range (i.e., 25th and 75th percentiles); y, year

DISCUSSION

This study shows that nearly half of all medical departments voluntarily implemented an Experience Feedback Committee (EFC). Healthcare professionals adhere to the method that is implanted in a wide variety of medical departments. Reported adverse events are analysed and corrective actions are decided by the committees.

However, this picture is mitigated by the problems maintaining this activity in the routine of the healthcare teams. The number of meetings varied over time and from one department to another; some even had long periods of inactivity. Healthcare professionals explained these variations in activity by the departure of a leader who was not replaced and, above all, by the lack of time and resources. Indeed, carrying out investigations to identify the causes of events, as well as writing analysis reports, takes a lot of time for professionals who are already very busy.[14, 15]

The study also shows the sometimes significant deviations in practices compared to the theoretical functioning of the EFC. The EFC is based on a systems approach to patient safety and it provides a formal method for the root cause analysis (RCA) of adverse events. The Orion method, based on the Reason model, is close to the ALARM method and includes the same steps.[16] Initiated in civil aviation, the Orion method was adapted to the field of healthcare by aviation safety experts. It is simpler than the ALARM method and, a priori, easier to use by healthcare professionals not specialised in risk management. However, the analysis of events did not always follow all the steps defined by the Orion method and the search for contributing factors was often superficial. There was also a lack of planning the action selected by the committee and a failure to follow up the corrective actions decided previously.

This weakness found in the practice of RCA in the field of healthcare is reported by many authors.[14, 17-19] Earlier studies of RCA reports had shown that this analysis often lacks depth and rigor, or that the method is rarely adequately applied.[17, 19] Overly simple or poorly designed action plans are insufficient to prevent the recurrence of incidents, and may even generate new risks.[17] In addition, action plans are often not followed up, and when this monitoring exists, only part of the actions decided are effectively implemented.[14, 17]

This lack of rigor in the application of the RCA may be partly explained by the healthcare professionals' lack of training.

When interviewing healthcare professionals who participated in RCA training programs, paradoxical responses were obtained.[14, 15, 19] On the one hand, these people express a very positive opinion

on the method, which contributes to improving the safety of care and induces cultural changes.[14] On the other hand, the same people express difficulty in using RCA in practice. In addition to the lack of time and resources, healthcare professionals brought up difficulties related to interprofessional relations.[14, 15, 19] Indeed, the RCA interacts with a complex sociocultural context in which the investigation of a care-associated adverse event can be misunderstood.[17, 20] To preserve good interprofessional relations and avoid hierarchical tensions, the investigators remain on the surface of the issue and conceal certain profound sociopolitical and organisational problems.[17, 18, 20] Respect for the non-blame dogma, consubstantial with the systems approach, can lead to underestimating the human causes of an incident such as the transgression of rules or negligence.[17]

RCA effectively improves safety in various industries such as civil aviation. Because of this potential, RCA has become an important part of all healthcare safety management programs around the world.[15, 17] There is, however, no scientific evidence that RCA improves the safety of care.[15, 21] Studies that measured the impact of incident reporting systems or RCA use did not show any effect or only anecdotal effects.[21, 22] For example, Percarpio et al. studied 139 Veterans Affairs Medical Centers and found an association between RCA practice and the level of some safety indicators for postoperative complications,[23] but the study design could not assert that the relationship was causal. RCA's limitations are also illustrated by examples of incidents that occurred after an identical incident was analysed and an action plan decided within the same department.[17]

In agreement with the literature, we found that the practice of RCA in the EFC is imperfect and that it would be illusory to expect short-term effects on the incidence of adverse events and patient safety. However, we hope that the EFCs will have long-term effects by increasing safety culture and learning by error. Indeed, the EFC is a particular mode of implementation of RCA in the field of healthcare that aims to directly involve members of the staff in the management of adverse events affecting their department. A person who reports an incident is invited to participate in the analysis and solution development and can then see the implementation of corrective actions and observe their effects. In this context, the professionals receive feedback on the reports, within a short feedback loop. This can help reduce a traditional barrier of incident reporting that is related to the opacity of reporting systems and the lack of feedback to the reporters about actions decided following such reports.[24] We support the hypothesis that the direct involvement of healthcare professionals in a learning-by-error system can be a strategy for the acquisition of values and behaviours that make up the safety culture.[24] In another study, we analyzed the association between patient safety culture, as measured by the Hospital Survey on Patient Safety Culture (HSOPS), and the care provider involvement in EFC activities.[25] We showed that EFC participants had a more developed patient safety culture, with nine of the twelve HSOPS dimension scores significantly higher than EFC non-participants.

We also hypothesise that the system-based approach to manage adverse events within a team is likely to improve the perception of the collective dimension of healthcare and thus foster teamwork. The social and cultural functions of the EFCs can be compared to those of the mortality and morbidity conferences when they are used to improve the quality and safety of care.[26, 27] Mortality and morbidity conferences oriented towards patient safety are more often multidisciplinary and use methods of systemic analysis of adverse events based on lighter versions of RCA.[26, 28, 29] Like the EFCs, the mortality and morbidity conferences can lead to the implementation of improvement action plans.[28, 29] Their contribution to fostering teamwork and enhancing the safety culture of healthcare professionals has been acknowledged.[26, 27, 30, 31]

The main limitation of this study is that it concerns only one hospital. It is probably not representative of the functioning of all EFCs in all French hospitals. However, this study is exploratory, it is the first one that analyses the functioning of several EFCs implanted in different medical specialties. It shows the difficulties of running a risk management system on a regular basis and opens up leads for carrying out broader studies on several hospitals.

We can also discuss the relevance of the criterion based on the number of actions decided to estimate the effectiveness of EFCs. This is an intermediate criterion, indicative of EFC functioning. To evaluate EFC's ability to improve care safety, it would be necessary to verify the actual implementation of these actions and their effect on adverse event incidence.

Another limitation is related to the retrospective collection of data, including missing data due to absent or incomplete reports. However, this lack of traceability is itself a result confirming how difficult it is for teams to follow a rigorous method.

In spite of these limitations, this study provides leads to improve the functioning of EFCs. Although the method seems simple, its implementation requires training and it will be necessary to strengthen the training of healthcare professionals and to offer long-term methodological support.

Conclusion

The EFC is a way to involve healthcare professionals in system-based analysis of adverse events associated with medical care. The study identifies the limitations of this type of activity, which requires time for professionals and skills that are not easy to acquire. But the main barriers to implementing RCA in healthcare teams are psychological, social and cultural. However, we observe the approval of professionals persisting over time, and we remain hopeful that RCA will contribute to improving the safety culture of healthcare professionals.

Acknowledgements:

We acknowledge Linda Northrup for her assistance in editing the manuscript.

This work was supported by the French Ministry of Health [PREPS 12-026-0001 2012]

REFERENCES

1. Baines R, Langelaan M, de Bruijne M, et al. How effective are patient safety initiatives? A retrospective patient record review study of changes to patient safety over time. *BMJ Qual Saf* 2015;24:561-71.
2. Weaver SJ, Lubomksi LH, Wilson RF, et al. Promoting a culture of safety as a patient safety strategy: a systematic review. *Ann Intern Med* 2013;158:369-74.
3. Zoutman DE, Ford BD. Quality improvement in hospitals: barriers and facilitators. *Int J Health Care Qual Assur* 2017;30:16-24.
4. Kemper PF, de Bruijne M, van Dyck C, et al. Crew resource management training in the intensive care unit. A multisite controlled before-after study. *BMJ Qual Saf* 2016;25:577-87.
5. Duclos A, Peix JL, Piriou V, et al. Cluster randomized trial to evaluate the impact of team training on surgical outcomes. *Br J Surg* 2016;103:1804-1814.
6. Weld LR, Stringer MT, Ebertowski JS, et al. TeamSTEPPS Improves Operating Room Efficiency and Patient Safety. *Am J Med Qual* 2016;31:408-14.
7. Lartigau E, Vitoux A, Debouck F. Crex and Orion analysis in radiation oncology: towards a mutualisation of corrective actions. *Cancer Radiother* 2009;13:458-60.
8. Debouck F, Rieger E, Petit H, et al. ORION((R)): A simple and effective method for systemic analysis of clinical events and precursors occurring in hospital practice. *Cancer Radiother* 2012;16:201-8. .
9. Lecoanet A, Sellier E, Carpentier F, et al. Experience feedback committee in emergency medicine: a tool for security management. *Emerg Med J* 2014;31:894-8.
10. Boussat B, Bougerol T, Detante O, et al. Experience Feedback Committee: a management tool to improve patient safety in mental health. *Ann Gen Psychiatry* 2015;14:23.
11. Caporossi A, Brudieu E, Lehmann A, et al. An Experience Feedback Committee for Improving Medication Process Safety: An Observational Study in a Hospital Pharmacy Department. *J Pharma Care Health Sys* 2014;S1-010.
12. Francois P, Sellier E, Imburchia F, et al. Experience feedback committee: a method for patient safety improvement. *Rev Epidemiol Sante Publique* 2013;61:155-61.
13. Sherman H, Castro G, Fletcher M, et al. Towards an International Classification for Patient Safety: the conceptual framework. *Int J Qual Health Care* 2009;21:2-8.
14. Braithwaite J, Westbrook MT, Mallock NA, et al. Experiences of health professionals who conducted root cause analyses after undergoing a safety improvement programme. *Qual Saf Health Care* 2006;15:393-9.
15. Bowie P, Skinner J, de Wet C. Training health care professionals in root cause analysis: a cross-sectional study of post-training experiences, benefits and attitudes. *BMC Health Serv Res* 2013;13:50.
16. Vincent C, Taylor-Adams S, Chapman EJ, et al. How to investigate and analyse clinical incidents: clinical risk unit and association of litigation and risk management protocol. *BMJ* 2000;320:777-81.
17. Peerally MF, Carr S, Waring J, et al. The problem with root cause analysis. *BMJ Qual Saf* 2017;26:417-422.
18. Wu AW, Lipshutz AK, Pronovost PJ. Effectiveness and efficiency of root cause analysis in medicine. *JAMA* 2008;299:685-7.
19. Wallace LM, Spurgeon P, Adams S, et al. Survey evaluation of the National Patient Safety Agency's Root Cause Analysis training programme in England and Wales: knowledge, beliefs and reported practices. *Qual Saf Health Care* 2009;18:288-91.
20. Nicolini D, Waring J, Mengis J. Policy and practice in the use of root cause analysis to investigate clinical adverse events: mind the gap. *Soc Sci Med* 2011;73:217-25.
21. Percarpio KB, Watts BV, Weeks WB. The effectiveness of root cause analysis: what does the literature tell us? *Jt Comm J Qual Patient Saf* 2008;34:391-8.
22. Stavropoulou C, Doherty C, Tosey P. How Effective Are Incident-Reporting Systems for Improving Patient Safety? A Systematic Literature Review. *Milbank Q* 2015;93:826-66.
23. Percarpio KB, Watts BV. A cross-sectional study on the relationship between utilization of root cause analysis and patient safety at 139 Department of Veterans Affairs medical centers. *Jt Comm J Qual Patient Saf* 2013;39:32-7.
24. Anderson JE, Kodate N, Walters R, et al. Can incident reporting improve safety? Healthcare practitioners' views of the effectiveness of incident reporting. *Int J Qual Health Care* 2013;25:141-50.
25. Boussat B, Seigneurin A, Giai J, et al. Involvement in root cause analysis and patient safety culture among hospital care providers. *J Patient Saf*;in press.
26. Szekendi MK, Barnard C, Creamer J, et al. Using patient safety morbidity and mortality conferences to promote transparency and a culture of safety. *Jt Comm J Qual Patient Saf* 2010;36:3-9.
27. Lecoanet A, Vidal-Trecan G, Prate F, et al. Assessment of the contribution of morbidity and mortality conferences to quality and safety improvement: a survey of participants' perceptions. *BMC Health Serv Res* 2016;16:176.
28. Tignanelli CJ, Embree GGR, Barzin A. House staff-led interdisciplinary morbidity and mortality conference promotes systematic improvement. *J Surg Res* 2017;214:124-130.
29. Francois P, Prate F, Vidal-Trecan G, et al. Characteristics of morbidity and mortality conferences associated with the implementation of patient safety improvement initiatives, an observational study. *BMC Health Serv Res* 2016;16:35.
30. Szostek JH, Wieland ML, Loertscher LL, et al. A systems approach to morbidity and mortality conference. *Am J Med* 2010;123:663-8.

31. Pelieu I, Djadi-Prat J, Consoli SM, et al. Impact of organizational culture on preventability assessment of selected adverse events in the ICU: evaluation of morbidity and mortality conferences. *Intensive Care Med* 2013;39:1214-20.

3. CHAPITRE II: Mesure de la culture sécurité

3.1. Introduction

Afin d'évaluer l'association entre la participation à un CREX et la culture de sécurité des professionnels de santé, il était nécessaire de choisir un outil de mesure valide. La culture de sécurité, telle que définie par la société européenne pour la qualité des soins (European Society for Quality in Health Care), désigne « un ensemble cohérent et intégré de comportements individuels et organisationnels, fondés sur des croyances et des valeurs partagées, qui cherche continuellement à réduire les dommages aux patients, lesquels peuvent être liés aux soins » (van Everdingen JJ 2006). Il s'agit donc d'un ensemble complexe dont la question de la mesure a fait l'objet de nombreux travaux (Scott, Mannion et al. 2003, Pronovost and Sexton 2005, Occelli 2010). Plusieurs outils d'évaluation ont été développés, variant dans la conceptualisation de la culture de sécurité et dans les dimensions de celle-ci, mais la plupart établissent comme des dimensions incontournables le soutien du management pour la sécurité des soins, la communication au sein des équipes et entre les services, ou encore la réponse non punitive à l'erreur (Colla, Bracken et al. 2005). Le questionnaire auto-administré Hospital Survey on Patient Safety Culture (HSOPS), élaboré et validé aux Etats Unis par l'AHRQ (AHRQ Rockville 2004), est largement utilisé à travers le monde, notamment en Europe après avoir été approuvé par l'European Union Network for Patient Safety, et avoir été traduit dans de nombreuses langues (AHRQ Rockville 2015). Ce questionnaire HSOPS est largement utilisé par les équipes de recherche, aussi bien pour analyser les relations entre la culture de sécurité et les caractéristiques des répondants, que pour évaluer des interventions visant à améliorer la culture de sécurité (AHRQ Rockville 2014). Chaque année, l'AHRQ publie l'Hospital Survey on Patient Safety Culture Comparative Database qui sert à mesurer l'évolution de la culture sécurité dans des établissements de santé aux Etats-Unis, et à effectuer des comparaisons inter

établissements (Sorra, Famolaro et al. 2014). En 2014, 653 hôpitaux ont participé à cette enquête nationale, et 405 281 professionnels de santé ont été interrogés.

Le questionnaire HSOPS a été initialement développé par l'AHRQ à partir d'une revue de la littérature, et validé sur un échantillon de 1437professionnels de santé travaillant dans 21 hôpitaux des Etats-Unis (AHRQ Rockville 2004). Ce questionnaire comporte 42 items, dont la réponse est classée selon une échelle de Likert à 5 points, l'accord variant entre les réponses « pas du tout d'accord » à « tout à fait d'accord », et la fréquence entre « jamais » et « toujours ». Les 42 items sont ensuite regroupés selon 12 dimensions : perception globale de la sécurité ; fréquence du signalement des événements indésirables ; attentes et actions des supérieurs hiérarchiques pour la sécurité des soins ; organisation apprenante et amélioration continue ; travail d'équipe dans le service ; liberté d'expression ; retour et communication sur les erreurs ; réponse non punitive à l'erreur ; ressources humaines ; soutien du management pour la sécurité des soins ; travail d'équipe entre les services de l'établissement ; continuité des soins. A noter que les 42 items sont présentés en 5 sections (« votre service » ; « votre supérieur hiérarchique immédiat » ; « communication » ; « fréquence des événements indésirables » ; « votre établissement de santé »). Ces sections ne rentrent pas dans le calcul des scores dimensionnels et ne servent que de présentation et mise en page. Le questionnaire comporte également deux items supplémentaires ne rentrant pas dans le calcul des scores dimensionnels : un premier item concernant le niveau global de sécurité des soins perçu, classé selon 5 modalités variant de « Excellent » à « Défaillant », et un second concernant le nombre de fiches d'événements indésirables signalés, classé selon 6 modalités variant entre « Aucune » à « Plus de 20 fiches ». Enfin, le questionnaire comporte des items servant à recueillir les caractéristiques sociodémographiques des participants (profession, genre, âge, ancienneté), et une question ouverte destinée à recueillir des commentaires libres.

L'AHRQ rapporte une liste de 71 pays ayant utilisé le questionnaire HSOPS, et référence 32 traductions de l'instrument original (AHRQ Rockville 2015). Pour les enquêtes de culture de sécurité effectuées en langue française, il existe deux adaptations différentes, toutes deux référencées comme des versions françaises sur le site de l'AHRQ (AHRQ Rockville 2014). La première a été effectuée par une équipe de recherche en Belgique, et a été notamment utilisée pour réaliser des enquêtes nationales dans les établissements psychiatriques en Belgique (Vlayen, Hellings et al. 2012). Cette version (version « Vlayen ») est également utilisée en Suisse, où elle a fait l'objet d'une étude de validation externe menée auprès des professionnels de santé d'un centre hospitalier universitaire (Perneger, Staines et al. 2014). La seconde traduction de ce questionnaire en langue française a été conduite par les équipes du Comité de coordination de l'évaluation clinique et de la qualité en Aquitaine (CCECQA) (Occelli, Quenon et al. 2011). Cette version (version « Occelli ») a fait l'objet d'une diffusion en France par la Haute Autorité de Santé. Elle a récemment été utilisée en Tunisie, dans une enquête de culture sécurité auprès des professionnels de santé des blocs opératoire (Mallouli, Tlili et al. 2017).

3.2. Enquête culture sécurité

L'objectif de cette enquête était de mesurer les dimensions de la culture de sécurité de l'ensemble du personnel soignant du CHU de Grenoble, afin d'étudier les relations entre les CREX et la culture de sécurité des professionnels de santé.

3.2.1. Méthode

Schéma d'étude : il s'agissait d'une enquête transversale par questionnaire auto-administré auprès des professionnels des services cliniques et médicotechniques du CHU de Grenoble.

Population : l'enquête concernait tous les professionnels, (médicaux, paramédicaux, hôteliers, techniques et administratifs) travaillant au moins 50% du temps, dans les services médicaux du

CHU, depuis plus de 6 mois, présents au moment de l'enquête et acceptant de répondre au questionnaire.

Etaient exclus les agents à temps partiel inférieur à 50%, les agents ayant moins de 6 mois d'ancienneté dans l'établissement et les agents absents plus d'un mois au moment de l'enquête.

Questionnaire : nous avons utilisé la version « Occelli » du questionnaire HSOPS, reprenant les 42 items de la version originale américaine, à laquelle nous avons ajouté un item interrogeant les répondants sur leur participation à un CREX (Annexe).

Recueil des données : le recueil a été réalisé entre avril 2013 et septembre 2014. Comme recommandé par l'AHRQ, une lettre a été diffusée à l'ensemble des personnels éligibles, leur présentant les objectifs de l'étude. Le recueil a été organisé service par service en collaboration avec le cadre de santé du service. Une assistante de recherche établissait avec le cadre de santé la liste de tous les agents du service, distribuait les questionnaires et mettait en place une urne pour le recueil, munie d'une liste d'émargement. Les répondants retournaient le questionnaire dans l'urne scellée et signaient la liste des employés. Afin d'assurer la confidentialité, les questionnaires étaient anonymes. Plusieurs relances étaient envoyées aux professionnels n'ayant pas émargés sur la liste afin de maximiser les taux de réponses. L'objectif était d'atteindre pour chaque service un taux de réponse de 70%.

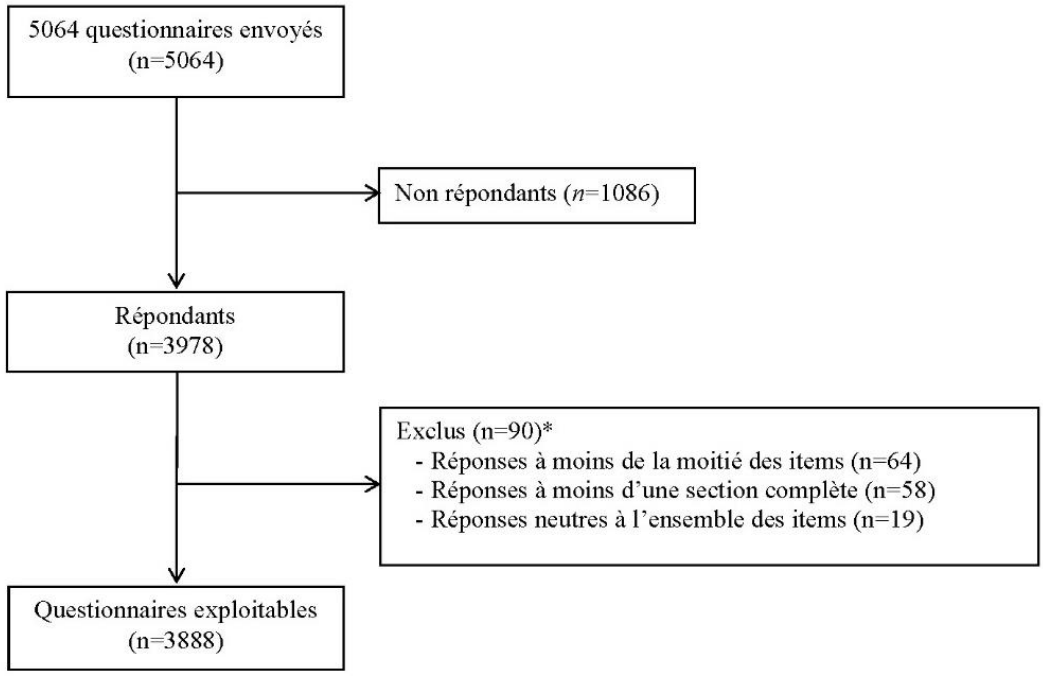
Conformément aux recommandations de l'AHRQ, les questionnaires étaient exclus lorsque moins de la moitié des 42 items étaient renseignés, qu'aucune section du questionnaire n'était complète, ou que l'ensemble des réponses étaient la valeur neutre 3.

Calcul des scores : avant le calcul des scores des 12 dimensions, les réponses aux items ayant une formulation négative étaient inversées, afin que les plus grands scores correspondent à une plus forte culture de sécurité. Les scores des 12 dimensions étaient ensuite calculés selon la méthode recommandée par l'AHRQ, à savoir le pourcentage des réponses positives, définies

comme les valeurs 4 (« d'accord » ou « la plupart du temps ») et 5 (« tout à fait d'accord » ou « toujours »). Pour obtenir les scores dimensionnels, la première étape était de calculer pour chaque item, le pourcentage de participants ayant donné une réponse positive sur le nombre total de participants ayant répondu à l'item. La seconde étape était de calculer la moyenne non pondérée des pourcentages de réponse positive des items composant chaque dimension, aboutissant à un score compris entre 0 et 100 pour chacune des 12 dimensions.

3.2.2. Résultats

Sur les 5 064 professionnels de santé éligibles de l'établissement auxquels ont été distribués le questionnaire, 3 978 ont répondu, soit 78,6 % (figure 1). L'application des critères de validité a conduit à exclure 90 questionnaires : 64 participants avaient répondu à moins de 21 items, 58 n'avait pas répondu complètement à au moins une section, et 19 n'avait répondu que des réponses neutres. Au total, 3 888 questionnaires étaient exploitables pour l'analyse (soit 76,8% des questionnaires distribués).



* Un questionnaire pouvait avoir plusieurs critères d'exclusion

Figure 1. Diagramme de flux des questionnaires

La majorité des participants étaient des femmes (80%), avaient moins de 45 ans (63%) et travaillaient au contact direct des patients (88%) (Tableau 1).

Tableau 1. Caractéristiques des répondants

Caractéristiques, *n* (%)*	(*n*=3,888)	
Femme	3,016	(80.6)
Age, *année*		
< 35	1,406	(37.5)
35–44	966	(25.7)
45–54	967	(25.8)
≥ 55	415	(11.1)
Profession		
Infirmier	1,386	(36.3)
Aide-soignant	708	(18.6)
Médecin	436	(11.4)
Autre soignant	124	(3.3)
Administratif	331	(8.7)
Technique	378	(9.7)
Autre	450	(11.8)
Pôle hospitalier		
Médecine et pédiatrie	1,159	(29.8)
Chirurgie et gynécologie	833	(21.4)
Médicotechnique (pharmacie, bloc opératoire, imagerie, laboratoire)	795	(20.4)
Réanimation, urgence, and anesthésie	688	(17.7)
Plusieurs secteurs	413	(10.6)
Contact direct avec les patients	3,357	(88.0)
Participation à un CREX	440	(11.3)

* Des réponses étaient manquantes pour le genre (*n*=147), l'âge (*n*=134), la profession (*n*=75), and le contact direct avec les patients (*n*=55).

Les scores les plus élevés étaient observés pour les dimensions « travail en équipe dans le service » et « liberté d'expression ». Les scores les plus faibles étaient observés pour les dimensions « soutien du management pour la sécurité des soins » et « transmissions et transferts » (Tableau 2).

Tableau 2. Résultats des scores par dimensions de la culture de sécurité

Dimension	Score sur 100	Classement
1. Perception globale de la sécurité des soins	42	7
2. Fréquence de signalement des évènements indésirables	50	5
3. Attentes et actions des supérieurs hiérarchiques	57	3
4. Organisation apprenante et amélioration continue	50	5
5. Travail d'équipe dans le service	62	1
6. Liberté d'expression	59	2
7. Retour d'information et communication des erreurs	55	4
8. Réponse non punitive à l'erreur	29	10
9. Ressources humaines	32	9
10. Soutien du management pour la sécurité des soins	20	12
11. Travail d'équipe entre les services	35	8
12. Transmissions et transferts	25	11

Avant d'analyser les relations entre le CREX et la culture de sécurité des professionnels de santé, nous avons effectué plusieurs travaux portant sur les propriétés métrologiques du questionnaire HSOPS. Tout d'abord, la version « Occelli » du questionnaire HSOPS que nous avons utilisée n'avait était validée que de façon interne, sur un petit échantillon de répondants (Occelli, Quenon et al. 2013). Il nous est apparu nécessaire de compléter l'étude des propriétés psychométriques du questionnaire. Nous avons réalisé une analyse factorielle confirmatoire afin de valider la possibilité de réaliser des analyses de scores sur les 12 dimensions de la culture sécurité. Ce travail était également l'occasion de pouvoir comparer les propriétés psychométriques de la version « Occelli » avec celles de la version « Vlayen », version développée en Belgique et également référencée sur le site de l'AHRQ (AHRQ Rockville 2015).

Par ailleurs, nous avons retrouvé dans la littérature 3 méthodes différentes d'agrégation des scores par dimension. Malgré la recommandation par l'AHRQ d'effectuer des moyennes non

pondérées de pourcentage de réponses positives aux items, de nombreuses études utilisaient des moyennes de scores individuelles, sans dichotomisation préalable des items. Le site web de l'AHRQ référence ces études, et la validation initiale des propriétés psychométriques du questionnaire original effectuée par une équipe de l'AHRQ avait également utilisé cette méthode de moyenne individuelle (AHRQ Rockville 2004, AHRQ Rockville 2014). Il nous est apparu intéressant d'explorer l'influence du mode de calcul des scores sur le résultat final. Nous présentons un article étudiant la variabilité des méthodes d'agrégation des scores, dans lequel nous avons étudié l'impact de ces méthodes sur le classement de la culture de sécurité.

Enfin, la question de la gestion des données manquantes pouvant avoir un impact sur les analyses, nous avons réalisé une large étude de simulation, à partir des questionnaires complets de notre enquête, afin de comparer les performances de plusieurs méthodes d'imputation des données manquantes. A notre connaissance, cette question pour les données manquantes du HSOSP n'avait fait l'objet d'aucun travail publié, conduisant là aussi à une hétérogénéité de choix méthodologiques pour les équipes de recherche utilisant le questionnaire HSOPS, certaines effectuant des imputations simples, et d'autres des imputations multiples (Smits, Christiaans-Dingelhoff et al. 2008, Hammer, Ernstmann et al. 2011, Ito, Seto et al. 2011, Feng, Bobay et al. 2012, Vlayen, Schrooten et al. 2015).

3.3. Propriétés métrologiques du questionnaire HSOPS

A partir des données de l'enquête de culture de sécurité réalisée au niveau de notre établissement, nous avons donc analysé les propriétés métrologiques de la version française du questionnaire HSOPS, développée par le CCECQA (version Occelli) : propriétés psychométriques, variabilité des méthodes de d'agrégation des scores des 12 dimensions, performance des méthodes imputation des données manquantes. Ces analyses ont donné lieu à 3 articles, deux publiés et un soumis, présentés dans cette thèse.

3.3.1. Etude de validation externe du questionnaire : article publié – Inconsistencies between two cross-cultural adaptations of the Hospital Survey on Patient Safety Culture into French

Considérant que la version française de l'HSOPS développée par le CCECQA (version « Occelli ») n'avait été validée que de façon interne sur un échantillon de taille limitée, cette étude visait à réaliser la validation externe de l'outil, à partir d'un large échantillon de professionnels de santé (Occelli, Quenon et al. 2013). De plus, nous avons effectué une comparaison post-hoc de nos résultats avec une précédente étude de validation de la version française concurrente (version « Vlayen ») réalisée aux Hôpitaux Universitaires de Lausanne, en Suisse (Perneger, Staines et al. 2014). En effet, aucune comparaison n'avait été effectuée entre la version « Occelli » et la version « Vlayen », pourtant toutes deux citées et référencées en tant que « French version » officielles sur le site de l'AHRQ (AHRQ Rockville 2015).

Objectif : L'objectif de cette étude était de valider de façon externe la version « Occelli » du questionnaire HSOPS, et de comparer ses propriétés psychométriques à celles rapportées pour l'instrument original et pour la version « Vlayen ».

Méthode : Nous avons utilisé les données de l'enquête culture sécurité menée au CHU de Grenoble. Nous avons examiné l'acceptabilité, la consistance interne, la structure factorielle et la validité de construit de la version « Occelli » du questionnaire HSOPS, et nous avons comparé ces résultats avec ceux précédemment rapportés par Thomas Perneger dans l'étude de validation externe de la version « Vlayen » (Perneger, Staines et al. 2014).

Résultats : Nous avons obtenu des taux de données manquantes et des scores dimensionnels inférieurs à ceux rapportés pour la version « Vlayen », et des scores dimensionnels inférieurs. Les coefficients alpha de Cronbach étaient inférieurs pour la version « Occelli » (médiane, 0.64 ; étendue, 0.56-0.84) comparativement à la version « Vlayen » (médiane, 0.73 ; étendue 0.57-0.86) et à la version originale de l'HSOPS (médiane, 0.78 ; étendue, 0.63-0.84). Les

résultats de l'analyse factorielle confirmatoire étaient satisfaisants, et consistants entre les versions « Occelli » et « Vlayen », indiquant la validité de la structure à 12 dimensions pour ces deux versions.

Conclusions : Cette étude montrait que les deux versions françaises pouvaient être utilisées de façon interchangeable pour réaliser des enquêtes de culture sécurité auprès des professionnels de santé francophones. Nous avons discuté plusieurs hypothèses pour justifier les écarts observés dans les propriétés des deux versions, pourtant toutes deux développées en suivant les mêmes procédés d'adaptation transculturelle, tels que recommandés par l'AHRQ. On peut notamment citer l'absence de contre traduction, probablement à l'origine de différences non négligeables dans la formulation de certains items entre les deux versions.

Inconsistencies Between Two Cross-Cultural Adaptations of the Hospital Survey on Patient Safety Culture Into French

Bastien Boussat, MD,† Patrice François, MD, PhD,*† Gérald Gandon, MSc,* Joris Giai, MD,‡ Arnaud Seigneurin, MD, PhD,*† Thomas Perneger, MD, PhD,§ and José Labarère, MD, PhD*†||*

Objectives: Two cross-cultural adaptations of the 12-dimension Hospital Survey on Patient Safety Culture (HSOPSC) into French coexist: the Occelli and Vlayen versions. The objective of this study was to assess the psychometric properties of the Occelli version in comparison with those reported for the Vlayen and the original US versions of this instrument.

Methods: Using the original data from a cross-sectional study of 5,064 employees at a single university hospital in France, we examined the acceptability, internal consistency, factorial structure, and construct validity of the Occelli version of the HSOPSC.

Results: The response rate was 76.8% (n = 3888). Our study yielded lower missing value rates (median, 0.4% [range, 0.0%–2.4%] versus 0.8% [range, 0.2%–11.4%]) and lower dimension scores (median, 3.19 [range, 2.67–3.54] versus 3.42 [range, 2.92–3.96]) than those reported for the Vlayen version. Cronbach alphas (median, 0.64; range, 0.56–0.84) compared unfavorably with those reported for the Vlayen (median, 0.73; range, 0.57–0.86) and original US (median, 0.78; range, 0.63–0.84) versions. The results of the confirmatory factor analysis were consistent between the Vlayen and Occelli versions, making it possible to conduct surveys from the 12-dimensional structure with both versions.

Conclusions: The inconsistencies observed between the Occelli and Vlayen versions of the HSOPSC may reflect either differences between the translations or heterogeneity in the study population and context. Current evidence does not clearly support the use of one version over the other. The two cross-cultural adaptations of the HSOPSC can be used interchangeably in French-speaking countries.

Key Words: patient safety culture, cross-cultural adaptation, HSOPSC, Occelli version, Vlayen version

(*J Patient Saf* 2017;00: 00–00)

Medical errors are the third leading cause of death in the United States, despite patient safety being a priority for health-care systems for more than 15 years.[1] The usual perception of medical errors as failures by individual staff members is outdated, with full awareness that adverse events result from multiple organizational failures.[2] The greater awareness of these failures has demonstrated the need to move toward a collaborative teamwork culture, where all actors share the same vision and goals and consider themselves responsible for improving patient safety. One of the ways to prevent adverse events is the development of a patient safety culture shared by all health-care professionals.[3–5] The need to measure patient safety culture in health care seems crucial and has been the subject of numerous studies.[6–9]

Different tools have been developed to measure safety culture in various settings. Among them, the most widely used in hospital settings is the Hospital Survey on Patient Safety Culture (HSOPSC). The HSOPSC is a self-administered questionnaire designed by the Agency for Healthcare Research and Quality (AHRQ) to assess 12 dimensions of patient safety culture from the perspective of hospital staff in the United States. In 2014, approximately 405,281 hospital staff in 653 US hospitals participated in the sixth nationwide HSOPSC questionnaire-based survey.[10]

Multiple-language versions of the HSOSPC are needed for international comparison purposes and in international multicenter clinical trials. This questionnaire has been translated into 29 different languages and used worldwide since its original development and validation.[11] Two research teams have independently conducted cross-cultural adaptations of the HSOPSC into French, leading to the coexistence of concurrent versions (hereby designated as the Vlayen and Occelli versions) with substantial variations in item wordings and presentation.[12–14] To our knowledge, these two versions have never been compared. In a recent study conducted in Tunisia, the patient safety culture in operating rooms was surveying using the Occelli version, without explaining the reasons for this choice.[15] Indeed, the AHRQ Web site does not distinguish the two concurrent versions, listing both the Occelli and the Vlayen versions as "French translation of the HOSPSC".[16] Although the Vlayen version was primarily intended for nationwide surveys of French-speaking hospital employees in Belgium,[14,17,18] its psychometric properties have also been evaluated in 1,171 staff members working at a multisite hospital in Switzerland.[19] The Occelli version, which was primarily designed for internal use throughout a regional hospital network,[12] yielded promising psychometric properties in 401 staff members recruited at seven hospitals in the Aquitaine region of France.[13] Yet this study had a relatively limited sample size, and to our knowledge, no external validation was conducted on an independent sample.

The aims of this study were to externally validate the psychometric properties of the Occelli version of the HSOPSC and evaluate whether these properties were consistent with those reported for the Vlayen and the original US versions of this instrument.

METHODS

Study Design

We conducted a cross-sectional survey of hospital staff, using a self-administered questionnaire. The present study complied with guidelines for reporting measurement properties of questionnaires.[20]

From the *Quality of Care Unit, Grenoble University Hospital; †Grenoble Alps University, TIMC UMR 5525 CNRS, France; ‡Service de Biostatistique, Hospices Civils de Lyon, Laboratoire de Biométrie et Biologie Evolutive, Lyon, France; §Division of Clinical Epidemiology, Geneva University Hospitals and Faculty of Medicine, University of Geneva, Geneva, Switzerland; and ||INSERM CIC 1406, Grenoble, France.

Correspondence: Bastien Boussat, MD, Quality of Care Unit, Pavillon Taillefer, CHU CS 10 217, F-38043, Grenoble Cedex, France (e-mail: bboussat@chu-grenoble.fr).

The authors disclose no conflict of interest.

This work was supported by a grant from the French Ministry of Health (Programme de recherche sur la performance du système de soins-PREPS 12-026-0001).

This work was developed within the Grenoble Alpes Data Institute, which is supported by the French National Research Agency under the "Investissements d'avenir" program (ANR-15-IDEX-02).

Supplemental digital contents are available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.journalpatientsafety.com).

Copyright © 2017 Wolters Kluwer Health, Inc. All rights reserved.

The project was submitted for advisory purposes to an independent protection committee (IRB00006705) and was approved by the French data protection authority (CNIL).

Study Site

The study was conducted at a single university-affiliated hospital with an 1,836-bed capacity serving a predominantly urban population of 675,000 inhabitants in France. The study site reported 135,999 stays in 2014. The hospital staff comprised 4,422 registered health-care providers and 642 board-certified physicians, the vast majority of them specialty hospitalists.

Participants

Eligible participants were full-time or part-time (half-time or more) employees with at least six months of employment in the hospital's clinical, laboratory/pathology, radiology, or pharmacy departments.

As recommended,[21] HSOPSC questionnaires were excluded from the analytical sample if respondents answered less than one entire section, fewer than half of the items, or every item with the same non-neutral response.

HSOPSC Questionnaire

The HSOPSC questionnaire was originally developed by the AHRQ on the basis of a literature review, refined in accordance with psychometric theory, and supported by psychometric analyses conducted on 1,437 hospital staff working in 21 US hospitals in 2004.[22] The cross-cultural adaptation procedure of the HSOPSC into French and its preliminary psychometric properties were reported in detail by Occelli et al.[13] Briefly, the original HSOPSC questionnaire was independently translated by two panels consisting of qualified translators, methodologists, health-care providers, and researchers with field expertise in patient safety. Reconciliation of the two independent forward translations resulted in a version translated into French including all 42 items of the original US version, with the addition of three further items measuring a supplementary dimension labeled "training and organizational learning." However, this a priori dimension achieved poor internal consistency and was not verified as a 13th dimension in the internal structure. Consequently, the present survey questionnaire did not include these three additional items. No backward-translation was done, and the resulting survey questionnaire was not submitted to the original developers of the HSOPSC to assess its consistency with the original US version.

Each item was rated on a five-point Likert scale to determine agreement, ranging from "strongly disagree" to "strongly agree," and frequency, ranging from "never" to "always," in accordance with the original HSOPSC. The present analysis focused on the 42 items covering the 12 dimensions that were included in the original US version of the HSOPSC questionnaire (online supplementary appendix 1, http://links.lww.com/JPS/A127).

Employees were asked to give an overall grade on patient safety in their department using a five-point response scale, ranging from excellent to failing, and to indicate how many events they had reported within the past 12 months. Although these two single-item outcome variables were part of the HSOPSC, they were not used to compute the 12 dimension scores.

Information on respondent background characteristics was collected at the end of the questionnaire, including age group, sex, position, seniority in the department, and working time (expressed as a percentage of full-time).

Data Collection

The data were collected anonymously on a volunteer basis, department by department, between April 2013 and September 2014.

Safety Culture Scores

Items were coded using a five-point response scale, with higher values denoting a better rating, after reverse coding for negatively worded items.[21] For each respondent, we computed 12 dimension scores, each of them corresponding to a key patient safety culture dimension. Each dimension score was computed by the arithmetic mean of the individual items and ranged from 1 to 5.[23]

Consistent with the external validation study of the Vlayen version,[19] dimension scores were coded as missing if all corresponding items were unanswered. In a separate analysis, we coded dimension scores as missing if one or more items were unanswered, to compare missing value rates with those reported in the preliminary validation study of the Occelli version.[13]

Statistical Analysis

The background respondent characteristics were reported as numbers and percentages and composite scores as means along with standard deviations. Using established criteria,[24,25] we analyzed the psychometric properties of the Occelli version of the HSOPSC, including its acceptability, internal consistency and structure, and construct validity.

Acceptability was assessed through the numbers and percentages of missing values for each item and dimension score. We also reported the number of questionnaires with missing values for one or more items. The numbers and percentages of the responses on anchor points for items and dimension scores were examined to detect floor and ceiling effects. Floor and ceiling effects lower than 15% for dimension scores were considered acceptable.[24]

Internal consistency was evaluated through average interitem correlation, item-rest correlation, and the Cronbach alpha coefficient. The internal consistency criterion was fulfilled for the item-rest correlation greater than 0.40, and Cronbach alpha greater than 0.70 was considered satisfactory.[24]

The internal structure of the HSOPSC was verified using a confirmatory factor analysis (CFA). To determine whether the survey data fit with the internal structure of the original US version of the HSOPSC, we performed structural equation modeling (SEM) of 12 dimensions with the 42 items assigned to the intended dimensions. The SEM corresponded to an external model representing the relationships between the dimension (latent variables) and their related items (manifest variables). The root mean square error of approximation, the comparative fit index, the non-normed fit index, the standardized root mean residual, and the coefficient of determination were obtained. These results were compared with the results of the external validation study of the Vlayen version and with the recommended cutoff: we considered a root mean square error with an approximation value of 0.10 or less, a standardized root mean residual value of 0.06 or less, a goodness-of-fit index value of 0.95 or more, a comparative fit index, and a non-normed fit index value of 0.90 or more as indicative of satisfactory model fit.[26]

We carried out exploratory factor analysis (EFA) to examine a possible alternative structure of the Occelli version of the HSOPSC. The principal component factor method was used along with orthogonal varimax rotation of factors with eigenvalues higher than 1.00. The Kaiser-Meyer-Olkin measure of sampling adequacy (>0.6 was considered satisfactory) and the Bartlett test of independence were used to verify the factorability of the sample. Primary loadings on intended dimensions

© 2017 Wolters Kluwer Health, Inc. All rights reserved.

higher than 0.40 with cross-loadings lower than 0.30 were considered satisfactory.

Consistent with the external validation study of the Vlayen version,[19] construct validity was assessed by comparing mean dimension score values for each dimension between subgroups of respondents defined by the two single-item outcome variables. We hypothesized that mean dimension scores were higher for respondents who rated safety as excellent or very good and for those who had reported one or more incidents within the past 12 months. The mean differences between the groups were quantified using the Cohen *d* effect size.

Differences in categorical variables were analyzed using χ^2 tests or Fisher exact tests where appropriate, and continuous variables were compared using Student *t* tests. Two-sided *P* values lower than 0.05 were considered statistically significant. All analyses were performed using Stata version 11.0 (Stata Corporation, College Station, TX).

RESULTS

Respondent Characteristics

Of 5,064 eligible employees, 3,978 (78.6%) participated in the study. After the exclusion of 90 questionnaires due to the discovery of exclusion criteria (online supplementary appendix 2, http://links.lww.com/JPS/A127), the sample consisted of 3,888 survey questionnaires (76.8%). The response rate was consistent with that of the preliminary validation study of the Occelli version (76.8% versus 76.5%, $P = 0.24$)[13] and higher than that observed for the external validation study of the Vlayen version (76.8% versus 74.0%, $P < 0.001$).[19] Overall, most participants were female (80%), younger than 45 years (63%), and likely to report contact with patients (88%) (online supplementary appendix 3, http://links.lww.com/JPS/A127).

Acceptability

Missing value rates ranged from 0.5% to 3.3% for the items (online supplementary appendix 4, http://links.lww.com/JPS/A127) and from 0.0% to 2.4% for the dimension scores (Table 1). A nonsignificant trend toward higher percentages of missing values was observed for 18 negatively worded items in comparison to 24 positively worded items (on average, 2.1% versus 1.6%, $P = 0.06$). Overall, 843 questionnaires (21.7%) yielded missing values for one or more items.

Missing value rates did not differ for 9 of 12 dimension scores between the preliminary validation study of the Occelli version and the present study (online supplementary appendix 5, http://links.lww.com/JPS/A127). In comparison with the external validation study of the Vlayen version, our study found lower missing value rates (median, 0.4% [range, 0.0% 2.4%] versus 0.8% [range, 0.2% 11.4%]) (Table 1). The largest discrepancy was found for the "frequency of event reporting" dimension, which yielded a lower missing value rate (2.4% versus 11.4%, $P < 0.001$) and a higher mean score (3.38 versus 3.22, $P < 0.001$) in the present study (Table 2).

Dimension Score Distributions

Mean dimension scores ranged from 2.67 to 3.54 (Table 2), with no floor or ceiling effects observed. Yet the highest values (5/5) for eight items and the lowest values (1/5) for two other items accrued more than 15% of the respondents (online supplementary appendix 4, http://links.lww.com/JPS/A127). In comparison with the external validation study of the Vlayen version, the mean dimension scores were lower for all unit- and hospital-level process dimensions (range, 2.67 3.54 versus 2.92-3.96) (Table 2).

Internal Consistency

The Cronbach alphas (median, 0.64; range, 0.56 0.84) were consistent with that observed in the preliminary validation study of the Occelli version (0.63; range, 0.46 0.84) but compared unfavorably with those reported in the external validation of the Vlayen version (median, 0.73; range, 0.57 0.86) and the original US development (median, 0.78; range, 0.63 0.84) studies (Table 3). Cronbach alphas were below 0.70 for one, five, and nine dimension

TABLE 1. Number (Percentages) of Missing Values* for the Occelli and Vlayen Versions of the Hospital Survey on Patient Safety Culture

Dimension Score	No. Items	Present Study† (n = 3,888)		External Validation Study of the Vlayen Version[9] (n = 1,171)		*P*
Outcomes						
Overall perception of safety	4	6	(0.2)	4	(0.3)	0.25
Frequency of event reporting	3	94	(2.4)	133	(11.4)	<0.001
Unit-level processes						
Supervisor/manager expectations and actions	4	33	(0.8)	7	(0.6)	0.40
Organizational learning	3	7	(0.2)	3	(0.3)	0.71
Teamwork within hospital units	4	0	(0.0)	2	(0.2)	0.05
Communication openness	3	14	(0.4)	11	(0.9)	0.01
Feedback and communication about error	3	16	(0.4)	12	(1.0)	0.01
Nonpunitive response to error	3	7	(0.2)	6	(0.5)	0.09
Staffing	4	0	(0.0)	3	(0.3)	0.01
Hospital-level processes						
Hospital management support	3	48	(1.2)	18	(1.5)	0.42
Teamwork across hospital units	4	47	(1.2)	19	(1.6)	0.27
Hospital handoffs and transitions	4	67	(1.7)	26	(2.2)	0.27

*Dimension scores were coded as missing if all corresponding items were unanswered (see Methods).

†The present study used the Occelli version of the Hospital Survey on Patient Safety Culture.

Original (US) data were unavailable.

© 2017 Wolters Kluwer Health, Inc. All rights reserved.

TABLE 2. Mean (Standard Deviation) Dimension Scores for the Occelli and Vlayen Versions of the Hospital Survey on Patient Safety Culture

Dimension Score	No. Items	Present Study* (n = 3,888)		External Validation Study of the Vlayen Version[9] (n = 1,171)		P
Outcomes						
Overall perception of safety	4	3.21	(0.71)	3.20	(0.74)	0.68
Frequency of event reporting	3	3.38	(0.76)	3.22	(0.99)	<0.001
Unit-level processes						
Supervisor/manager expectations and actions	4	3.50	(0.81)	3.72	(0.68)	<0.001
Organizational learning	3	3.40	(0.64)	3.73	(0.61)	<0.001
Teamwork within hospital units	4	3.54	(0.76)	3.96	(0.68)	<0.001
Communication openness	3	3.50	(0.71)	3.84	(0.76)	<0.001
Feedback and communication about error	3	3.38	(0.74)	3.59	(0.87)	<0.001
Nonpunitive response to error	3	2.94	(0.75)	3.30	(0.74)	<0.001
Staffing	4	2.88	(0.74)	3.36	(0.75)	<0.001
Hospital-level processes						
Hospital management support	3	2.67	(0.74)	2.92	(0.82)	<0.001
Teamwork across hospital units	4	3.05	(0.58)	3.22	(0.62)	<0.001
Hospital handoffs and transitions	4	2.88	(0.61)	3.02	(0.70)	<0.001

*The present study used the Occelli version of the Hospital Survey on Patient Safety Culture.
Original (US) data were unavailable.

scores of the original US, Vlayen, and Occelli versions of the HSOPSC, respectively. In the present study, the internal consistency criterion was not fulfilled for 13 items with an item-rest correlation lower than 0.40 (online supplementary appendix 4, http://links.lww.com/JPS/A127).

Internal Structure

In CFA including 3,045 observations without missing values, SEM of 42 items apportioned in 12 latent factors yielded a root mean square error of approximation, standardized root mean residual, comparative fit index, and goodness-of-fit index estimates achieving recommended thresholds and consistent with those reported in the external validation study of the Vlayen version (Table 4). These results were less advantageous than those of the original instrument.

Alternative Structure

The Kaiser-Meyer-Olkin measure of sampling adequacy was satisfactory (90.4%), showing a high proportion of variance among variables, and the Bartlett test of sphericity was rejected as recommended, allowing the factorability of the sample. The EFA of the 42 items departed from the hypothesized 12-dimension structure: the Occelli and Vlayen versions of the HSOPSC questionnaire yielded 10 principal components with eigenvalues higher than 1.0, but with substantial differences in factor structure (online supplementary appendix 6, http://links.lww.com/JPS/A127).

TABLE 3. Cronbach Alpha for the Occelli, Vlayen, and Original (US) Versions of the Hospital Survey on Patient Safety Culture

Dimension Scores	No. Items	Occelli Version of the HSOPSC		External Validation Study of the Vlayen Version[9] (n = 1,171)	Original (US)[1] (n = 1,437)
		Present Study* (n = 3,888)	Preliminary Validation Study[5] (n = 401)		
Overall perception of safety	4	0.65 (0.63–0.67)	0.67	0.68	0.74
Frequency of event reporting	3	0.84 (0.82–0.85)	0.84	0.86	0.84
Supervisor expectations and actions	4	0.81 (0.80–0.82)	0.83	0.75	0.75
Organizational learning	3	0.62 (0.59–0.64)	0.59	0.57	0.76
Teamwork within hospital units	3	0.71 (0.69–0.72)	0.63	0.80	0.83
Communication openness	3	0.63 (0.61–0.66)	0.62	0.67	0.72
Feedback and communication about error	4	0.66 (0.64–0.68)	0.64	0.77	0.78
Nonpunitive response to error	3	0.60 (0.58–0.63)	0.57	0.60	0.79
Staffing	4	0.60 (0.57–0.62)	0.46	0.61	0.63
Hospital management support	3	0.66 (0.64–0.68)	0.73	0.79	0.83
Teamwork across hospital units	4	0.56 (0.53–0.59)	0.59	0.71	0.80
Hospitals handoffs and transitions	4	0.60 (0.58–0.63)	0.66	0.78	0.80

*Values are Cronbach alpha point estimates (95% bootstrap confidence interval).

© 2017 Wolters Kluwer Health, Inc. All rights reserved.

TABLE 4. Confirmatory Factor Analysis for the 42 items for the Original (US), Vlayen, and Occelli Versions of the Hospital Survey on Patient Safety Culture Questionnaire

	Original (US)[10]	External Validation Study of the Vlayen Version[18]	Preliminary Validation Study of the Occelli Version[14†]	Present Study*	Recommended Criteria of Good Fit
Root mean square error of approximation	0.04	0.04	0.05	0.04	<0.10
Comparative fit index	0.94	0.89	0.85	0.90	>0.90
Non-normed fit index	0.93	0.88	—	0.89	>0.90
Standardized root mean residual	0.04	0.05	—	0.04	<0.06
Goodness-of-fit index	—	>0.99	—	>0.99	>0.95

*The present study used the Occelli version of the Hospital Survey on Patient Safety Culture.

†The non-normed fit index, standardized root mean residual, and goodness-of-fit index were not available for the preliminary validation study of the Occelli version.

Compared with the external validation study of the Vlayen version, the 10 principal components explained a lower percentage of overall variance (54% versus 58%), and fewer items had primary factor loadings greater than 0.40 (29 versus 41 out of 42; online supplementary appendix 7, http://links.lww.com/JPS/A127). None versus three of the 42 items yielded cross-loadings greater than 0.30 for the preliminary validation of the Occelli version and the external validation of the Vlayen version.

Construct Validity

Consistent with the findings from the external validation study of the Vlayen version, all mean dimension scores were significantly higher for respondents who reported excellent or very good safety grades compared with those who reported poor, fair, or good safety grades (Table 5). In contrast, reporting one or more events within the past 12 months was associated with comparable or significantly lower mean values for all but two dimension scores. The exceptions were the "frequency of event reporting" and the "organizational learning" mean dimension scores, which were significantly higher for respondents who reported one or more events within the past 12 months (3.41 versus 3.34, $P = 0.004$, and 3.42 versus 3.38, $P = 0.04$, respectively).

DISCUSSION

The present study provides additional evidence on the psychometric properties of the French-language version of the HSOPSC questionnaire, which was translated and underwent preliminary validation as reported in Occelli et al.[13] The results of this study are in line with previous studies that consistently reported worse performance for cross-cultural adaptations of the HSOPSC in

TABLE 5. Comparison of Hospital Survey on Patient Safety Culture Dimension Scores According to Safety Grade and Event Reporting

	Safety Grade				Cohen *d* Effect Size		Reported an Event Within the Past 12 Months				Cohen *d* Effect Size	
Dimension	Excellent or very good		Good, fair, or poor		Present study*‡	Swiss survey[9†‡]	One or more		None		Present study*	Swiss survey[9†]
Overall perceptions of safety	3.67	(0.57)	2.93	(0.67)	1.17	1.31	3.16	(0.75)	3.29	(0.67)	−0.18¶	−0.40¶
Frequency of event reporting	3.57	(0.72)	3.28	(0.76)	0.39	0.46	3.41	(0.77)	3.34	(0.75)	0.10#	0.13
Supervisor expectations and actions	3.82	(0.71)	3.31	(0.82)	0.65	0.41	3.47	(0.86)	3.54	(0.75)	−0.09#	0.09
Organizational learning	3.67	(0.56)	3.24	(0.65)	0.70	0.62	3.42	(0.67)	3.38	(0.61)	0.07#	0.08
Teamwork within hospital units	3.82	(0.65)	3.38	(0.76)	0.61	0.50	3.56	(0.75)	3.55	(0.76)	0.01	0.08
Communication openness	3.75	(0.61)	3.36	(0.72)	0.57	0.58	3.52	(0.72)	3.48	(0.69)	0.06	−0.02
Feedback and communication about error	3.67	(0.65)	3.21	(0.73)	0.65	0.72	3.37	(0.75)	3.40	(0.71)	−0.04	−0.05
Nonpunitive response to error	3.15	(0.74)	2.83	(0.73)	0.43	0.55	2.96	(0.77)	2.94	(0.73)	0.02	0.07
Staffing	3.19	(0.73)	2.70	(0.70)	0.68	0.76	2.88	(0.76)	2.93	(0.71)	−0.07#	−0.19#
Hospital management support	2.93	(0.74)	2.49	(0.70)	0.61	0.79	2.56	(0.74)	2.80	(0.71)	−0.32¶	−0.18#
Teamwork across hospital units	3.24	(0.56)	2.93	(0.57)	0.54	0.51	3.00	(0.59)	3.11	(0.58)	−0.18¶	0.03
Hospital handoffs and transitions	3.02	(0.62)	2.78	(0.59)	0.40	0.47	2.83	(0.62)	2.93	(0.59)	−0.17¶	0.03

*The present study used the Occelli version of the Hospital Survey on Patient Safety Culture.

†The Swiss survey used the Vlayen version of the Hospital Survey on Patient Safety Culture (external validation study of the Vlayen version).

‡$P < 0.001$ for all dimensions.

#$P < 0.05$.

¶$P < 0.001$.

© 2017 Wolters Kluwer Health, Inc. All rights reserved.

comparison with the original instrument.[27,28] Yet this study also highlights differing acceptability, internal consistency, and alternative factor structure between two competing adaptations of the HSOPSC into French,[19] with potential implications for interpreting survey results on patient safety culture.

Overall, the Occelli version of the HSOPSC was developed on 401 staff members and was validated on 4,289 hospital employees across eight academic and nonacademic hospitals in France. When restricting the analysis to the 42 items covering the 12 dimensions of the original HSOPSC, acceptability and internal consistency estimates were comparable to those reported by the preliminary validation study of the Occelli version of this instrument.[13] Yet Occelli et al generated three additional items intended to compose a 13th dimension.[13] These substantial adjustments to the original HSOPSC structure might partly explain the differences in EFA findings between the preliminary validation study of the Occelli version and the present study.

The use of standardized validated instruments is advocated for surveying patient safety culture. Validated instruments are likely to accurately reflect the concept to be measured, whereas standardized instruments allow large-scale international comparisons of survey results.[29,30] In this context, not only the coexistence of two competing French-language adaptations of the HSOPSC but also the substantial variations observed in their psychometric properties raise questions. Several potential reasons may explain differing psychometric properties between the Vlayen and Occelli versions of the HSOPSC.

First, these two competing adaptations show numerous and substantial discrepancies in item wording. Overall, none of the 42 items are worded identically in the Vlayen and Occelli versions of the HSOPSC. Several differences in item wording likely reflect conceptual inconsistencies between the two versions. For instance, the words *error* and *mistake* were translated as *incident* (which is nearly equivalent to "adverse event") for 8 of 10 items in the Vlayen version, whereas the word *erreur* used in the Occelli version was a more literal translation. Other differences may stem from variations in translation of an idiomatic expression with a figurative meaning. The sentence "Things fall between the cracks," which has no equivalent in French, was therefore translated as "*des dysfonctionnements surviennent*" in the Occelli version ("dysfunctions occur"), whereas the Vlayen version used an expression with a more restrictive meaning "*des informations ne sont pas transmises*" ("information is not forwarded"). These observations may result from different translation and cultural adaptation procedures, although the two research teams complied with the approach advocated by the translation guidelines for AHRQ surveys on patient safety.[16,21] Accordingly, the two research teams did not perform backward translation, which may have contributed to inconsistencies in item wording. Yet evidence is currently lacking, and recommendations conflict on the added value of backward translation, like many other aspects of cross-cultural instrument adaptation, including the optimal number of forward translations, reconciliation processes, and expert committees.[31,32] It is also unclear to what extent inconsistencies in item wording may alter the psychometric properties of survey instruments. Indeed, a previous study reported almost identical psychometric properties for two competing cross-cultural adaptations of the SF36 health survey into French despite substantial differences in item wordings.[33]

Second, our findings may reflect more general problems with the original version of the HSOPSC, as previously noted by others.[19] Because the HSOPSC was originally developed and validated in the United States, the conceptual framework and factor structure may lack generalizability or robustness and therefore may not apply equally well to other countries. Overfitting is another potential issue that might contribute to optimistic performance estimates for the original version of the HSOPSC.[34] These hypotheses would explain why psychometric properties are not as good for cross-cultural adaptations of the HSOPSC in comparison with the original instrument.[19]

Third, discrepancies in acceptability estimates between the Vlayen and Occelli versions of the HSOPSC might be partly explained by true variations in context between the two external validation study sites. Typically, a formal incident-reporting system had not been implemented in all parts of the Swiss hospital at the time of the survey,[19] in contrast to our study site. This likely explains the higher missing value rate for the frequency of the event-reporting dimension (11.4% versus 2.4%), assuming that respondents were less familiar with incident reporting in the Swiss hospital.[19]

Fourth, the patient safety culture construct may also vary across health-care systems, explaining inconsistencies in EFA results between the Vlayen and Occelli versions. For instance, two items conceptually related to staffing or oversaturated departments, namely A7 ("we use more agency/temporary staff than is best for patient care") and A11 ("when one area in this unit gets really busy, others help out") yielded primary loadings lower than 0.40 across dimensions. Although speculative, one potential explanation for this observation may relate to the shortage of nurses, lack of investments, and growing discouragement of health-care providers in public hospitals in France over the last decade.[35] Despite a less than stellar EFA in our study, the results of the CFA were consistent between the Vlayen and Occelli versions. The CFA results were obviously less satisfactory than those of the original instrument, but this issue is raised for the majority of the instrument's translated versions, given the relative overfitting of the original instrument to its specific context. However, the CFA was relatively successful for the two surveys, making it possible to perform surveys from the 12-dimensional structure with both versions.

Fifth, complex relationships between response rates and item response patterns have been reported for health surveys.[36] Accordingly, cross-cultural adaptation studies with lower (median, 32%) response rates[27,37–39] reported a trend toward higher HSOPSC dimension scores (range, 3.40–3.48 versus 3.33–3.46) and Cronbach alphas (range, 0.69–0.79 versus 0.64–0.73) compared with studies with higher (median, 75%) response rates.[19,28,40,41] Yet it is unlikely that the inconsistencies observed in psychometric properties relate to slight, although statistically significant, differences in response rates between the external validation study of the Vlayen version and present survey (74.0% versus 76.8%).

Sixth, differences in respondent characteristics might have confounded the comparisons of psychometric properties between the Vlayen and Occelli versions of the HSOPSC. Missing value rates, mean values, and Cronbach alphas for dimension scores varied according to baseline respondent characteristics in our study (data not shown). No external validation study was conducted in Belgium, and large-scale head-to-head comparisons of the psychometric properties for the Vlayen and Occelli versions of the HSOPSC is currently lacking. Only two-arm parallel-group or cross-over randomized controlled trials conducted across different French-speaking countries can address this question adequately.

The limitations of our study deserve mention. First, we did not collect information on nonrespondent characteristics, and therefore, we could not examine the potential for selection bias. Second, construct validity was evaluated using self-reported rather than routinely collected data, a fact that might explain the unexpected associations between event reporting and lower HSOPSC scores. Third, we did not assess test/retest reliability,

© 2017 Wolters Kluwer Health, Inc. All rights reserved.

although a previous study reported intraclass correlation coefficient point estimates ranging from 0.11 to 0.77 across dimension scores.[13]

CONCLUSIONS

Using validated cross-cultural adaptations of the HSOPSC is essential to assess initiatives aiming to promote patient safety. Consistent with the external validation study of the Vlayen version, the 12-factor internal structure of the Occelli version was confirmed. This suggests that patient safety researchers and health-care providers in French-speaking countries can interchangeably use both of the adaptations of the HSOPSC. Although consistent with the preliminary validation study by Occelli et al., the acceptability, internal consistency, and factor structure of the Occelli version of the HSOPSC differed from those reported for the Vlayen version of this instrument. Whether the inconsistencies observed between these two cross-cultural adaptations of the HSOPSC into French reflect differing psychometric performance or heterogeneity in the study population and context remains unclear. Current evidence does not support the use of one French version of the HSOPSC over the other for surveying patient safety culture in French-speaking countries. We believe that only a direct comparison through a randomized controlled trial can definitely settle this issue. Hence, we would recommend using interchangeably one of the two versions of the HSOPSC if a patient safety culture survey was planned in France. Another implication of our study is the need for performing backward translation and reconciliation of the two French-language versions with the original US version.

ACKNOWLEDGMENTS

The authors thank Linda Northrup for assistance in editing the manuscript in English.

REFERENCES

1. Makary MA, Daniel M. Medical error-the third leading cause of death in the US. *BMJ*. 2016;353:i2139.
2. Reason J. Human error: models and management. *BMJ*. 2000;320:768 770.
3. Boussat B, Bougerol T, Detante O, et al. Experience Feedback Committee: a management tool to improve patient safety in mental health. *Ann Gen Psychiatry*. 2015;14:23.
4. Leape LL. Reporting of adverse events. *N Engl J Med*. 2002;347: 1633 1638.
5. Leape LL, Berwick DM, Bates DW. What practices will most improve safety? Evidence-based medicine meets patient safety. *JAMA*. 2002;288: 501 507.
6. Burlison JD, Quillivan RR, Kath LM, et al. A multilevel analysis of U.S. hospital patient safety culture relationships with perceptions of voluntary event reporting. *J Patient Saf*. 2016. Epub ahead of print.
7. Danielsson M, Nilsen P, Rutberg H, et al. A national study of patient safety culture in hospitals in Sweden. *J Patient Saf*. 2017. Epub ahead of print.
8. Pimentel MPT, Choi S, Fiumara K, et al. Safety culture in the operating room: variability among perioperative healthcare workers. *J Patient Saf*. 2017. Epub ahead of print.
9. Raeissi P, Reisi N, Nasiripour AA. Assessment of patient safety culture in Iranian Academic hospitals: strengths and weaknesses. *J Patient Saf*. 2015. Epub ahead of print.
10. Sorra J, Famolaro T, Dyer N, et al. Hospital Survey on Patient Safety Culture: 2014 User Comparative Database Report. Available at: http://www.ahrq.gov/sites/default/files/wysiwyg/professionals/quality-patient-safety/patientsafetyculture/hospital/2014/hsops14pt1.pdf. Accessed 2014.
11. Agency for Healthcare Research and Quality R, MD. Surveys on Patient Safety Culture Research Reference List. Available at: http://www.ahrq.gov/professionals/quality-patient-safety/patientsafetyculture/resources/index.html. Accessed 2014.
12. Occelli P, Quenon JL, Hubert B, et al. Development of a safety culture: initial measurements at six hospitals in France. *J Healthc Risk Manag*. 2011;30:42 47.
13. Occelli P, Quenon JL, Kret M, et al. Validation of the French version of the Hospital Survey on Patient Safety Culture questionnaire. *Int J Qual Health Care*. 2013;25:459 468.
14. Vlayen A, Hellings J, Claes N, et al. A nationwide hospital survey on patient safety culture in Belgian hospitals: setting priorities at the launch of a 5-year patient safety plan. *BMJ Qual Saf*. 2012;21:760 767.
15. Mallouli M, Tlili MA, Aouicha W, et al. Assessing patient safety culture in Tunisian operating rooms: a multicenter study. *Int J Qual Health Care*. 2017;29:176 182.
16. Agency for Healthcare Research and Quality, Rockville, MD. International Use of the Surveys on Patient Safety Culture. Available at: http://www.ahrq.gov/professionals/quality-patient-safety/patientsafetyculture/pscintusers.html. Accessed 2015.
17. Hellings J, Schrooten W, Klazinga N, et al. Challenging patient safety culture: survey results. *Int J Health Care Qual Assur*. 2007;20: 620 632.
18. Vlayen A, Schrooten W, Wami W, et al. Variability of patient safety culture in Belgian acute hospitals. *J Patient Saf*. 2015;11:110 121.
19. Perneger TV, Staines A, Kundig F. Internal consistency, factor structure and construct validity of the French version of the Hospital Survey on Patient Safety Culture. *BMJ Qual Saf*. 2014;23:389 397.
20. Mokkink LB, Terwee CB, Gibbons E, et al. Inter-rater agreement and reliability of the COSMIN (COnsensus-based Standards for the selection of health status Measurement Instruments) checklist. *BMC Med Res Methodol*. 2010;10:82.
21. Sorra J, Nieva VF. *Hospital Survey on Patient Safety Culture. (Prepared by Westat, under Contract No. 290-96-0004). AHRQ Publication No. 04 0041*. Rockville, MD: Agency for Healthcare Research and Quality; 2004.
22. Nieva VF, Sorra J. Safety culture assessment: a tool for improving patient safety in healthcare organizations. *Qual Saf Health Care*. 2003; 12:ii17 23.
23. Giai J, Boussat B, Occelli P, et al. Hospital survey on patient safety culture (HSOPS): variability of scoring strategies. *Int J Qual Health Care*. 2017;29:685 692.
24. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60:34 42.
25. Nunnally JC, Bernstein IH. *Psychometric Theory: McGraw-Hill; 3rd edition (January 1, 1994)*. 1994.
26. Kline RB. *Principles and practice of structural equations modelling, 2nd ed*. New York: The Guilford Press; 2005.
27. Hammer A, Ernstmann N, Ommen O, et al. Psychometric properties of the Hospital Survey on Patient Safety Culture for hospital management (HSOPS_M). *BMC Health Serv Res*. 2011;11:165.
28. Ito S, Seto K, Kigawa M, et al. Development and applicability of Hospital Survey on Patient Safety Culture (HSOPS) in Japan. *BMC Health Serv Res*. 2011;11:28.
29. Gjersing L, Caplehorn JR, Clausen T. Cross-cultural adaptation of research instruments: language, setting, time and statistical considerations. *BMC Med Res Methodol*. 2010;10:13.
30. Smits M, Christiaans-Dingelhoff I, Wagner C, et al. The psychometric properties of the 'Hospital Survey on Patient Safety Culture' in Dutch hospitals. *BMC Health Serv Res*. 2008;8:230.

© 2017 Wolters Kluwer Health, Inc. All rights reserved.

31. Epstein J, Santo RM, Guillemin F. A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *J Clin Epidemiol*. 2015;68:435 441.

32. Epstein J, Osborne RH, Elsworth GR, et al. Cross-cultural adaptation of the Health Education Impact Questionnaire: experimental study showed expert committee, not back-translation, added value. *J Clin Epidemiol*. 2015;68:360 369.

33. Perneger TV, Leplege A, Etter JF. Cross-cultural adaptation of a psychometric instrument: two methods compared. *J Clin Epidemiol*. 1999;52:1037 1046.

34. Labarere J, Renaud B, Fine MJ. How to derive and validate clinical prediction models for use in intensive care medicine. *Intensive Care Med*. 2014;40:513 527.

35. Batailler P, Francois P, Dang VM, et al. Trends in patient perception of hospital care quality. *Int J Health Care Qual Assur*. 2014;27: 414 426.

36. Labarere J, Francois P, Bertrand D, et al. [Evaluation of inpatient satisfaction. Comparison of different survey methods]. *Presse Med*. 2000;29:1112 1114.

37. Eiras M, Escoval A, Grillo IM, et al. The hospital survey on patient safety culture in Portuguese hospitals: instrument validity and reliability. *Int J Health Care Qual Assur*. 2014;27:111 122.

38. Sarac C, Flin R, Mearns K, et al. Hospital survey on patient safety culture: psychometric analysis on a Scottish sample. *BMJ Qual Saf*. 2011;20:842 848.

39. Waterson P, Griffiths P, Stride C, et al. Psychometric properties of the Hospital Survey on Patient Safety Culture: findings from the UK. *Qual Saf Health Care*. 2010;19:e2.

40. Haugen AS, Softeland E, Eide GE, et al. Patient safety in surgical environments: cross-countries comparison of psychometric properties and results of the Norwegian version of the Hospital Survey on Patient Safety. *BMC Health Serv Res*. 2010;10:279.

41. Nie Y, Mao X, Cui H, et al. Hospital survey on patient safety culture in China. *BMC Health Serv Res*. 2013;13:228.

© 2017 Wolters Kluwer Health, Inc. All rights reserved.

3.3.2. Variabilité des méthodes d'agrégation des scores dimensionnels : article publié – Hospital survey on patient safety culture (HSOPS): variability of scoring strategies

Au travers d'une étude de la littérature, nous avons recensé 3 différentes méthodes d'agrégation des scores dimensionnels de l'HSOPS. Cette hétérogénéité est susceptible d'entrainer des biais dans les résultats des enquêtes de culture sécurité et d'altérer les comparaisons inter-établissements, ou internationales, publiées dans la littérature (Fujita, Seto et al. 2013, Sorra, Famolaro et al. 2014).

Nous avons réalisé une étude comparant les performances des 3 méthodes d'agrégations utilisées dans la littérature, au travers de leur impact sur les classements des pôles hospitaliers (Giai, Boussat et al. 2017).

Objectif : L'objectif de cette étude était d'évaluer la variabilité des scores dimensionnels de culture de sécurité et leurs classements en fonction des trois méthodes d'agrégation utilisées pour le questionnaire HSOPS.

Méthodes : Nous avons utilisé les données de l'enquête culture sécurité menée au CHU de Grenoble. Nous avons agrégé les scores dimensionnels par pôle hospitalier selon 3 méthodes différentes. 1/ M1 : agrégation par la moyenne des pourcentages de réponses positives (dichotomisation des réponses, considérant que les réponses positives étaient celles strictement supérieures à 3, calcul du pourcentage de réponses positives pour chaque item, moyenne non pondérée des pourcentages de réponses positives aux 3 ou 4 items constituant une même dimension, le score varie donc entre 0 et 100) ; 2/ agrégation par la moyenne individuelle des réponses (moyennes individuelles des réponses aux 3 ou 4 items d'une même dimension, puis somme de ces moyennes individuelles divisée par le nombre de répondant, le score varie donc entre 1 et 5) ; 3/ agrégation par la moyenne des sommes individuelles (sommes des réponses aux 3 ou 4 items d'une même dimension, puis moyenne de ces sommes individuelles, le score varie donc entre 3 et 15 pour les dimensions à 3 items et entre 4 et 20 pour les dimensions à 4

items). Pour pouvoir comparer les 3 méthodes, nous avons standardisé les scores obtenus via M2 et M3 en utilisant une transformation linéaire simple pour obtenir des scores variant entre 0 et 100, et nous avons estimé par bootstrap des intervalles de confiance à 95 %. Pour chacune des 12 dimensions du HSOPS, nous avons établi un classement des 14 pôles de l'hôpital selon les 3 méthodes d'agrégation. On obtenait ainsi 3 classements pour la dimension « travail en équipe » par exemple, dans lesquels figuraient les 14 pôles. En théorie, le classement devrait être le même avec les 3 méthodes. Nous avons également examiné si les différences entre les méthodes avaient un impact au sein d'un service, par exemple s'il y avait des variations entre les dimensions. Les comparaisons de classements étaient également effectuées par genre et par groupes professionnels.

Résultats : Les résultats différaient sensiblement en fonction du type d'agrégation, et ce à plusieurs égards. La méthode M1 produisait des scores plus élevés en valeur absolue que M2 et M3. Pour la dimension « soutien du management à la sécurité des soins » par exemple, un pôle avait un score de 15% avec la méthode M1, contre 43% avec M2 et 48% avec M3. Les variations étaient plus ou moins marquées selon les dimensions. En outre, les classements variaient fortement en fonction du type d'agrégation. Un pôle était par exemple classé au 4^{e} rang sur 14 avec la méthode M1, alors qu'il figurait à la 13^{e} place avec la méthode M2. Les différences les plus importantes étaient observées entre la méthode 1 et les méthodes 2 ou 3, ce qui s'explique par la réduction de la variance liée à la dichotomisation dans la méthode M1, alors que les méthodes M2 et M3 conservaient toute l'information disponible.

Conclusions : Dans cette étude, nous avons montré que la procédure utilisée pour agréger les scores des dimensions des enquêtes de culture sécurité, a une influence significative sur les résultats et leur interprétation. La méthode d'agrégation par la moyenne des réponses positives, bien que recommandée par l'AHRQ, peut être considérée d'un œil critique. Considérant que les études comparant la culture de sécurité entre plusieurs établissements ou plusieurs pays sont

régulièrement publiées, nous pensons qu'un consensus quant au choix de la méthode d'agrégation est nécessaire, afin de garantir la validité et l'équité de ces comparaisons.

ISQua®

International Journal for Quality in Health Care, 2017, 1–8
doi: 10.1093/intqhc/mzx086
Article

OXFORD

Article

Hospital survey on patient safety culture (HSOPS): variability of scoring strategies

JORIS GIAI[1], BASTIEN BOUSSAT[2,3], PAULINE OCCELLI[4,5], GERALD GANDON[2], ARNAUD SEIGNEURIN[2,3], PHILIPPE MICHEL[4], and PATRICE FRANÇOIS[2,3]

[1]Service de Biostatistique, Hospices Civils de Lyon, Laboratoire de biométrie et biologie évolutive, UMR 5558 CNRS, F-69003 Lyon, France, [2]Quality of Care Unit, Grenoble University Hospital, F-38043 Grenoble, France, [3]TIMC UMR 5525 CNRS, Université Grenoble Alpes, France, [4]Quality and Patient Safety Department, Pôle IMER, University Hospital of Lyon, F-69003 Lyon, France, and [5]EA 7425 Health Services and performance Research (HESPER), Université Claude Bernard Lyon 1, France

Address reprint requests to: Bastien Boussat, Quality of Care Unit, Pavillon Taillefer, CHU CS 10-217, F-38043, Grenoble Cedex, France. Tel: +33-4-76-76-69 -96; Fax: +33-4-76-76-88-31; E-mail: bboussat@chu-grenoble.fr

Editorial Decision 1 July 2017; Accepted 3 July 2017

Abstract

Objective: To assess the variability of safety culture dimension scores and their associated rankings depending on three different scoring strategies using the Hospital Survey On Patient Safety Culture (HSOPS).
Design: Cross-sectional study using a self-administered questionnaire.
Setting: The study was conducted in an 1836-bed acute-care French university hospital with an annual volume of 135 999 stays, between April 2013 and November 2014.
Participants: All caregivers and technical-administrative staff with at least 6 months of employment, spending at least half of their working time in the hospital, were asked to participate.
Intervention: None.
Main outcome measure: The variability of the HSOPS results using three different scoring methods: the percentage of positive responses recommended by the Agency for Healthcare Research and Quality, the averaged individual means and the averaged individual sums.
Results: The response rate was 78.6% ($n = 3978$). The percentage of positive responses resulted in lower scores compared to averaged individual means and averaged individual sums in the six least developed dimensions, and gave more widely spread scores and greater 95CIs in the six most developed dimensions. Department rankings also varied greatly depending on the scoring methods.
Conclusion: The values of the HSOPS scores and their corresponding rankings greatly depended on the computation method. This finding shows how important it is to agree on the use of the same scoring strategies, before broadly comparing results within and across organizations.

Key words: patient safety, healthcare benchmarking, safety management, survey

Introduction

Patient safety culture, as defined by the Agency for Healthcare Research and Quality (AHRQ), is 'the product of individual and group values, attitudes, perceptions, competencies and patterns of behavior that determine the commitment to, and the style and proficiency of, an organization's health and safety management' [1].

© The Author 2017. Published by Oxford University Press in association with the International Society for Quality in Health Care. All rights reserved.
For permissions, please e-mail: journals.permissions@oup.com

The need to measure patient safety in healthcare appears central and has been the subject of numerous studies [2–6].

Numerous tools have been designed to measure safety culture in various designs and settings. Among them, the most widely used in hospital-wide designs is the Hospital Survey on Patient Safety Culture (HSOPS), set up in 2004 by the US Agency for Healthcare Research and Quality (AHRQ) [1]. The HSOPS is a self-administered questionnaire including 42 items used to calculate composite scores for 12 dimensions of safety culture. Although the main purpose of this questionnaire is the assessment of quality improvement, the results of the HSOPS could have substantial policy implications when used to conduct comparisons within and across organizations. In 2014, 653 US hospitals participated in the sixth nationwide HSOPS [7]. The managers received a feedback report comparing their results to the comparative database, helping them identify strengths and opportunities for improvement in their hospital's patient safety culture. In 2010, the French National Health Authority (Haute Autorité de Santé, HAS) introduced a new criterion relative to the patient safety culture in the national accreditation process of all French hospitals, demonstrating the growing importance of the survey results for healthcare decision-makers [8].

Using the HSOPS, many studies have reported culture safety variations among healthcare facilities, departments or occupational categories of healthcare workers in North America [9–12], Europe [13–22], Asia [23–27] and the Middle East [28–30]. Surprisingly, three different aggregation methods were used to compound dimension-level scores in these studies. Whatever the context or the purpose of the studies, the three methods were used without specific explanation Briefly, dimension scores were calculated (i) as recommended by AHRQ guidelines by dividing the total number of positive responses in the dimension by the total number of responses, (ii) by averaging individual means of the items within a dimension, or (iii) by averaging individual sums of the items within a dimension. See the Methods section for further details.

Several studies found substantial variations for HSOPS results between hospitals and between countries [7, 22, 27]. While HSOPS score differences undoubtedly reflected true variations in patient safety culture, the use of different aggregation rules for the composite scores could bias the comparisons. Although never considered for the HSOPS questionnaire, inconsistencies relating to the scoring methods have already been reported for the use of other composite scores in the field of health service performance studies [31, 32]. Knowing that comparison of HSOPS results could impact safety culture enhancement policies, the possible variability between the three scoring methods raises questions.

The main goal of this study was to assess the variability of safety culture dimension scores and their associated rankings depending on three different scoring strategies. We expect that our findings will guide researchers who plan to compare patient safety culture levels in their units or hospitals based on the most appropriate scoring method.

Methods

Study design

We conducted a cross-sectional study using a self-administered questionnaire. The project was submitted for advisory purposes to an independent protection committee (IRB00006705), and received its approval from the French National Commission for Information Technology and Civil Liberties (CNIL).

Setting

The study was conducted at a single university-affiliated hospital with a capacity of 1836 beds. The study site reported 135 999 stays in 2014. The hospital staff comprised 4422 registered paramedical staff and 642 board-certified physicians, the vast majority of them hospital specialists.

Participants

Eligible participants were full-time or part-time (half-time or more) employees with at least 6 months of employment in clinical, laboratory/pathology, radiology or pharmacy departments.

Data collection

The data collection was conducted anonymously on a volunteer basis, department by department, between April 2013 and September 2014.

Questionnaire and variables

The French version of the HSOPS was used to assess the level of safety culture in our setting. The present analysis takes into account the 42 items covering the 12 dimensions that were included in the original HSOPS questionnaire. Eighteen items were negatively worded

Each item was answered on a five-point Likert scale, from 'Strongly disagree' (1 point) to 'Strongly agree' (5 points), or from 'Never' to 'Always' when relevant. A global safety grade between 'poor' and 'excellent' and the numbers of reported incidents in the past 12 months were also assessed but were not used to compute any of the 12 dimension scores.

Scoring methods

Before analysis, the coding of negatively worded items were reversed so that high-value answers would always be associated with a more highly developed safety culture. Then three methods were compared: (1) the percentage of positive responses, (2) the averaged individual means and (3) the averaged individual sums.

1) Percentage of positive responses (M1)

The first dimension score computation method was the percentage of positive answers defined as values of 4 ('Agree') or 5 ('Strongly agree'). Following AHRQ guidelines, the first step was to obtain, for each item, the percentage of respondents who answered it positively among non-missing answers [1]. Then unweighted averages of those percentages were computed for each dimension. By definition this method was a mean of percentages, hence resulting in dimension scores ranging from zero to 100.

2) Averaged individual means (M2)

The second method first required the case-wise computation of individual means across the three or four items in a dimension, followed by an averaging within dimensions. Given that each individual item response ranged from one to five, M2 dimension scores varied to the same extent.

3) Averaged individual sums (M3)

The last method we found was only present in three studies [23, 24, 27]. First, individual sums were assessed over dimensions with missing

responses accounting for zero. Those were then averaged within dimensions. Thus, three-item dimensions scores ranged from 1 to 15, whereas four-items dimension scores ranged from 1 to 20.

Statistical analysis

The usual descriptive statistics were used to describe the sample characteristics. Qualitative variables were reported as numbers and percentages, while quantitative variables were reported as means with standard deviations or medians with the range.

Dimension scores were assessed with M1, M2 and M3 within each of the 14 departments. Rankings were then obtained for each dimension and each scoring method by simply ordering the computed scores, with the first rank attributed to the highest score. Absolute rank differences were calculated for each department in each dimension as the absolute value of the difference between its rank using a method and its rank using another method. The sum of such absolute rank differences between two methods x and y across all departments is noted $SARD_{x,y}$. These SARDs can range from zero (no difference in ranks between the two methods) to trunc$\left(\frac{c^2}{2}\right)$ (the two methods give opposite rankings) for a variable with c categories. Dimension scores, rankings and SARDs were also computed for the other respondent characteristics included, for illustrative purposes.

For comparability, dimension scores obtained with the M2 and M3 methods were rescaled in order to vary from zero to 100. Since this was done using linear transformations only, they did not impair the ranking process.

Seeking a single way to calculate 95% confidence intervals (95CIs) for dimension scores of all three scoring strategies, we estimated them using the percentile bootstrap method from 5000 samples. Bootstrap estimators of M2 and M3 95CIs where compared with those obtained from normal distribution approximation.

Analyses were performed using Stata software release 11 (StataCorp. 2009. College Station, TX, USA: StataCorp LP).

Results

Respondent characteristics

The overall response rate was 78.6% ($n = 3978$). After the exclusion of 90 questionnaires due to the discovery of exclusion criteria (Fig. 1), the analytical sample consisted of 3888 survey questionnaires. Females accounted for 80.6% of the responses (Table 1).

Scores variability

M2 and M3 strategies globally attributed higher dimension scores than M1. According to M1, the two most developed dimensions were 'teamwork within hospital units' (M1 = 62.7%; M2 = 63.4% and M3 = 65.2%) and 'communication openness' (M1 = 59.4%; M2 = 62.4% and M3 = 62.9%), while the two least developed dimensions were 'hospital handoffs and transitions' (M1 = 25.4%; M2 = 46.9% and M3 = 51.7%) and 'hospital management support' (M1 = 20.2%; M2 = 41.7% and M3 = 46.1%).

In all 12 dimensions, score values varied across scoring methods for a given department. For instance, in the most developed dimension, the score of the worst department (F) increased by more than 10% from M1 to M3, increasing from 46.9% to 57.7% (Fig. 2). In contrast, for the best ranked departments in the same dimension, scores tended to decrease from M1 to M3 on a smaller scale: department G for example decreased its score from 72.4% to 70.2%. This pattern of variability reduction was also found in five other dimensions (see additional figures in Supplemental Digital Content A, which show caterpillar-type plots with M1, M2 and M3 department score point estimates and bootstrap 95CIs for all dimensions). Notably, these six dimensions were the most developed ones in the

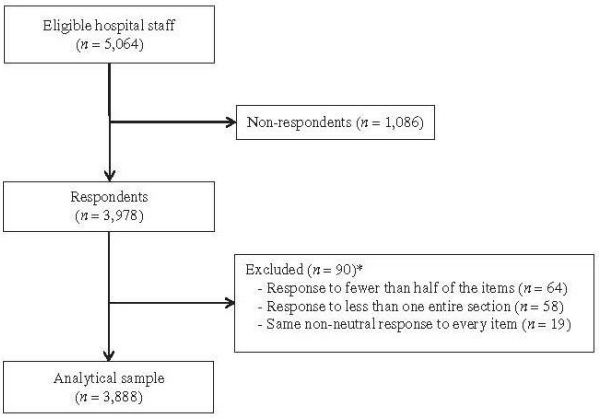


Figure 1 Survey participation.

Table 1 Characteristics of the participants ($n = 3888$)

Characteristics	*N* (%)
Female	3016 (80.6)
Age classes (years)	
Up to 35	1406 (37.5)
36–45	966 (25.7)
46–55	967 (25.7)
56 or older	415 (11.1)
Seniority in current unit (years)	
Up to 1	460 (12.1)
1–2	652 (17.2)
3–5	982 (25.9)
6 or more	1698 (44.8)
Healthcare occupation	2665 (69.9)
Nurse	1323 (34.7)
Nursing assistant	706 (18.5)
Physician	457 (12.0)
Other healthcare	179 (4.7)
Non-healthcare occupation	1148 (30.1)
Technical	775 (20.3)
Administrative	347 (9.1)
Other non-healthcare	26 (0.7)
Department	
Couple–children	591 (15.2)
Orthopedics, traumatology	466 (12.0)
Multidisciplinary medicine	415 (10.7)
Anesthesia, reanimation	337 (8.7)
Thorax, vessels	310 (8.0)
Biology	303 (7.8)
Emergency, acute medicine	302 (7.8)
Gastroenterology, uronephrology	300 (7.7)
Psychiatry, neurology	266 (6.9)
Oncology	158 (4.1)
Operating rooms	153 (3.9)
Imaging	150 (3.9)
Pharmacy	97 (2.5)
Management, home care	32 (0.8)

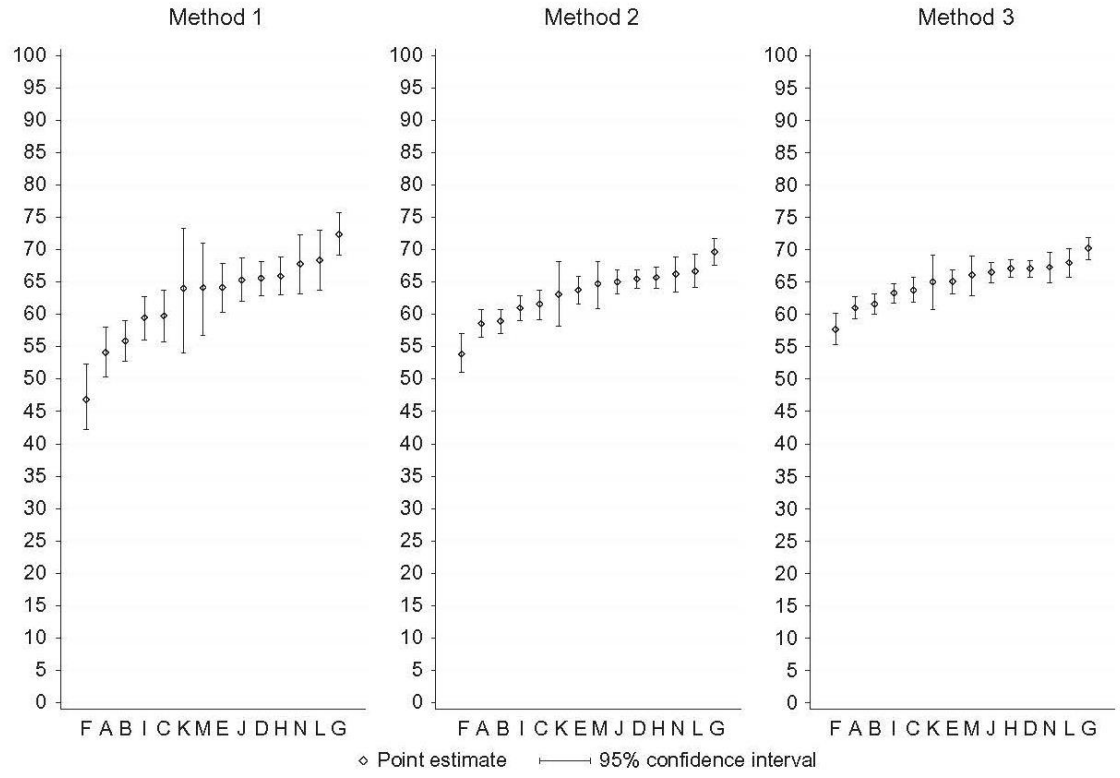


Figure 2 'Teamwork within hospital units' dimension scores. *X*-axis: ordered de-identified departments for each method. *Y*-axis: rescaled score.

study according to M1 scores from the AHRQ, all above the median of 42%.

In the other six dimensions, M1 scores were always lower than their M2 and M3 counterparts. For example, in the least developed 'hospital management support' dimension, the last department according to M1 more than tripled its score, increasing from 14.6% with M1 to 47.5% with M3 (Fig. 3).

Ranking variability

The greatest department ranking discordances between methods in terms of SARD were found in the 'hospital handoffs and transitions' dimension (Table 2). As depicted, SARDs could reach moderately high values with $SARD_{1,2} = 48$, $SARD_{1,3} = 48$ and $SARD_{2,3} = 4$. The closeness shown between M2 and M3 rankings, compared to M1 and M2 or M1 and M3 was observed in all dimensions (see additional tables in Supplemental Digital Content B, which show subgroup scores using M1, M2 and M3 with their associated ranks and SARD for all dimensions).

As for score values, variability patterns were different for the six most developed dimensions and the six least developed dimensions. In the six most developed dimensions cited above, $SARD_{1,2}$ for departments ranged from 2 in 'teamwork within hospital units' to 20 in 'frequency of event reporting,' with a median of 12 (mean = 11.0). Although for the six least developed dimensions these sums were much higher and increased from 12 in 'staffing' to 48 in 'hospital management support,' with a median of 14 (mean = 24.7).

Finally, bootstrap estimations of 95CIs with the percentile method were consistent with normal approximation estimations for M2 and M3 (data not shown). In all dimensions, M1 always had the widest 95CIs while they were quite similar in width between M2 and M3. Moreover, the 95CIs of the scores were almost systematically overlapping within each scoring method. Thus, few departments could be confidently placed in the top or bottom quartiles of the league table.

Discussion

This study highlights the heterogeneity of the results obtained by the three scoring methods used to assess levels of perceived safety culture from the HSOPS, showing that dimensional score values as well as their associated rankings could vary substantially across these methods. In particular, the recommended method, M1, gave lower scores in the six least developed dimensions and overall varied more. In conjunction with the overlapping 95CIs, this led to substantial ranking variations across methods for the 14 departments.

We found seven studies conducted in various settings (three in Europe, three in the Middle East, one in the USA) that computed scores using M1 and M2 [12–14, 18, 28–30]. Many discrepancies between M1 and M2 could be observed: for instance, in the study reported by El-Jardali *et al.* [29], who analyzed a comparable number of questionnaires ($n = 2572$) in a similar setting, the 'overall perception of safety' dimension, which ranked fourth with M1, dropped to the eighth place with M2, while the 'frequency of event reporting' dimension, which ranked eighth, rose to fifth place.

Unweighted means of M1, M2 and M3 dimensional scores along with several summary statistics at a continental level are presented in the Supplemental Digital Content C. Interestingly, the same pattern as in our study appeared. In all studies, greater variability was noted in the six most developed dimensions according to M1, while in the six others it attributed lower dimensional scores [13, 14, 18,

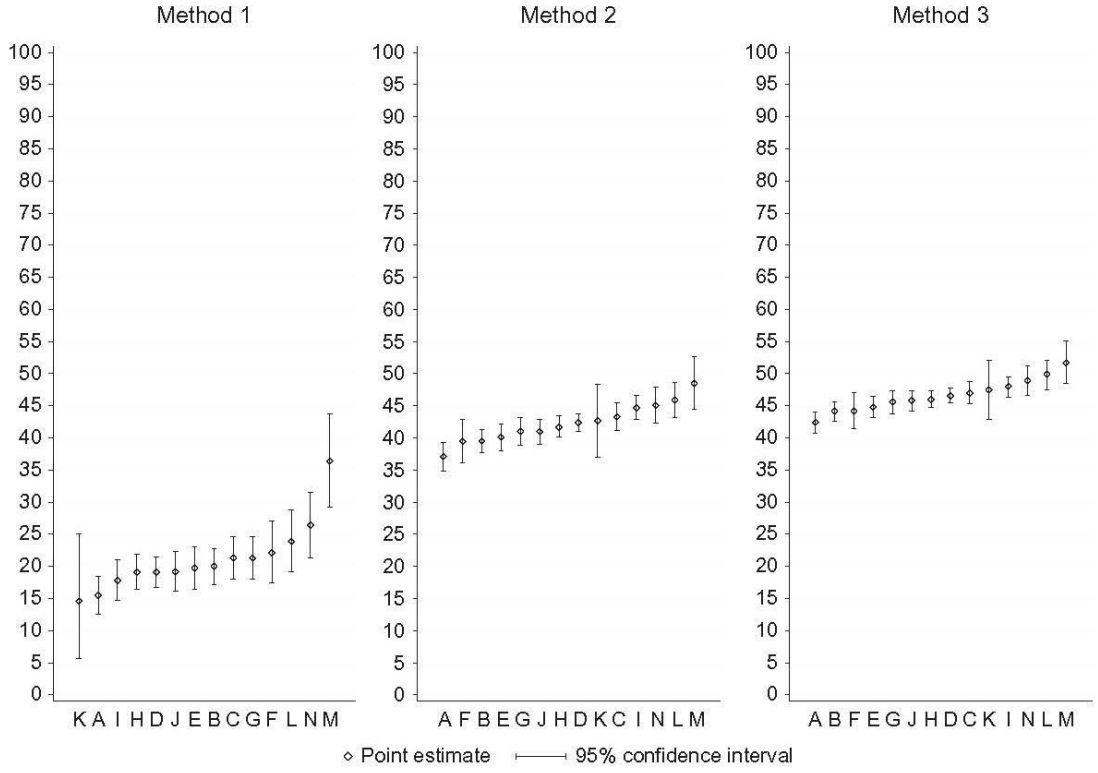


Figure 3 'Hospital management support' dimension scores for each department. *X*-axis: ordered de-identified departments for each method. *Y*-axis: rescaled score.

28, 29]. As an example from the study of Ballangrud *et al.* [14], even though the 'frequency of event reporting' dimension was the least developed one according to both M1 and M2, its associated scores were 18% and 42.5%, respectively. As shown in the Supplemental Digital Content C, this remained true for continental scores as well, despite the diversity of study designs and settings.

Variations in score values and rankings can be explained by the specific properties of the three scoring methods. The main discrepancies were found when comparing M1 with M2 and M1 with M3. M1 is a percentage of positive responses obtained by a prior dichotomization of item responses. This process undoubtedly leads to a loss of information and consequently an increase in variance, which could help explain the lower accuracy illustrated by the widest 95CI. Neutral responses are considered as non-positive, which could provide lower scores for the least developed dimensions. Both M2 and M3 do not transform item responses before computing the scores. Indeed, M2 provides an averaged individual mean and M3 an averaged individual sum of all the items in each dimension. Those two methods preserve all the information from the items and consequently yield more accurate dimensions scores, with narrower 95CIs.

Even if close results between M2 and M3 were found, several arguments tip the scale in favor of M2. M3 is found in only three studies, whereas individual means are already widely used. This method is recommended and validated for most other self-administered questionnaires, in the same field [33] or in other similar depression scales [34]. Finally, the psychometric properties of the original HSOPS questionnaire and most of their translated versions were validated using M2 [1, 16, 17, 21, 24, 35].

Although the HSOPS questionnaire can be used for various purposes, several studies compared HSOPS results between hospitals and between countries. In 2014, 405 281 hospital staff respondents from 653 US hospitals participated in the sixth nationwide HSOPS questionnaire-based survey [7]. This database identifies areas for internal improvement for each hospital and also provides hospital rankings using M1. In 2013, Fujita *et al.* compared HSOPS results between Japan, Taiwan and the USA using M1 and M3 [23]. The development of such benchmarking studies demonstrates the need to agree on the use of the same scoring strategies. These comparisons may impact the policy-making process, including external evaluation such as accreditation, purchasing or incentive policies. Indeed, the ranking of hospitals based on performance composite indicators has become common worldwide. In Great Britain, the National Health Service (NHS) disseminated league tables, rating hospitals according a composite score calculated with 47 indicators, to the public. In 2016, the NHS published the 'Learning from mistakes league' based on a nationwide hospital staff survey, and produced a ranking for 230 hospitals [36]. In this survey, the 30 last hospitals were publicly noted as having a 'Poor reporting culture.' If similar studies used the HSOPS as a methodology, the impact of the variability of the scoring method might have a huge impact for the hospitals standing at the bottom of the league. In the USA, the National Committee for Quality Assurance (NCQA) recently developed comparable hospital rankings based on composite scores up to 100, which reflect member satisfaction and success in preventing and treating illness compared with other Medicare plans [37]. Parallel to this, the Centers for Medicare & Medicaid Services (CMS) and the Joint Commission on the Accreditation of Healthcare Organizations (JCAHO)

Table 2 'Hospital handoffs and transitions' subgroup scores using M1[a], M2[b] and M3[c] with their associated ranks and sums of absolute rank differences (SARD)

Variable	*N* (%)	Scores			Ranks			Sums of absolute rank differences (SARD)		
		M1[a]	M2[b]	M3[c]	M1[a]	M2[b]	M3[c]	$SARD_{1,2}$	$SARD_{1,3}$	$SARD_{2,3}$
Age classes (years)								2	2	4
45–54	967 (25.8)	27.0	47.7	52.1	1	1	2			
<35	1406 (37.5)	26.3	47.1	52.2	2	2	1			
35–44	966 (25.7)	24.0	46.1	51.2	3	4	3			
≥55	415 (11.1)	22.2	46.6	50.8	4	3	4			
Occupation								14	12	4
Nursing assistant	706 (18.5)	34.7	48.9	53.6	1	1	1			
Nurse	1323 (34.7)	29.0	46.6	51.7	2	3	3			
Other non-healthcare	26 (0.7)	24.1	45.1	48.2	3	6	7			
Other healthcare	179 (4.7)	23.7	45.7	51.3	4	5	4			
Physician	457 (12.0)	23.2	45.0	49.9	5	7	6			
Technical	775 (20.3)	19.2	47.6	51.8	6	2	2			
Administrative	347 (9.1)	11.4	46.2	51.2	7	4	5			
Department								36	34	14
G	310 (8.0)	31.4	47.1	52.4	1	6	4			
D	591 (15.2)	29.0	48.5	53.0	2	1	2			
F	153 (3.9)	28.7	47.6	52.3	3	4	6			
B	415 (10.7)	28.6	46.8	51.9	4	8	8			
H	466 (12.0)	28.5	47.1	52.2	5	7	7			
J	337 (8.7)	28.4	48.3	52.8	6	2	3			
E	300 (7.7)	26.7	46.0	51.3	7	9	9			
N	158 (4.1)	25.9	47.9	53.2	8	3	1			
A	302 (7.8)	24.8	47.5	52.3	9	5	5			
C	266 (6.9)	21.9	45.3	50.4	10	11	10			
K	32 (0.8)	21.1	41.2	47.3	11	14	14			
M	97 (2.5)	14.4	44.6	48.7	12	12	12			
L	150 (3.9)	13.6	43.7	48.8	13	13	11			
I	303 (7.8)	9.4	45.5	48.5	14	10	13			
Gender								2	2	0
Female	3016 (80.6)	25.6	46.9	51.7	1	2	2			
Male	725 (19.4)	24.6	47.2	52.1	2	1	1			
Speciality								2	0	2
Home care, nursing home, follow-up care	378 (9.7)	32.4	48.5	53.5	1	2	1			
Anesthesia, reanimation, emergency	771 (19.8)	28.8	48.8	53.3	2	1	2			
Surgery, gynecology, obstetrics	990 (25.5)	28.6	47.1	52.1	3	3	3			
Medicine, psychatry	1163 (29.9)	24.9	45.8	51.1	4	4	4			
Biology, imaging, pharmacy	586 (15.1)	11.5	45.2	49.0	5	5	5			

Ordered according to the AHRQ M1 score, $n = 3888$.

[a]M1: Percent positive scores.

[b]M2: Average individual mean.

[c]M3: Average individual sum.

produced several rankings based on composite indicators publicly disseminated on their website as 'Hospital Compare' [38] and 'Quality Check' [39]. One part of the CMS ranking is produced with the results of the composite score of the Survey of patients' experiences (HCAHPS). The use of similar ranking based on the HSOPS may also be considered by similar organizations. In France, the HAS has produced several patient safety indicators for many years and published the hospitals' ranking based on the composite scores on a public website. Moreover, the results of these indicators were recently included in the financial incentive program for quality improvement in healthcare, implemented in 2016 [40]. For the moment, the HSOPS results have not been included in this program, but we can presume a future use, given the promotion of the HSOPS and its use in the French hospitals' national accreditation process [8]. Overall, given our findings on the huge variability according to the scoring methods, healthcare decision-makers should consider development of similar ranking based on the HSOPS with great caution.

This study has several limitations. First, participants were included on a volunteer basis and even with high answer rates this may have created a selection bias. However, it is very unlikely that the extent of this bias could have been heterogeneous among departments. Moreover, this inclusion method has been used by every HSOPS-related study so far. Second, this survey was monocentric. Although the results might apply less broadly, they are still valuable at least for similar settings. As we estimated the dimensional scores for the three strategies on the same sample, this should not challenge the corresponding findings to a significant extent. Third, our study was cross-sectional and therefore could not assess the impact of scoring methods on the trending in patient safety culture over time.

The percentage of positive responses method might be best suited for this purpose, considering the widest range of M1 scores for the 12 dimensions observed in the survey. However, only another survey with a longitudinal design can address this question adequately.

Conclusion

The values of the HSOPS scores and their corresponding rankings greatly depended on the computation method used. This finding shows that agreement must be reached on the use of the same scoring strategies, before broadly comparing results within and across organizations. Overall, the variability of the three scoring methods raises questions about the usefulness of the HSOPS results.

Supplementary material

Supplementary material is available at *International Journal for Quality in Health Care* online.

Acknowledgements

This work was developed within the framework of the Grenoble Alpes Data Institute, which is supported by the French National Research Agency under the 'Investissements d'avenir' program (ANR-15-IDEX-02). We acknowledge Linda Northrup for her assistance in editing the manuscript in English.

Funding

This work was supported by a grant from the French Ministry of Health (Programme de recherche sur la performance du système de soins-PREPS 12-026-0001)

References

1. Sorra J, Nieva VF. *Hospital Survey on Patient Safety Culture*. Rockville, MD: Agency for Healthcare Research and Quality, 2004. (Prepared by Westat, under Contract No. 290-96-0004). AHRQ Publication No. 04-0041.
2. An J, Kim SJ, Park S *et al.* The effects of patient education on patient safety: can we change patient perceptions and attitudes? Lessons from the Armed Forces Capital Hospital in Korea. *Int J Qual Health Care* 2017. doi:10.1093/intqhc/mzx037.
3. Mallouli M, Tlili MA, Aouicha W *et al.* Assessing patient safety culture in Tunisian operating rooms: a multicenter study. *Int J Qual Health Care* 2017;29:176–82. doi:10.1093/intqhc/mzw157.
4. Tereanu C, Smith SA, Sampietro G *et al.* Experimenting the hospital survey on patient safety culture in prevention facilities in Italy: psychometric properties. *Int J Qual Health Care* 2017;29:269–75. doi:10.1093/intqhc/mzx014.
5. Tricarico P, Castriotta L, Battistella C *et al.* Professional attitudes toward incident reporting: can we measure and compare improvements in patient safety culture? *Int J Qual Health Care* 2017;29:243–9. doi:10.1093/intqhc/mzx004.
6. Zecevic AA, Li AH, Ngo C *et al.* Improving safety culture in hospitals: facilitators and barriers to implementation of Systemic Falls Investigative Method (SFIM). *Int J Qual Health Care* 2017. doi:10.1093/intqhc/mzx034.
7. Sorra J, Famolaro T, Dyer N *et al. Hospital Survey on Patient Safety Culture: 2014 User Comparative Database Report* http://www.ahrq.gov/sites/default/files/wysiwyg/professionals/quality-patient-safety/patientsafetyculture/hospital/2014/hsops14pt1.pdf (2014, date last accessed).
8. Occelli P. La culture de sécurité des soins: du concept à la pratique https://www.has-sante.fr/portail/upload/docs/application/pdf/2011-02/culture_de_securite_des_soins__du_concept_a_la_pratique.pdf (2010, date last accessed).
9. Bowman C, Neeman N, Sehgal NL. Enculturation of unsafe attitudes and behaviors: student perceptions of safety culture. *Acad Med* 2013;88:802–10. doi:10.1097/ACM.0b013e31828fd4f4.
10. Jasti H, Sheth H, Verrico M *et al.* Assessing patient safety culture of internal medicine house staff in an academic teaching hospital. *J Grad Med Educ* 2009;1:139–45. doi:10.4300/01.01.0023.
11. Jones KJ, Skinner AM, High R *et al.* A theory-driven, longitudinal evaluation of the impact of team training on safety culture in 24 hospitals. *BMJ Qual Saf* 2013;22:394–404. doi:10.1136/bmjqs-2012-000939.
12. Sorra JS, Dyer N. Multilevel psychometric properties of the AHRQ hospital survey on patient safety culture. *BMC Health Serv Res* 2010;10:199. doi:10.1186/1472-6963-10-199.
13. Agnew C, Flin R, Mearns K. Patient safety climate and worker safety behaviours in acute hospitals in Scotland. *J Safety Res* 2013;45:95–101. doi:10.1016/j.jsr.2013.01.008.
14. Ballangrud R, Hedelin B, Hall-Lord ML. Nurses' perceptions of patient safety climate in intensive care units: a cross-sectional study. *Intensive Crit Care Nurs* 2012;28:344–54. doi:10.1016/j.iccn.2012.01.001.
15. Bodur S, Filiz E. A survey on patient safety culture in primary healthcare services in Turkey. *Int J Qual Health Care* 2009;21:348–55. doi:10.1093/intqhc/mzp035.
16. Hammer A, Ernstmann N, Ommen O *et al.* Psychometric properties of the Hospital Survey on Patient Safety Culture for hospital management (HSOPS_M). *BMC Health Serv Res* 2011;11:165. doi:10.1186/1472-6963-11-165.
17. Occelli P, Quenon JL, Kret M *et al.* Validation of the French version of the Hospital Survey on Patient Safety Culture questionnaire. *Int J Qual Health Care* 2013;25:459–68. doi:10.1093/intqhc/mzt047.
18. Robida A. Hospital Survey on Patient Safety Culture in Slovenia: a psychometric evaluation. *Int J Qual Health Care* 2013;25:469–75. doi:10.1093/intqhc/mzt040.
19. van Noord I, de Bruijne MC, Twisk JW. The relationship between patient safety culture and the implementation of organizational patient safety defences at emergency departments. *Int J Qual Health Care* 2010;22:162–9. doi:10.1093/intqhc/mzq013.
20. Verbeek-Van Noord I, Wagner C, Van Dyck C *et al.* Is culture associated with patient safety in the emergency department? A study of staff perspectives. *Int J Qual Health Care* 2014;26:64–70. doi:10.1093/intqhc/mzt087.
21. Vlayen A, Hellings J, Claes N *et al.* Measuring safety culture in belgian psychiatric hospitals: validation of the dutch and French translations of the hospital survey on patient safety culture. *J Psychiatr Pract* 2015;21:124–39. doi:10.1097/01.pra.0000462605.17725.48.
22. Wagner C, Smits M, Sorra J *et al.* Assessing patient safety culture in hospitals across countries. *Int J Qual Health Care* 2013;25:213–21. doi:10.1093/intqhc/mzt024.
23. Fujita S, Seto K, Ito S *et al.* The characteristics of patient safety culture in Japan, Taiwan and the United States. *BMC Health Serv Res* 2013;13:20. doi:10.1186/1472-6963-13-20.
24. Ito S, Seto K, Kigawa M *et al.* Development and applicability of Hospital Survey on Patient Safety Culture (HSOPS) in Japan. *BMC Health Serv Res* 2011;11:28. doi:10.1186/1472-6963-11-28.
25. Liu C, Liu W, Wang Y *et al.* Patient safety culture in China: a case study in an outpatient setting in Beijing. *BMJ Qual Saf* 2014;23:556–64. doi:10.1136/bmjqs-2013-002172.
26. Nie Y, Mao X, Cui H *et al.* Hospital survey on patient safety culture in China. *BMC Health Serv Res* 2013;13:228. doi:10.1186/1472-6963-13-228.
27. Wu Y, Fujita S, Seto K *et al.* The impact of nurse working hours on patient safety culture: a cross-national survey including Japan, the United States and Chinese Taiwan using the Hospital Survey on Patient Safety Culture. *BMC Health Serv Res* 2013;13:394. doi:10.1186/1472-6963-13-394.
28. El-Jardali F, Jaafar M, Dimassi H *et al.* The current state of patient safety culture in Lebanese hospitals: a study at baseline. *Int J Qual Health Care* 2010;22:386–95. doi:10.1093/intqhc/mzq047.

29. El-Jardali F, Sheikh F, Garcia NA *et al.* Patient safety culture in a large teaching hospital in Riyadh: baseline assessment, comparative analysis and opportunities for improvement. *BMC Health Serv Res* 2014;14:122. doi:10.1186/1472-6963-14-122.
30. Hamdan M, Saleem AA. Assessment of patient safety culture in Palestinian public hospitals. *Int J Qual Health Care* 2013;25:167–75. doi:10.1093/intqhc/mzt007.
31. Minvielle E. Analyse critique du développement d'indicateurs composites: le cas de l'infarctus du myocarde après la phase aiguë In: Haute Autorité de Santé; 2010.
32. Jacobs R, Goddard M, Smith PC. How robust are hospital ranks based on composite performance measures? *Med Care* 2005;43:1177–84.
33. Pronovost PJ, Weast B, Holzmueller CG *et al.* Evaluation of the culture of safety: survey of clinicians and managers in an academic medical center. *Qual Saf Health Care* 2003;12:405–10.
34. Zung WW. A Self-Rating Depression Scale. *Arch Gen Psychiatry* 1965; 12:63–70.
35. Perneger TV, Staines A, Kundig F. Internal consistency, factor structure and construct validity of the French version of the Hospital Survey on Patient Safety Culture. *BMJ Qual Saf* 2014;23:389–97. doi:10.1136/bmjqs-2013-002024.
36. Transparency data. Learning from mistakes league https://www.gov.uk/government/publications/learning-from-mistakes-league (2016, date last accessed).
37. Shetty P, Rivas R, Hristidis V. Correlating ratings of health insurance plans to their providers' attributes. *J Med Internet Res* 2016;18:e279. doi:10.2196/jmir.6475.
38. Hospital Compare. https://www.medicare.gov/hospitalcompare/about/what-is-HOS.html (2017, date last accessed).
39. Quality Check. https://www.jointcommission.org/facts_about_quality_check_and_quality_reports/ (2017, date last accessed).
40. Incitation financière à l'amélioration de la qualité (IFAQ) https://www.has-sante.fr/portail/jcms/c_493937/fr/incitation-financiere-a-l-amelioration-de-la-qualite-ifaq-generalisation (2016, date last accessed).

3.3.3. Méthodes d'imputation des données manquantes : article soumis – Dealing with missing data in the Hospital Survey on Patient Safety Culture: a simulation study

La question de la gestion des données manquantes n'a jamais été évaluée pour le questionnaire HSOPS, aboutissant à une hétérogénéité de choix méthodologiques dans la littérature (Smits, Christiaans-Dingelhoff et al. 2008, Hammer, Ernstmann et al. 2011, Ito, Seto et al. 2011, Feng, Bobay et al. 2012, Vlayen, Schrooten et al. 2015). En effet, nous avons retrouvé plusieurs études publiées utilisant, sans justification, des méthodes d'imputations par la moyenne, ou d'imputations multiples. Considérant que les données manquantes, et les méthodes d'imputation utilisées peuvent avoir des conséquences sur l'interprétation des résultats des enquêtes de culture sécurité, nous avons effectué une étude de simulation à partir des questionnaires complets de notre base de données (questionnaires ne comportant aucune donnée manquante sur les 42 items). Les résultats sont présentés dans l'article suivant (soumis).

Objectif : L'objectif de cette étude était de comparer les performances des méthodes d'imputation des données manquantes dans le cadre d'enquêtes de culture sécurité effectuées à partir du questionnaire HSOPS.

Méthode : Utilisant les données produites à partir de l'enquête de culture de sécurité menée au CHU de Grenoble, nous avons simulé des données sur deux niveaux. Dans un premier temps, nous avons simulé 1000 bases de données en utilisant une procédure de tirage au sort avec remise à partir de la base de données des 3045 questionnaires comportant des réponses complètes aux 42 items du HSOPS. Dans un second temps, nous avons dégradé les 1000 bases de données complètes selon 6 scénarios de données manquantes, obtenant ainsi 6000 bases de données comportant des données manquantes. Les scénarios choisis étaient 1/ manquant complètement aléatoirement avec 10% de données manquantes MCAR10, 2/ MCAR20, 3/ MCAR30, 4/ MCAR50, 5/ manquant aléatoirement (en assignant plus de probabilité de données manquantes aux répondant de sexe féminin) MAR et 6/ manquant non aléatoirement (en

assignant plus de probabilité de donnée manquante aux répondants ayant un pourcentage de réponse positive inférieur à 50%) MNAR. Ensuite, pour chacune des 6000 bases, les données manquantes étaient imputées selon 5 méthodes d'imputations : 1/ par une sélection aléatoire (chiffre compris entre 1 et 5), 2/ par la moyenne de réponse à l'item, 3/ par la moyenne de réponse de l'individu, 4/ par une méthode d'imputation multiple (predictive matching method) et 5/ par une méthode de factorisation de matrice utilisée fréquemment en machine learning (sparse non negative matrix factorization, sNMF). Les performances pour chaque méthode d'imputation étaient évaluées selon deux niveaux de biais, pour chacun des 6 scénarios de données manquantes. Nous avons tout d'abord analysé les erreurs quadratiques moyennes (Root Mean Square Error, RMSE), montrant le biais entre les réponses aux items imputées et leurs valeurs réelles dans les bases complètes. Les RMSE étaient également produits par valeur de réponse à l'item, montrant les performances des 5 méthodes pour imputer les valeurs 1, 2, 3, 4 et 5 de l'échelle de Likert. Enfin, nous avons analysé les écarts en termes de pourcentages de réponses positives pour chacune des 12 dimensions de culture de sécurité.

Résultats : La méthode d'imputation par sélection aléatoire produisait sans surprise les pires résultats en termes de RMSE. Les quatre autres méthodes produisaient des RMSE proches, avec tout de même un avantage pour l'imputation multiple. Globalement, toutes les méthodes d'imputations avaient de meilleures performances pour imputer les réponses de valeur 3, alors que les pires performances étaient observées pour l'imputation des réponses extrêmes, 1 et 5. Les différences entre les méthodes étaient plus importantes au regard des RMSE spécifiques de ces valeurs extrêmes, avec un très net avantage pour l'imputation multiple. En termes de scores, les deux méthodes d'imputation par la moyenne produisaient les pourcentages de réponses positives les plus éloignés de ceux calculés à partir des 1000 bases complètes, pour les 12 dimensions, alors que l'imputation multiple et la sNMF avaient de bien meilleurs performances, pour les 6 scénarios. Par exemple, pour le scénario MCAR50, les biais de scores variaient entre

(-20.8 ; +5.9) et (-19.8 ; +2.7) pour l'imputation par la moyenne des items et l'imputation par la moyenne de réponses individuelles, respectivement. Alors que pour ce même scénario, les biais étaient plus faibles pour l'imputation multiple (-6.3 ; -0.8) et la méthode sNMF (-2.1 ; +6.5).

Conclusions : Au travers de cette large étude de simulation, nous avons mis en évidence que l'imputation multiple avait d'excellentes performances pour imputer les données manquantes du questionnaire HSOPS, et ce même pour des pourcentages très élevés de délétion. Bien que l'utilisation de l'imputation multiple soit habituellement réservée aux scénarios « manquant aléatoirement » MAR, les performances de cette méthode étaient bonnes pour l'ensemble des scénarios simulés dans notre étude. Nos résultats sont cohérents avec les travaux de Faris qui démontrait de façon empirique les bonnes performances de l'imputation multiple pour imputer des données épidémiologiques, quel que soit le scénario utilisé (Faris, Ghali et al. 2002). L'autre résultat principal de ce travail concerne les mauvaises performances des techniques d'imputation par la moyenne, en termes de scores dimensionnels. Compte-tenu des écarts observés entre les scores après imputation et les scores des bases complètes, l'imputation par la moyenne, pourtant utilisée dans certaines études de culture sécurité publiées, devrait être évitée lors des enquêtes réalisées à partir du questionnaire HSOPS.

Dealing with missing data in the Hospital Survey on Patient Safety Culture: a simulation study

Bastien Boussat[1,2], Julien Viotti[1], Olivier François[2], Arnaud Seigneurin[1,2], Joris Giai[3], José Labarère[1,2,4], Patrice François[1,2]

ABSTRACT

Background: The handling of missing data has never been assessed for the Hospital Survey on Patient Safety culture (HSOPS).

Objectives: Through a computer-intensive simulation study, we aimed to evaluate the accuracy of imputation methods to manage missing data in the HSOPS

Methods: Using the original data from a cross-sectional survey of 5,064 employees at a single university hospital in France, we produced simulation data on two levels. First, we simulated 1,000 completed data based on the original 3,045 complete responses using a bootstrap procedure. Second, missing values were simulated in these 1,000 complete cases data for comparison purposes, using 6 different missing data scenarios. Third, missing values were imputed using 5 different imputation methods (1/ random selection, 2/ item mean, 3/ individual mean, 4/ multiple imputation, 5/ sparse nonnegative matrix factorization). The performances of each imputation method were assessed using root mean square error (RMSE) and dimension scores bias.

Results: The five imputation methods yielded close RMSE, with an advantage for the multiple imputation. The bias differences were higher regarding the dimension scores, with a clear advantage for the multiple imputation. The worst performances were achieved by the mean imputation methods.

Discussion and conclusions: We recommend the use of multiple imputation to handle missing data in HSOPS-based surveys, while mean imputation methods should be avoided. Overall, our results may suggest the possibility to optimize the HSOPS instrument, which should be reduced without loss of global information.

Keywords: Patient safety, Safety Culture, Missing data, Multiple imputation.

INTRODUCTION

Self-report instruments are widely used in quality and safety research programs [1 2]. Depending on the properties of the instrument used, missing responses often represent a pronounced problem hard to handle without clear and detailed guidelines [3]. Among these instruments, the Hospital Survey on Patient Safety Culture (HSOPS) is designed to evaluate the patient safety culture in the health services. This self-administered questionnaire, founded by the Agency of Healthcare Research and Quality (AHRQ), include 42 questions rated on a five-point Likert scale of agreement used to compute 12 dimensions scores [4]. These dimensions scores are calculated through a custom aggregation method, i.e. dividing the number of positives responses on all dimensions' items (numerator) by the total number of responses to all dimensions' items (denominator), thus not taking missing data into consideration. Beside the fact that missing data exclusion could reduce data amount and study power, it can also promote biases in analysis and impact the accuracy of the calculated scores [5-7].

Different methods widely used in the epidemiologic research field were designed to manage missing values. Given the specificity of the data produced from self-report instruments, some studies have tried to re-evaluate the classic imputation methods [8 9]. In a simulation study based on a depression scale instrument, Shrive et al. concluded that multiple imputation was the most accurate method but individual's mean imputation was also appropriate [10]. This methodological aspect has never been assessed for the HSOPS questionnaire, while some studies used mean imputation or multiple imputation before assessing the psychometric properties of the instrument or the factors associated with patient safety culture [11-15].

Given that the assumption scenario of missing data has an impact on the handling method, we produced 6 missing data simulation from a complete dataset obtained in a patient safety culture survey. We then compared 5 methodological approaches of missing data imputation and evaluated the accuracy of each of them.

The objective of this study was to evaluate the accuracy of some of the most used imputations methods to manage missing data in the HSOPS.

METHODS

Study Design and data collection

This study was performed on a HSOPS conducted at a single university-affiliated hospital with a capacity of 1836 beds and serving a population of 675,000 inhabitants in France. The HSOPS was conducted anonymously on a volunteer basis between April 2013 and September 2014. Eligible participants were full-time or part-time employees with at least 6 months of employment in clinical, laboratory/pathology, radiology, or pharmacy departments. Of 5,064 eligible employees, 3,978 (78.6%) participated in the study.

HSOPS questionnaire and safety culture scores

The HSOPS is a self-administered 42-item questionnaire designed to assess 12 dimensions of patient safety culture from the perspective of hospital staff in the United States (US). The HSOPS questionnaire was originally developed by the AHRQ on the basis of a literature review, refined in accordance with psychometric theory, and supported by psychometric analyses performed on 1,437 hospital staff working in 21 US hospitals in 2004. We used the French version of the HSOPS questionnaire, which yielded promising psychometric properties in 4,289 hospital employees across eight academic and nonacademic hospitals in France [16 17].

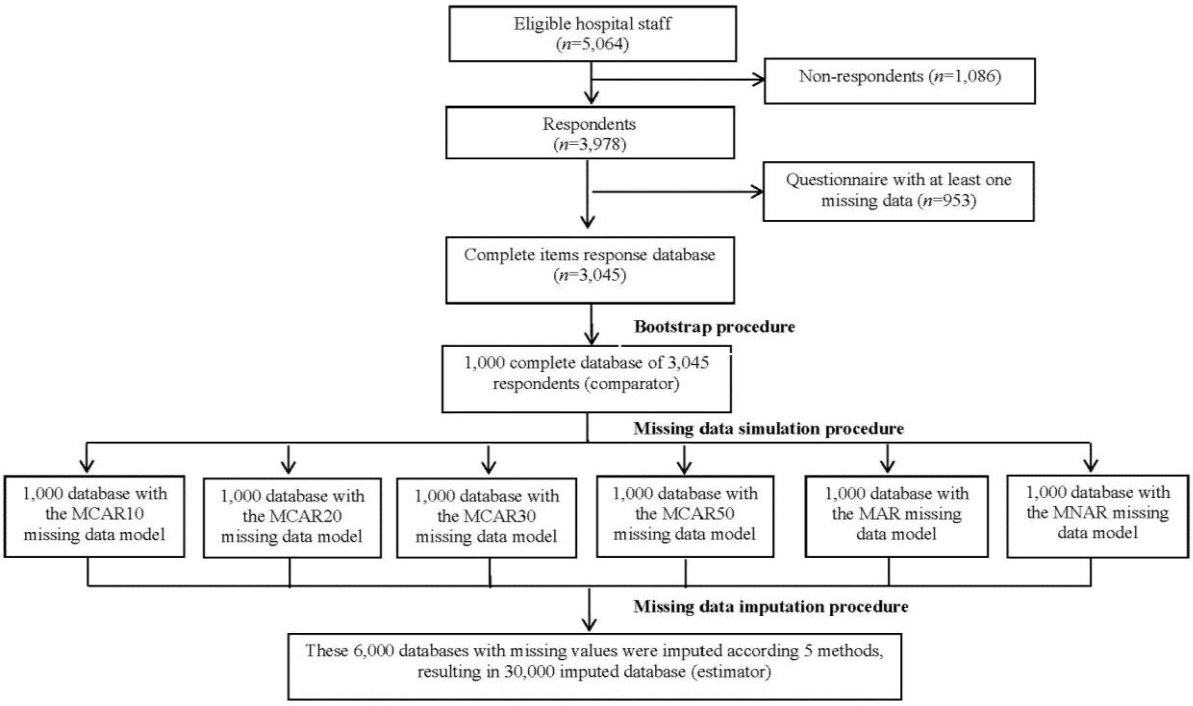
Each item was rated on a five-point Likert scale of agreement, ranging from "strongly disagree" to "strongly agree," or frequency, ranging from "never" to "always,". Items were coded using a 5-point response scale, with higher values denoting a better rating, after reverse coding for negatively worded items [4]. The dimension score computation method was the percentage of positive answers defined as values of 4 ("Agree" or "Most of the time") or 5 ("Strongly agree" or "always"). Following AHRQ guidelines, the first step was to obtain, for each item, the percentage of respondents who answered it positively among non-missing answers [4]. Then unweighted averages of those percentages were computed for each dimension. By definition, this method was a mean of percentages, hence resulting in dimension scores ranging from zero to 100 [18].

Missing data models

Of the 3,979 respondents, 3,045 completed all items in the HSOPS questionnaire. We produced simulation data on two levels. First, we simulated 1,000 completed data based on the original 3,045 using a bootstrap procedure. Second, missing values were simulated in these 1,000 complete cases data for comparison purposes, using 6 different missing data scenarios (Fig 1).

1) Missing completely at random 10% (MCAR10)
For each response to the 42 items, we generated a random number between 1 and 10 randomly selected from the uniform distribution (1,10); each number between 1 and 10 has an equal probability of being assigned. Item-responses assigned a value of 1 were deleted simulating a sample with 10% of the collected item-responses missing.
2) Missing completely at random 20% (MCAR20)
Item-responses assigned a value of 2 or less were deleted simulating a sample with 20% of the collected item-responses missing.
3) Missing completely at random 30% (MCAR30)
Item-responses assigned a value of 3 or less were deleted simulating a sample with 30% of the collected item-responses missing.
4) Missing completely at random 50% (MCAR50)
Item-responses assigned a value of 5 or less were deleted simulating a sample with 50% of the collected item-responses missing.
5) Missing at random (MAR)
The probability of missing was linked to the gender of the participants: females were assigned a 20% probability of non-response, while male participants a 10% probability of non-response.
6) Missing not at random (MNAR)
We used the individual dimension scores to assign the missing response probability. Participants with a dimension score above or equal to 50% were assigned a 10% probability of non-response, while participants with a dimension score lower than 50% a 20% probability of non-response.

Figure 1: Missing data simulation and imputation process



Imputation methods
Five methods of imputation were compared: 1) random selection, 2) Item mean, 3) individual mean, 4) multiple imputation, 5) sparse nonnegative matrix factorization (sNMF).
1) Random Selection
Missing responses were imputed with a randomly selected value from 1 to 5.
2) Item mean
We imputed the missing responses with the overall mean of the related item.
3) Individual mean
Individual mean method imputed the mean of the subject's complete responses to other items.
4) Multiple imputation
Missingness was addressed using the MICE (Multivariate Imputation by Chained Equations) algorithm. Five imputations were performed, using the Predictive Mean Matching method performed on the 42 items of the HSOPS. First, we estimated for cases with no missing data a linear regression for each variable based on all the other item responses, producing a set of coefficients b. Then, a random draw from the posterior distribution of b was performed and using this b we generated predicted value for all cases (missing and non-missing). For each case with missing data, we identified a set of cases with observed data whose predicted values were closed to the predicted values for the case with missing data (set of cases = 10). Then a value was randomly chosen from these set of case and substitute for the missing value.
5) Sparse non-negative matrix factorization
Finally, we imputed the missing responses using the sparse nonnegative matrix factorization (sNMF). This approach implements a nonnegative matrix factorization matrix introduced by Kim et al. [19]. Initially developed for exploratory analysis, the factorization matrix methods were adapted to missing data imputation. The implementation is built on a regularized version of alternate least mean-square algorithm for nonnegative matrix factorization. We used the software R implementation proposed by the package LEA [20].

Statistical analysis
The number and percentage of missing data was calculated for each item and for all the 12 dimensions. Then we assessed the number and percentage of participant with at least one missing item response. To evaluate the mechanism of the missing data, we performed comparison of the distribution of participants with

at least one missing data versus participants with complete responses according to baseline characteristics, using chi-square tests.

We performed a 2-steps analysis to assess the global performance of the five imputation methods, for each of the 6 missing data models. Firstly, we assessed the bias between the 1000 complete dataset and their related imputed dataset using the root mean square error (RMSE). Finally, we calculated the mean of the 1,000 RMSE for each missing data model and for each imputation method.

Secondly, we estimated the bias between the dimensions scores of the complete dataset and the related scores of the imputed dataset. We computed the mean difference between the 12 dimensions scores of the complete dataset and the related dimension scores of the imputed dataset.

Analyses were performed using Stata version 14.0 (Stata Corporation, College Station, TX), with the exception of the sNMF which was performed using R software version 3.2.1 (2015 The R Fundation for Statistical Computing) with the LEA package version 1.0.

RESULTS

Table 1 showed the number and percentage of missing data in the original dataset. Among the 42 items, the number of missing data was bounded between 18 (A4. `People treat each other with respect') and 140 (A7. `Use more agency/tempory staff'), representing respectively 0.5% and 3.6%. Overall, 843 (21.7%) participants had at least one missing item response.

Table 1: Distribution of missing data in the original HSOPS dataset

Missing Data, N (%)	Study population (n=3888)	
Dim 1: Overall Perceptions of Safety	**168**	**4.3**
A10r. Just by chance no more serious mistakes	64	1,6
A15. Safety never sacrificed	57	1.5
A17r. We have safety problems	73	1,9
A18. Good at preventing errors	53	1,4
Dim 2: Frequency of Event Reporting	**135**	**3.5**
D1. Mistake caught and corrected	106	2,7
D2. Mistake has no potential to harm	109	2,9
D3. Mistake could harm	120	3,1
Dim 3: Supervisor/manager expectations & actions promoting safety	**72**	**1,9**
B1. Supervisor says good word	45	1,2
B2. Supervisor considers staff suggestions	48	1,2
B3r. Supervisor wants faster work, shortcuts	47	1,2
B4r. Supervisor overlooks problems	47	1,2
Dim 4: Organizational Learning—Continuous improvement	**161**	**4.1**
A6. Doing this to improve safety	49	1.3
A9. Mistakes led to positive changes	63	1,6
A13. Evaluate effectiveness of changes	90	2.3
Dim 5: Teamwork Within Hospital Units	**110**	**2.8**
A1. People support one another	40	1,0
A3. Work together as a team	21	0,5
A4. People treat each other with respect	18	0.5
A11.Others help out when busy	46	1.2
Dim 6: Communication Openness	**65**	**1.7**
C2. Freely speak up	36	1.0
C4. Feel free to question decisions and actions	41	1,1
C6r. Staff afraid to ask questions	25	0,6
Dim 7: Feedback and Communication About Error	**87**	**2.2**
C1. Feedback about changes	62	1,6
C3. Informed about happened errors	40	1,0
C5. Discuss ways of preventing errors	31	0,8
Dim 8: Non-punitive Response To Error	**120**	**3.1**
A8r. Mistakes held against them	57	1.5
A12r. Person written up when event reported	40	1.0
A16r. Mistakes kept in their personnel file	82	2.1
Dim 9: Staffing	**237**	**6.1**
A2. Enough staff to handle workload	36	0,9
A5r. Work longer hours	122	3.1
A7r. Use more agency/temporary staff	140	3.6
A14r. Work in "crisis mode"	90	2.3
Dim 10: Hospital Management Support for Patient Safety	**152**	**3.9**
F1. Management provides work climate	86	2,2
F8. Safety is a top priority	104	2,7
F9r. Interest only after adverse events	91	2,3
Dim 11: Teamwork Across Hospital Units	**159**	**4.1**
F2r. Units do not coordinate well	77	2.0
F4. Good cooperation among units	89	2,3
F6r. Unpleasant with other units	84	2,2
F10. Units work well together	101	2,6
Dim 12: Hospital Handoffs & Transitions	**203**	**5.2**
F3r. Things "fall between the cracks"	118	3.0
F5r. Lost information during shift changes	129	3.3
F7r. Problems in the exchange of information	113	3.0
F11r. Shrift changes are problematic	126	3,2
All dimensions	**873**	**21.7**

The relationship between missing data and baseline characteristics was presented in Table 2. Overall, feminine gender was significantly associated with higher missing data rates, and participants younger than 44 year old were significantly better respondents.

Table 2. Distribution of respondents with at least one missing value according to baseline characteristics

	All respondents		All items (n=3045)		At least one item missing (n=843)		
	n	%	n	%	n	%	*P*-value
Profession (n= 3912)							0.07
Nurse	1386	36.3	1108	36.9	278	34.4	
Nurse assistant	708	18.6	544	18.1	164	20.3	
Physician	436	11.4	359	12.0	77	9.5	
Other healthcare	124	3.3	104	3.5	20	2.5	
Administrative	331	8.7	254	8.5	77	9.5	
Technical	378	9.7	286	9.5	92	11.4	
Other	450	11.8	101	11.6	101	12.5	
Gender (n= 3741)							<0.001
Women	3036	80.7	2340	79.2	676	85.9	
Men	727	19.3	614	20.8	111	14.1	
Age, (n=3754)							<0.001
< 35	1410	37.3	1185	39.9	221	28.0	
35-44	975	25.8	771	26.0	195	24.8	
45-54	973	25.8	707	23.8	260	33.0	
≥ 55	419	11.1	304	10.2	111	14.1	
Seniority in hospital, (n=3802)							<0.001
< 1 year	209	5.4	174	5.8	33	4.1	
1 to 2 years	391	10.3	329	11.0	62	7.7	
3 to 5 years	697	18.3	568	19.0	129	16.0	
> 6 years	2507	65.9	1926	64.3	581	72.2	
Hospital sector							>0.05
Medicine and pediatrics	1159	29.8	888	29.2	271	32.1	
Surgery and gynecology	833	21.4	610	21.6	174	20.6	
Technical (pharmacy, operating rooms, imaging, laboratories)	795	20.4	659	20.0	185	21.9	
Intensive care, emergency and anesthesia	688	17.7	566	18.6	122	14.5	
Several or other	413	10.6	322	10.6	91	10.6	

Table 3 gives a summary comparing performances of the 6 imputation methods in term of RMSE of each imputed dataset versus the complete ones. The results are reported for all responses as well as split by true items' response values from one to five. As expected, the random selection method yielded the worse RMSE under all the missing data models. The performances of the four other methods were close. However, the lowest RMSE was observed when multiple imputation method was used. The individual mean and the item mean method produced similar RMSE, slightly lower than the sparse non-negative matrix factorization. Finally, the performances appeared to be little or no difference in mean of RMSE in either of missing data scenarios. Table 3 reports also the mean of RMSE according the five possible item response values. Overall, every imputation methods performed better when the missing response values were 3. The worse performances were found for the imputations of extremes response values (i.e. item response value = 1 or 5) whatever the missing data scenario and the imputation method. However, the differences between the methods were more pronounced for these extreme values, with a clear advantage for the multiple imputation method.

Table 4 provides the bias between the 12 HSOPS dimensions scores of the complete dataset and the related scores of the imputed dataset. The largest bias were found for the two mean imputation methods (item mean and individual mean methods), and the smallest bias were found for the multiple imputation and the SNMF method. For example, for the MCAR50 missing data scenario, the bias ranged between (-20.8; +5.9) and (-19.8; +2.7), using the item mean and the individual mean imputation methods, respectively. While for the same MCAR50 scenario, the bias were smaller when we imputed with the multiple imputation and the SNMF methods (-6.3; -0.8 and -2.1; +6.5, respectively).

Table 3. Mean of RMSE according to missing data scenarios

	Mean of RMSE					
	Global	Item R = 1	Item R = 2	Item R = 3	Item R = 4	Item R = 5
MCAR (P=0.10)						
Random Selection	1.763	2.449	1.732	1.414	1.731	2.449
Item Mean	1.008	1.954	1.094	0.492	0.875	1.599
Individual Mean	1.018	1.940	1.141	0.454	0.878	1.652
Multiple Imputation	0.916	1.542	1.007	0.652	0.804	1.300
sNMF	1.086	1.870	1.273	0.793	0.884	1.485
MCAR (P=0.20)						
Random Selection	1.763	2.449	1.733	1.411	1.732	2.448
Item Mean	1.007	1.953	1.094	0.492	0.875	1.599
Individual Mean	1.020	1.944	1.143	0.460	0.877	1.652
Multiple Imputation	0.926	1.568	1.018	0.656	0.812	1.320
sNMF	1.097	1.901	1.294	0.794	0.889	1.500
MCAR (P=0.30)						

	Mean of RMSE					
	Global	Item R = 1	Item R = 2	Item R = 3	Item R = 4	Item R = 5
Random Selection	1.763	2.449	1.733	1.414	1.732	2.450
Item Mean	1.007	1.952	1.094	0.492	0.875	1.599
Individual Mean	1.021	1.947	1.147	0.465	0.877	1.651
Multiple Imputation	0.937	1.595	1.028	0.659	0.820	1.342
sNMF	1.112	1.933	1.317	0.799	0.898	1.521
MCAR (P=0.50)						
Random Selection	1.763	2.449	1.732	1.414	1.732	2.450
Item Mean	1.007	1.955	1.095	0.493	0.873	1.598
Individual Mean	1.029	1.958	1.161	0.486	0.876	1.649
Multiple Imputation	0.962	1.659	1.054	0.667	0.834	1.398
sNMF	1.152	2.011	1.366	0.815	0.934	1.582
MAR (Sex)						
Random Selection	1.762	2.449	1.732	1.414	1.732	2.449
Item Mean	1.004	1.949	1.092	0.491	0.874	1.595
Individual Mean	1.019	1.958	1.139	0.453	0.880	1.654
Multiple Imputation	0.924	1.571	1.014	0.654	0.813	1.315
sNMF	1.096	1.915	1.289	0.790	0.893	1.494
MNAR						
Random Selection	1.722	2.448	1.732	1.414	1.733	2.448
Item Mean	1.024	2.045	1.162	0.519	0.813	1.543
Individual Mean	1.014	1.940	1.139	0.454	0.902	1.712
Multiple Imputation	0.927	1.496	0.968	0.639	0.910	1.477
sNMF	1.122	1.859	1.265	0.778	0.959	1.647

Table 4: Dimension score bias according to missing data scenarios

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
Number of items	**4**	**3**	**4**	**3**	**4**	**3**	**3**	**3**	**4**	**3**	**4**	**4**
Complete database	41.5	49.2	56.5	49.4	62.9	59.0	55.4	28.9	32.2	19.7	34.7	25.0
MCAR (P=0.10)												
Random Selection	-0.1	-0.9	-1.6	-0.9	-1.5	-1.9	+1.1	-2.3	+0.8	+2.0	+0.5	+1.5
Item Mean	-4.2	-3.7	-0.6	-1.9	-2.2	+0.8	-2.9	+1.2	-0.7	-2.0	-3.3	-2.5
Individual Mean	-2.0	-2.7	-3.5	-2.8	-3.4	-3.8	-0.6	-4.1	-0.9	+0.4	-1.2	-0.2
Multiple Imputation	-0.7	-0.5	-0.2	-0.4	-0.7	-0.4	-1.0	-0.5	-0.7	-0.9	-0.8	-1.2
sNMF	+0.7	+0.2	+1.4	+0.8	+0.6	+0.3	0.0	+0.6	-0.2	+0.1	+0.4	+0.1
MCAR (P=0.20)												
Random Selection	-0.3	-1.8	-3.3	-1.9	-3.1	-3.8	+2.2	-4.6	+1.6	+4.1	+1.1	+3.0
Item Mean	-8.3	-7.4	-1.3	-3.8	-4.4	+1.5	-5.8	+2.4	-1.5	-3.9	-6.5	-5.0
Individual Mean	-3.9	-5.4	-6.9	-5.5	-6.7	-7.4	-1.2	-8.2	-1.8	+0.8	-2.4	-0.3
Multiple Imputation	-1.5	-0.9	-0.4	-0.9	-1.5	-0.9	-2.1	-1	-1.4	-1.9	-1.6	-2.3
sNMF	+1.3	+0.3	+2.3	+1.4	+0.9	+0.4	+0.6	+0.7	0.0	+0.8	+0.9	+0.7
MCAR (P=0.30)												
Random Selection	-0.4	-2.8	-4.9	-2.8	-4.6	-5.7	+3.3	-6.9	+2.3	+6.1	+1.6	+4.5
Item Mean	-12.5	-11.0	-1.9	-5.9	-6.5	+2.3	-8.7	+3.6	-2.2	-5.9	-9.8	-7.5
Individual Mean	-5.7	-8.0	-10.3	-8.1	-9.9	-11.0	-1.7	-12.3	-2.6	+1.3	-3.5	-0.4
Multiple Imputation	-2.2	-1.4	-0.6	-1.3	-2.2	-1.3	-3.3	-1.4	-2.1	-3.0	-2.5	-3.6
sNMF	+2.0	+0.5	+2.6	+1.8	+0.9	+0.1	+1.6	+0.4	+0.5	+2.0	+1.7	+1.7
MCAR (P=0.50)												
Random Selection	-0.8	-4.6	-8.2	-4.7	-7.7	-9.5	+5.5	-11.5	+3.9	+10.2	+2.7	+7.5
Item Mean	-20.8	-18.1	-3.2	-10.4	-10.6	+3.8	-14.5	+5.9	-3.6	-9.8	-15.7	-12.5
Individual Mean	-8.8	-12.7	-16.5	-12.8	-15.8	-17.7	-2.2	-19.8	-3.7	+2.7	-5.1	0.0
Multiple Imputation	-3.8	-2.4	-0.8	-2.1	-3.7	-2.1	-5.8	-2.3	-3.8	-5.4	-4.3	-6.3
sNMF	+3.0	+0.8	+1.3	+2.1	-0.1	-1.8	+5.0	-2.1	+2.7	+6.5	+3.7	+5.4
MAR (Sex)												
Random Selection	-0.2	-1.6	-2.8	-1.6	-2.7	-3.3	+2.0	-4.1	+1.3	+3.6	+1.0	+2.7
Item Mean	-7.3	-6.5	-1.1	-3.2	-3.8	+1.4	-5.1	+2.0	-1.3	-3.4	-5.7	-4.4
Individual Mean	-3.5	-4.8	-6.2	-4.9	-5.9	-6.6	-1.1	-7.5	-1.8	+0.6	-2.1	-0.4
Multiple Imputation	-1.3	-0.8	-0.4	-0.8	-1.3	-0.8	-1.9	-0.9	-1.3	-1.7	-1.4	-2.1
sNMF	+1.2	+0.3	+2.1	+1.3	+0.8	+0.4	+0.4	+0.6	-0.2	+0.6	+0.9	+0.5
MNAR												
Random Selection	+2.0	+1.9	+0.7	+1.3	+0.6	+0.1	+3.6	-0.2	+2.9	+5.0	+2.6	+4.0
Item Mean	-3.0	+1.1	+2.2	+0.5	+2.0	+2.9	-2.9	+2.8	+0.7	-2.0	-0.2	-2.2
Individual Mean	-1.3	-0.7	-1.8	-1.7	-2.2	-2.4	+0.5	-2.2	0.0	+1.8	-0.5	+1.2
Multiple Imputation	+0.0	+0.1	+0.3	+0.7	+0.2	+0.6	-0.6	+0.1	+0.1	-0.8	+0.1	-0.7
sNMF	+1.9	+1.9	+2.5	+2.2	+1.6	+1.4	+1.6	+1.2	+1.2	1.6	+1.9	+1.5

DISCUSSION

This computer-intensive simulation study shows that multiple imputation is the most accurate imputation method in every of the six scenarios that we assessed in our original HSOPS dataset. By contrast, the simple imputation methods, individual mean and item mean imputations, are the less accurate imputation methods. The bias regarding the item response values were close between the multiple imputation and the two mean imputation methods, with a slightly advantage for the multiple imputation method. This bias was higher regarding the dimension scores, with a clear advantage for the multiple imputation and the sNMF method in comparison with the mean imputation methods. Surprisingly, the mean imputation methods

yielded higher dimension scores bias than the random selection method.

To our knowledge, the handling of missing data has never been assessed for the HSOPS-based surveys. This lack leads to heterogeneous methodologic choices from research teams conducting patient safety culture surveys. For example, we found three HSOPS-based surveys imputing the missing data with the mean of individual responses [11 13 14], while two research teams used multiple imputation [12 15]. Such methodological inconsistencies may alter the interpretability of the HSOPS results and avoid benchmark comparisons between different surveys. With our study, we provide empirical evidence for the use of multiple imputation and we show that mean imputation should be avoided to handling HSOPS missing data. Our results are in line with Shrive et al. who reported in 2006 better performances for the use of multiple imputation in comparison with individual mean response imputation in a simulation study based on a depression scale [10]. However, they concluded that mean imputation was the best approach given the attractive balance of accuracy and conceptual simplicity. Nowadays, the multiple imputation method is available in all standard statistical software and is easy to use following the software documentations that provide practical examples. The increase of computer performances allows also to perform multiple imputation very quickly. Hence we think that the actual balance of accuracy and conceptual simplicity had considerably changed, and the relative complexity of multiple imputation implementation can no longer stand.

The main problem with multiple imputation is the assumption of MAR mechanism. The concept of missing data mechanism is often theoretical and hard to assess in routine. Indeed, MAR is only an assumption that no statistical method can verified [21]. To find the type of our missing data, we compared the distribution of subject with at least one missing data according to socio-professional characteristics. Women were less likely to answers all the survey, like young subject and peoples working in surgery, gynecology, emergency, anesthesia and intensive care sector. On the over hand, the senior worker let more missing data on the questionnaire. As other variables of the survey could explain the missingness, we assumed that data were MAR. Clustering of items according to the fraction of missing data is another tools used to investigate the nature of missing value [22]. Our clustering pointed that items groups obtain were not random (data not shown). These clusters mainly had the same composition of HSOPS dimensions. As dimension are composed of items dealing with the same topic (event reporting, management or communication), it indicated that missing data could also be considered as MNAR, at least for a part of them. These results clearly pointed the difficulty to understand missing data mechanism. This difficulty directly impacts the choice of the missing data management method, because each method works better with one types of mechanism [23 24].

To deal with these problems, we chose to simulate different mechanism of missing data, MCAR, MNAR and MAR. Interestingly, the multiple imputation performed well for all the scenarios, whether for different percentages of missing data and for different mechanisms simulated. These results are in line with Faris et al. [25], who empirically demonstrated good performances of multiple imputation whatever the mechanisms simulated. Applied with real HSOPS data, our results showed that the multiple imputation was appropriate and performed better than simple imputation methods, even without evidence of the MAR assumption in our HSOPS-based survey.

Another way to overcome the MAR assumption is the use of the sNMF method. Indeed, this method, frequently used to solve machine learning problems, works independently of missing data mechanism. Applied to our data, this method yielded better performances regarding the dimension scores bias than mean imputation methods, but not as well as the multiple imputation approach. However, this method is not available with all statistical software and requires specific machine learning knowledge, making harder for research team who conduct HSOPS-based surveys. For now, the balance between conceptual complexity and accuracy also limits sNMF application in routine.

Mean imputation methods, individual mean and item mean, achieved the worse performance in our study, yielding huge dimension scores bias. If global RMSE were close in comparison with multiple imputation, this method had the worst performance to impute the extreme values of item response (i.e., item response = 1 or 5). The dimension scores are percentages of positive response obtained after a dichotomization of the items response values [18]. The worst performances of mean imputation for extreme values impacted strongly the percentage of positive responses through an increase of the regression mean. This could also explain the paradoxical results showing better performances for the random selection method in comparison with the mean imputation methods.

The excellent performance of the multiple imputation method, even for the highest percentages of deletion simulated in our study, raises question regarding the properties of the HSOPS questionnaire. Moreover, most of external validation studies reported poor psychometric properties of this instrument [12 13 17 26]. Multiple imputation method replaces the missing values through multivariate model using the information contained in non-missing responses. Achieving such good performances despite 50% of missing data may suggest that many of the 42 items are redundant. Hence it could be possible to reduce the pool of items to eliminate redundant items and arrange a shorter questionnaire. Our data collection period lasted 18 months to survey 5,000 staff members, with a 78% response rate. On average, 278 professionals were surveyed and 221 questionnaires were returned every month. Another recent HSOPS survey included 368 participants with a 71% response rate over a 7-month period [27]. With such long data collection periods, it cannot be excluded that many factors may have changed affecting the staffs' perception of patient safety culture in their hospital. Although speculative, using a shorter version of this questionnaire could also be helpful to reduce the data collection period and increase the response rates.

The limitations of our study deserve mentions. First, we did not collect information on nonrespondent characteristics, given the anonymity of the survey. Therefore we could not examine the potential for selection bias. Second, the random simulations may not be reflective of the patterns of missing data encountered in real situations. Third, the imputation methods that we chose were only a fraction of existing methods, and we cannot exclude that other perform better than multiple imputation.

CONCLUSION

In conclusion, we recommend the use of multiple imputation to handle missing data in HSOPS-based surveys. Notably, mean imputation methods should be avoided. Indeed, individual mean and item mean imputations methods yielded clear dimension scores bias, which could lead to misinterpretation of HSOPS results. Overall, it is necessary to take into account the imputation methods used before to compare different HSOPS surveys. The performance of the multiple imputation for all missing data mechanisms, up to 50% of missingness, may allow to optimize the original HSOPS instrument. Indeed, our results suggest that HSOPS-based surveys should be conducted after

reducing substantially the item number, without loss of information.

Funding

This work was supported by a grant from the French Ministry of Health (Programme de recherche sur la performance du système de soins-PREPS 12-026-0001).

REFERENCES

1. Etchegaray JM, Thomas EJ. Comparing two safety culture surveys: safety attitudes questionnaire and hospital survey on patient safety. BMJ Qual Saf 2012;21(6):490-8 doi: 10.1136/bmjqs-2011-000449.
2. Kemp K, Warren S, Chan N, et al. Qualitative complaints and their relation to overall hospital rating using an H-CAHPS-derived instrument. BMJ Qual Saf 2015 doi: 10.1136/bmjqs-2015-004371.
3. Johnson TP, Wislar JS. Response rates and nonresponse errors in surveys. JAMA 2012;307(17):1805-6 doi: 10.1001/jama.2012.3532.
4. Sorra J, Nieva VF. Hospital Survey on Patient Safety Culture. (Prepared by Westat, under Contract No. 290-96-0004). AHRQ Publication No. 04-0041. Rockville, MD: Agency for Healthcare Research and Quality. 2004. http://www.ahrq.gov/sites/default/files/publications/files/hospcult.pdf.
5. Demissie S, LaValley MP, Horton NJ, et al. Bias due to missing exposure data using complete-case analysis in the proportional hazards regression model. Stat Med 2003;22(4):545-57 doi: 10.1002/sim.1340.
6. Joseph L, Belisle P, Tamim H, et al. Selection bias found in interpreting analyses with missing data for the prehospital index for trauma. J Clin Epidemiol 2004;57(2):147-53 doi: 10.1016/j.jclinepi.2003.08.002.
7. Rotnitzky A, Wypij D. A note on the bias of estimators with missing data. Biometrics 1994;50(4):1163-70.
8. Hardouin JB, Conroy R, Sebille V. Imputation by the mean score should be avoided when validating a Patient Reported Outcomes questionnaire by a Rasch model in presence of informative missing data. BMC Med Res Methodol 2011;11:105 doi: 10.1186/1471-2288-11-105.
9. Resseguier N, Verdoux H, Giorgi R, et al. Dealing with missing data in the Center for Epidemiologic Studies Depression self-report scale: a study based on the French E3N cohort. BMC Med Res Methodol 2013;13:28 doi: 10.1186/1471-2288-13-28.
10. Shrive FM, Stuart H, Quan H, et al. Dealing with missing data in a multi-question depression scale: a comparison of imputation methods. BMC Med Res Methodol 2006;6:57 doi: 10.1186/1471-2288-6-57.
11. Feng X, Bobay K, Krejci JW, et al. Factors associated with nurses' perceptions of patient safety culture in China: a cross-sectional survey study. J Evid Based Med 2012;5(2):50-6 doi: 10.1111/j.1756-5391.2012.01177.x.
12. Hammer A, Ernstmann N, Ommen O, et al. Psychometric properties of the Hospital Survey on Patient Safety Culture for hospital management (HSOPS_M). BMC Health Serv Res 2011;11:165 doi: 10.1186/1472-6963-11-165.
13. Ito S, Seto K, Kigawa M, et al. Development and applicability of Hospital Survey on Patient Safety Culture (HSOPS) in Japan. BMC Health Serv Res 2011;11:28 doi: 10.1186/1472-6963-11-28.
14. Smits M, Christiaans-Dingelhoff I, Wagner C, et al. The psychometric properties of the 'Hospital Survey on Patient Safety Culture' in Dutch hospitals. BMC Health Serv Res 2008;8:230 doi: 10.1186/1472-6963-8-230.
15. Vlayen A, Schrooten W, Wami W, et al. Variability of patient safety culture in Belgian acute hospitals. J Patient Saf 2015;11(2):110-21 doi: 10.1097/PTS.0b013e31829c74a3.
16. Occelli P, Quenon JL, Kret M, et al. Validation of the French version of the Hospital Survey on Patient Safety Culture questionnaire. Int J Qual Health Care 2013;25(4):459-68 doi: 10.1093/intqhc/mzt047.
17. Boussat B, Francois P, Gandon G, et al. Inconsistencies Between Two Cross-Cultural Adaptations of the Hospital Survey on Patient Safety Culture Into French. J Patient Saf 2017 doi: 10.1097/PTS.0000000000000443.
18. Giai J, Boussat B, Occelli P, et al. Hospital survey on patient safety culture (HSOPS): variability of scoring strategies. Int J Qual Health Care 2017;29(5):685-92 doi: 10.1093/intqhc/mzx086.
19. Kim H, Park H. Nonnegative matrix factorization based on alternating nonnegativity constrained least squares and active set method. . SIAM J. Matrix Anal. & Appl. 2008;30(2):713-30 doi: 10.1137/07069239X.
20. Frichot E, Mathieu F, Trouillon T, et al. Fast and efficient estimation of individual ancestry coefficients. Genetics 2014;196(4):973-83 doi: 10.1534/genetics.113.160572.
21. Schafer JL, Graham JW. Missing data: our view of the state of the art. Psychol Methods 2002;7(2):147-77.
22. Tian J, Yu B, Yu D, et al. Clustering-based multiple imputation via gray relational analysis for missing data and its application to aerospace field. ScientificWorldJournal 2013;2013:720392 doi: 10.1155/2013/720392.
23. Eekhout I, de Vet HC, Twisk JW, et al. Missing data in a multi-item instrument were best handled by multiple imputation at the item score level. J Clin Epidemiol 2014;67(3):335-42 doi: 10.1016/j.jclinepi.2013.09.009.
24. Schlomer GL, Bauman S, Card NA. Best practices for missing data management in counseling psychology. J Couns Psychol 2010;57(1):1-10 doi: 10.1037/a0018082.
25. Faris PD, Ghali WA, Brant R, et al. Multiple imputation versus data enhancement for dealing with missing data in observational health care outcome analyses. J Clin Epidemiol 2002;55(2):184-91.
26. Perneger TV, Staines A, Kundig F. Internal consistency, factor structure and construct validity of the French version of the Hospital Survey on Patient Safety Culture. BMJ Qual Saf 2014;23(5):389-97 doi: 10.1136/bmjqs-2013-002024.
27. Mallouli M, Tlili MA, Aouicha W, et al. Assessing patient safety culture in Tunisian operating rooms: A multicenter study. Int J Qual Health Care 2017 doi: 10.1093/intqhc/mzw157.

Acknowledgements

This work was developed within the framework of the Grenoble Alpes Data Institute, which is supported by the French National Research Agency under the "Investissements d'avenir" program (ANR-15-IDEX-02).
We acknowledge Linda Northrup for her assistance in editing the manuscript in English.

3.4. Etude qualitative des commentaires libres : article soumis – The contribution of open comments to understanding the results from the Hospital Survey on Patient Safety Culture: a qualitative study

Les 42 questions du HSOPS permettent de recueillir les données de façon standardisée et de calculer des scores de culture de sécurité, qui peuvent ensuite être utilisés pour effectuer des comparaisons statistiques. Mais il parait également intéressant pour ce type d'enquête d'explorer les aspects qualitatifs de la culture des professionnels de santé. Le questionnaire HSOPS ouvre cette possibilité par le biais de l'analyse des commentaires, que les répondants sont libres d'écrire dans une section dédiée. Etonnamment, aucun des articles référencés par l'AHRQ ne s'est intéressé à l'analyse des commentaires libres du HSOPS. Dans l'article suivant (soumis, en révisions mineures), nous présentons les résultats de cette analyse des commentaires, que nous avons complétée d'entretiens semi-directifs de professionnels de santé.

Objectif : L'objectif de cette étude était d'analyser les données qualitatives recueillies dans la section des commentaires des questionnaires HSOPS et complétées par des entretiens réalisés auprès de professionnels de santé.

Méthode : Les commentaires des questionnaires recueillis dans le cadre de l'enquête de culture de sécurité ont été retranscrits et soumis à une triple lecture (BB, KK, PF). Les trois relecteurs ont conduit indépendamment une analyse thématique, permettant de classer les commentaires selon plusieurs thèmes. Les résultats étaient ensuite comparés et les désaccords discutés jusqu'à l'obtention d'un consensus. Secondairement, des entretiens semi-directifs ont été conduits auprès de 19 professionnels de santé de l'établissement. Ces entretiens ont été réalisés en s'appuyant d'un guide développé à partir de l'analyse thématique des commentaires. Ce guide comprenait une première question d'ordre générale au sujet de la perception de la sécurité des patients et des principaux facteurs l'influençant. Le professionnel interviewé répondait librement, parlant des thèmes de son choix. Si le professionnel n'évoquait pas spontanément les thèmes du guide, l'interviewer les suggérait. Selon les réponses, l'interviewer était libre de

de demander à l'interviewé de détailler, clarifier ou compléter ses réponses. L'interviewer adaptait donc ses questions selon le contexte de l'interview. Toutes les interviews ont été enregistrées et retranscrites. Ces données ont ensuite été analysées selon la méthode inductive. Les extraits des interviews étaient catégorisés en utilisant la classification thématique des commentaires libres.

Résultats : Un total de 284 commentaires a été recueilli à partir des 3978 questionnaires. Après exclusion de 36 commentaires (6 illisibles ou incompréhensibles et 30 non pertinents), l'échantillon comportait 247 commentaires analysables. Les caractéristiques des répondants ayant écrit un commentaire étaient comparables à l'ensemble des répondants. 19 professionnels de santé ont été interviewés (9 infirmiers diplômés d'état, 7 médecins, 2 pharmaciens et 1 aide-soignant). Les professionnels ayant commenté avait des scores de culture de sécurité plus faibles que les autres répondants (taille d'effet moyen des 12 dimensions = -0.23, IC95% -0.36 à -0.10, P<0.001). Les trois plus importantes différences concernaient les dimensions « perception globale de la sécurité des patients », « ressources humaines » et « soutien du management à la sécurité des patients ». L'analyse des 247 commentaires a mis en évidence 6 thèmes principaux : « questionnaire » (101 occurrences), « ressources humaines et soutien du management » (98 occurrences), « organisation » (41 occurrences), « signalement des événements indésirables et coordination de la gestion du risque » (25 occurrences), « matériel et équipement » (16 occurrences) et « sécurité du personnel » (5 occurrences). Le thème relatif au questionnaire concernait principalement des commentaires indiquant que le questionnaire était mal adapté à certaines catégories professionnelles (exemple secrétaires), ou rapportant une inquiétude quant au respect de l'anonymat. Le second thème concernait les ressources humaines et le soutien du management. De nombreux commentaires rapportaient un manque de personnel et une charge de travail trop importante, conduisant à un épuisement des professionnels. Les interviews confirmaient cette perception de lien direct entre sous-effectif et diminution de la

sécurité des patients. Le mécontentement à l'encontre des directeurs et managers était très souvent recueilli au travers des commentaires et des interviews. « Le sous-effectif montre que la sécurité n'est pas vraiment une priorité » (commentaire), « C'est un combat quotidien contre la direction : ils veulent réduire nos ressources tout en nous demandant d'augmenter l'activité » (médecin interviewé 2). Le troisième thème concernait l'organisation, et comportait des commentaires relatifs aux difficultés rencontrées lors des transferts de patients. Les interviews confirmaient que le manque de communication ou de transmission entre services pouvait avoir des conséquences négatives sur la sécurité des patients : « quand les patients arrivent des urgences, ils ne sont pas toujours réévalués après leur transfert … Et parfois nous sommes surpris car l'état clinique des patient n'est tel que nous avions discuté … » (médecin interviewé 4). Le quatrième thème relatif à au signalement des événements indésirables et de la coordination de la gestion du risque comportait des éléments permettant de comprendre certaines barrières au signalement des événements indésirables. Plusieurs commentaires indiquaient ne plus signaler d'événements à cause du manque de retour d'information de la part des gestionnaires du risque de l'établissement, et du manque de temps pour signaler. D'autres commentaires pointaient le signalement comme source de conflit interprofessionnel. Enfin, certaines pistes d'amélioration, comme la décentralisation de la gestion du risque au sein des équipes soignantes étaient évoquées, au travers de programmes tels que les RMM et les CREX. Cependant, certains commentaires déploraient encore le manque de temps pour participer à ces activités : « Pour une infirmière, il est malheureusement impossible de trouver le temps pour participer aux CREX ou aux RMM » (infirmier interviewé 2). Les deux derniers thèmes concernaient la vétusté des équipements comme source d'événements indésirables, ainsi que le manque de sécurité du personnel, faisant face à des violences physiques entrainant des arrêts de travail, principalement aux urgences.

Discussion et conclusions : Cette étude qualitative a permis d'identifier plusieurs problèmes structurels perçus comme associés à une diminution de la sécurité des patients. De façon intéressante, ces failles étaient proches des dimensions ayant les plus faibles scores dans l'analyse quantitative de la culture de sécurité. Le manque de ressources, humaines et matérielles, était en effet fortement associé aux dimensions « ressources humaines » et « soutien du management à la sécurité des soins », qui étaient les 2 dimensions comportant les scores les plus faibles de l'enquête culture de sécurité. Mettant en évidence des problèmes concrets d'organisation, de management et de coordination de la gestion des risques, l'analyse des commentaires permettait de mieux comprendre et de préciser les aspects quantitatifs de l'enquête. Trop rarement effectuée lors des études basées sur le questionnaire HSOPS, l'analyse des données qualitatives permet de préciser les réponses aux questions fermées et de cibler des défaillances concrètes de l'organisation. Mixer les approches quantitatives et qualitatives dans ce type d'études pourrait donc permettre de développer des actions correctives visant à améliorer la culture de sécurité des professionnels de santé.

The contribution of open comments to understanding the results from the Hospital Survey on Patient Safety Culture (HSOPS): a qualitative study

Bastien Boussat[1,2*¶], Kevin Kamalanavin[1,3¶], Patrice François[1,2¶]

Abstract

Introduction To develop high-quality and safe healthcare, a good safety culture is an important feature of healthcare-providing structures. The objective of this study was to analyze the qualitative data of the comments section of a Hospital Survey on Patient Safety (HSOPS) questionnaire to clarify the answers given to the closed questions.

Method Using the original data from a cross-sectional survey of 5,064 employees at a single university hospital in France, we conducted a qualitative study by analyzing the comments of a HSOPS survey and conducting in-depth interviews with 19 healthcare providers. We submitted the comments and the interviews to a thematic analysis.

Results A total of 3,978 questionnaires were returned, with 247 comments collected. The qualitative analysis identified several structural failures. The main categories of the open comments were concordant with the lowest dimension scores found in the quantitative analysis. The most frequently reported failures were related to the staffing and hospital management support dimensions. The healthcare professionals perceived the lack of resources, including understaffing, as the major barrier to the development of a patient safety culture. Concrete organizational issues related to hospital handoffs and risk coordination were identified, such as transfers from the emergency departments and the lack of feedback following self-reporting of incidents.

Conclusion The analysis of the open comments complemented the HSOPS scores, increasing the level of detail in the description of the hospital's patient safety culture. Combined with a classical quantitative approach used in HSOPS-based surveys, the qualitative analysis of open comments is useful to identify organizational weaknesses within the hospital.

Keywords Patient safety; Safety culture; Qualitative research; surveys and questionnaires; interview

Introduction

Medical errors cause an estimated 251,454 deaths every year in the United States, the third leading cause of death,[1] even though patient safety has been a priority for health systems for more than 15 years. In 2004, the World Health Organization created the World Alliance for Patient Safety. This alliance has developed several initiatives, programs, and guides to enhance patient safety and reduce the occurrence of medical errors.[2-4] This raises the suspicion that safety enhancement efforts may have been inhibited by structural causes.

Inspired by the experience of high-reliability organizations such as air transport and the nuclear industry,[5] safety culture is defined as "the product of individual and group values, attitudes, perceptions, competencies, and patterns of behaviour that determine the commitment to, and the style and proficiency of, an organisation's health and safety management".[6] Safety culture has become increasingly important, inciting both professionals and healthcare facilities to adopt behaviors and management tools promoting patient safety. The rising importance of safety culture has resulted in the need to assess it and measure it.

Among the several tools developed to assess the patient safety culture, one of the most widely used is the Hospital Survey on Patient Safety Culture (HSOPS). Created by the Agency for Healthcare Research and Quality (AHRQ), this survey consists of a self-administered questionnaire including 42 items formulated as closed questions and used to calculate composite scores for 12 dimensions of safety culture.[6,7] It also includes a comments section, but to our knowledge, few surveys using the HSOPS questionnaire have analyzed these comments.

Closed questions allow collecting standardized information that can be statistically exploited, but the information is consequently poorer. However, comments give more detailed information about the concerns of staff members. Furthermore, interviews with staff members and the qualitative method enable us to understand the logic of their behavior and the influence of work conditions on this behavior. Therefore, it may complement the answers to the questionnaire's closed questions, thus clarifying their meaning. It also provides further information and helps identify the staff's preoccupations and structural issues that have not been diagnosed by the survey.[8,9]

The main objective of this study was to analyze the qualitative data of the comments section of the HSOPS to refine the survey results. These data were produced by analyzing the comments collected during the HSOPS survey and the interviews conducted with front-line staff members.

Methods

Study design

This study examined a qualitative part of an HSOPS-based survey. We first exploited the comments left in the comments area of the HSOPS questionnaire. This analysis was used to prepare the interview guide of an interview-based qualitative study conducted to further the analysis. The institutional review board at Grenoble University Hospital (IRB 6705) reviewed

the study protocol and waived the need for informed participant consent. The study protocol was approved by the Advisory Board on Medical Research Data Processing (CCTIRS) and authorization by the French Personal Data Protection Authority (CNIL) was obtained before data processing started.

Study site

The study was conducted at a single university-affiliated hospital with a capacity of 1836 beds (including 1175 acute care beds and 661 long-term or subacute care beds), serving a predominantly urban population of 675,000 inhabitants in France. The study site reported 135,999 stays in 2014. The hospital staff comprised 4,422 registered paramedical staff and 642 board-certified physicians. The number of beds and the paramedical staff rate was similar to other French university-affiliated hospitals (2336 versus 1836 beds, and 2.4 versus 2.6 paramedical staff per bed). [10]

Population

The HSOPS was conducted anonymously on a volunteer basis, department by department, between April 2013 and September 2014. Eligible participants were full-time or part-time (half-time or more) employees with at least 6 months of employment in the clinical, laboratory/pathology, radiology, or pharmacy departments. In accordance with the HSOPS guide,[6] this study sample encompassed clinical and nonclinical staff who had direct contact or interaction with patients and hospital staff who might not have direct contact with patients but whose work directly affected patient care.

As recommended,[6] HSOPS questionnaires were secondarily excluded if the respondent did not complete at least one survey section, answered fewer than half of the items, or answered every item with the same non-neutral response.

The comments were collected in the specific field at the end of the questionnaires. Then exclusion criteria were applied separately for the comments: illegible or incomprehensible comments and nonpertinent comments (such as “No,” “Nothing to report,” “None,” etc.) were excluded.

The interviews were conducted with 19 medical and paramedical professionals from May to July 2016. We limited the eligibility to selected categories of professionals in several selected departments. We interviewed physicians and pharmacists, head nurses, nurses, nursing aides, and stretcher bearers.

Data collection

The questionnaires were distributed in all departments by an investigator, cooperatively with head nurses, who established the list of staff members to be included.

Comments

Comments were collected in the comments section of the HSOPS questionnaire. They were scripted and submitted to a triple reading before being analyzed.

Interviews

The interviews were conducted from May to July 2016 and lasted an average 40 min. Nineteen staff members were interviewed and the interviews were recorded and transcribed.

Following the analysis of the comments, we developed an interview guide to conduct semi-directive interviews with staff members. The interview guide was composed of the topics identified by the comment analysis. The interviewer agreed with the interviewees that recordings and transcriptions would not be disseminated.

The interviews were based on a thematic guide comprising one main question – “What can you tell me about patient safety in the hospital and about the main factors influencing it?” – and the topics identified by analyzing the comments. At first, the staff members interviewed were asked to speak freely on this question, raising the topics of their choice. When the staff members did not raise the topics on the guide, the interviewer suggested topics. Depending on the responses, the interviewer was free to ask the interviewee to give details, clarify, or complete the responses. The interviewer also tailored the questions to the interview’s context and to the staff member being interviewed.

Analysis

Comments analysis

To analyze the comments, three interviewers conducted a thematic analysis independently and sorted the comments by topic. They then compared their results and agreed on a classification of the topics. Some comments concerned several topics and subsequently had several occurrences within different topics. Once this first sorting had been completed, the surveyors attributed one or several key words to each comment within a topic.

Interview analysis

We conducted an inductive qualitative textual analysis. Topic categorization of the comments analysis was used and interviews were cut into extracts, which we categorized within these topics. Then we attributed one or several key words to the extract depending on more specific topics or subjects discussed by the interviewee in this extract.

Statistical analysis

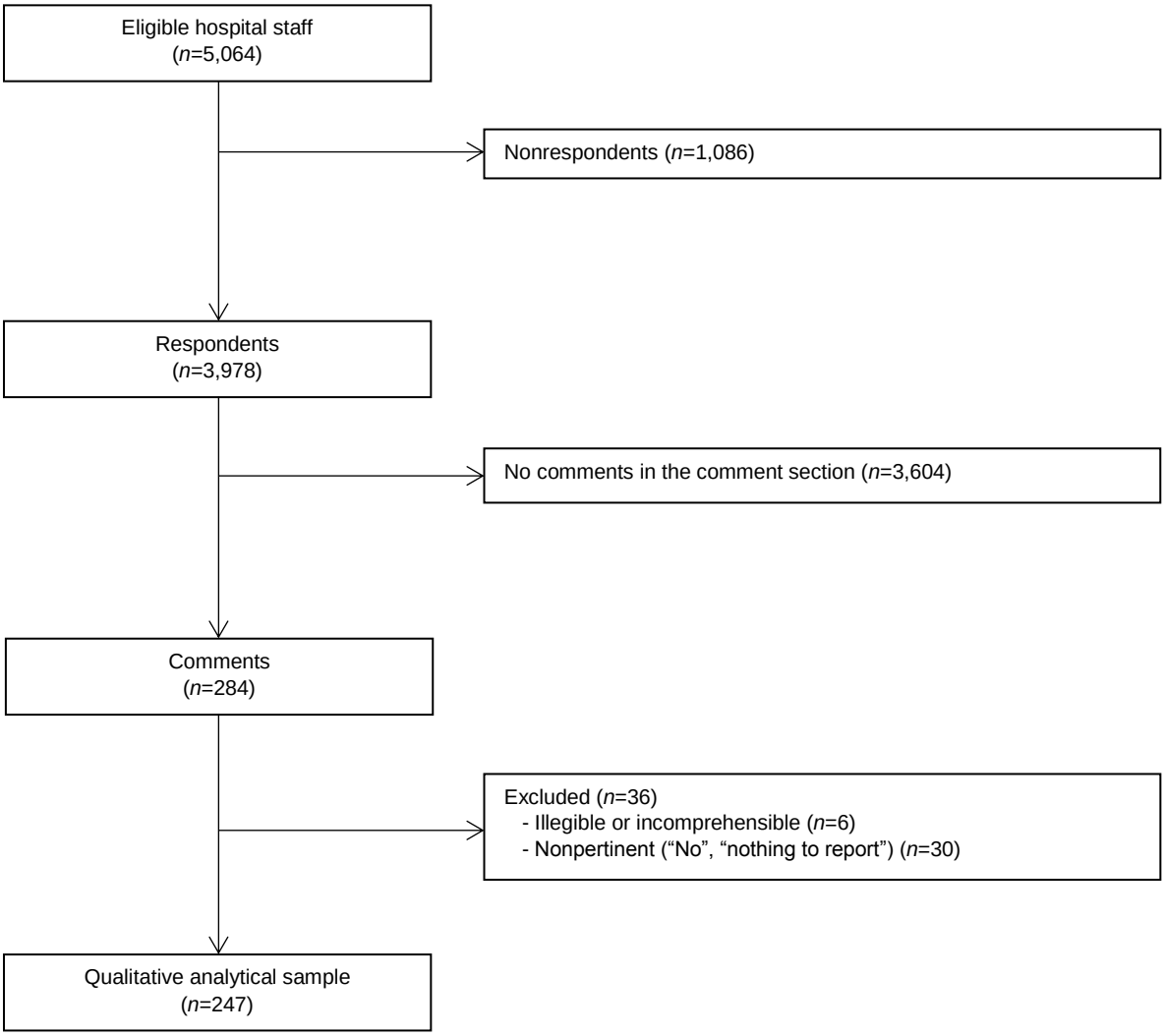
Background respondent characteristics and open comments were reported as numbers and percentages. We computed the individual means across three of the four items in a dimension to obtain the HSOPS dimension scores (range, 1–5). [11] Then we compared the HSOPS mean composite score values across subgroups of respondents defined by the presence or absence of an open comment. The differences between score means were tested using the Student *t*-test. In addition, we quantified the mean differences using the Cohen *d* effect size. Two-sided *P*-values lower than 0.05 were

considered statistically significant. All analyses were performed using Stata Version 14.0 (Stata Corporation, College Station, TX, USA).

Results

The response rate to the HSOPS survey was 78.6% (*n*=3,978). Among the questionnaires returned, 284 had text in the comments area; 36 questionnaires were excluded according to our exclusion criteria (Fig. 1).

Figure 1: Participation and comments in the HSOPS survey



The majority of the respondents (80.6% for the HSOPS survey, 87.5% for the comments, and 68.4% for the interviews) were women. In both the returned questionnaires and the comments, the majority of respondents were under 46 years of age (63.2% for the HSOPS, 60.9% for the comments). The subject's age was not requested during the interviews (Table 1).

Table 1. Respondent characteristics

Characteristics *N* (%)	HSOPS *n*=3978	Analyzable comment *n*=247	Interviews *n*=19
Female	3,016 (80.6)	209 (87.5)	13 (68.4)
Age class (years)			
Up to 35	1,406 (37.5)	82 (34.2)	
36–45	966 (25.7)	64 (26.7)	
46–55	967 (25.7)	66 (27.5)	
56 or older	415 (11.1)	28 (11.7)	
Occupational group			
Head nurse and nurse	1,386 (36.3)	94 (38.4)	9 (47.4)
Nursing assistant	708 (18.6)	61 (24.9)	1 (5.3)
Physician	436 (11.4)	17 (6.9)	7 (36.8)
Other healthcare provider	124 (3.3)	8 (3.3)	2 (10.5)
Administrative	331 (8.7)	33 (13.5)	0 (0)
Technical	378 (9.7)	3 (1.2)	0 (0)
Other	450 (11.8)	29 (11.8)	0 (0)

* Values were missing for gender (HSOPS, *n*=147; Analyzable comment, *n*=8); age (HSOPS, *n*=224; Analyzable comment, *n*=7); and occupational group (HSOPS, *n*=75; Analyzable comment, *n*=2).

HSOPS dimension scores

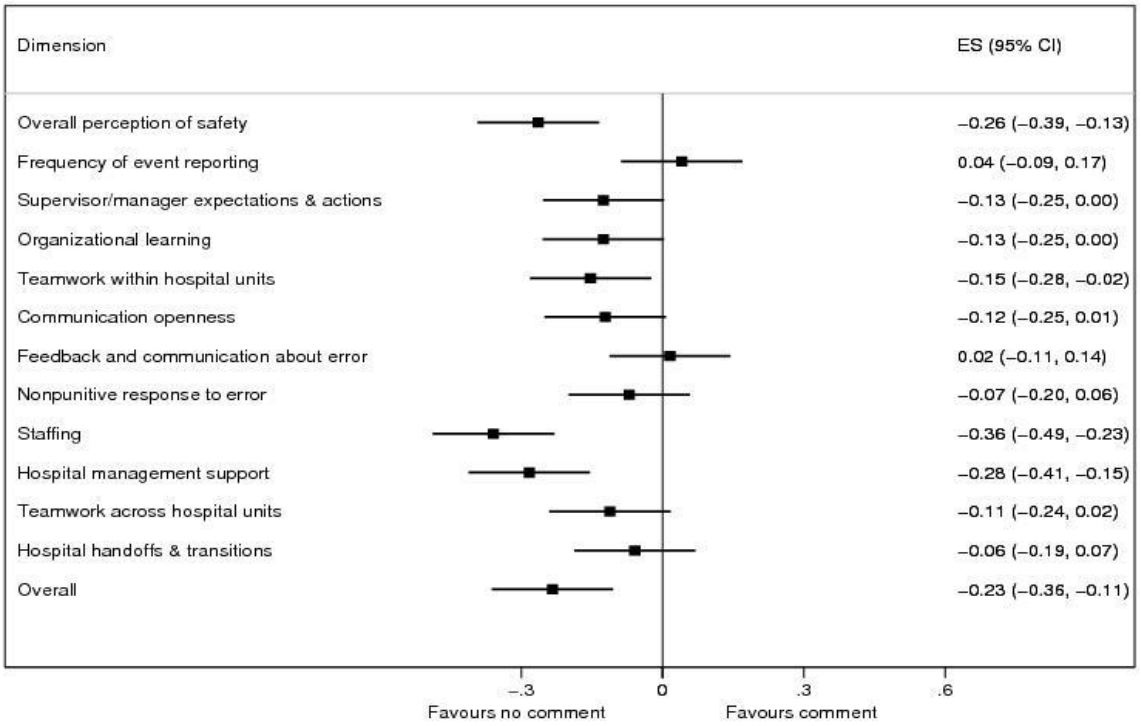
The mean dimension scores ranged from 2.67 for the dimension with the lowest score (hospital management support) to 3.54 positive answers for the dimension with the highest score (teamwork within hospital units) (Table 2). The two other dimensions with unfavorable mean scores concerned "staffing" and "hospital handoffs & transitions" (2.88).

Table 2. Results of the HSOPS

Dimension	Score	Rank
D1: Overall perceptions of safety	3.21	7
D2: Frequency of event reporting	3.38	5
D3: Supervisor expectations & actions	3.50	2
D4: Organizational learning	3.40	4
D5: Teamwork within hospital units	3.54	1
D6: Communication openness	3.50	2
D7: Feedback and communication about error	3.38	5
D8: Nonpunitive response to error	2.94	9
D9: Staffing	2.88	10
D10: Hospital management support	2.67	12
D11: Teamwork across hospital units	3.05	8
D12: Hospital handoffs & transitions	2.88	10

Figure 2 shows the association between the presence of an open comment and the patient safety culture dimensions scores. Compared with the absence of a comment, the presence of a comment was associated with a statistically significant lower patient safety culture for four of the 12 HSOPS dimensions (overall effect size, −0.23; 95% confidence interval, −0.36 to −0.10, *P*<0.001). The three greatest differences in HSOPS score were related to "overall perception of safety," "staffing," and "hospital management support."

Figure 2. Comparison of HSOPS scores according to the presence of open comments (effect size).



Classification of comments

The 247 analyzable comments were independently classified by three surveyors, who then agreed on the following classification: questionnaire (101 occurrences), staffing and hospital management support (98), organization (41), adverse event reporting and risk management support (25), premises and equipment (16) and staff safety (5). (Table 3)

Table 3. Categorization of comments (*n*=247) by topics (*n*=286) and key words

Subject	N	(%)	Key words (occurrence)
Questionnaire	101	(35)	Unsuitable questionnaire (51) Complementary information on an answer (17) Worries about anonymity (9) Questions about the objectives of the survey (8) Problems responding (7) About the questionnaire: satisfaction (5) dissatisfaction (4)
Staffing and hospital management support	98	(34)	Fatigue (16) and stress (11) including exhaustion (5) and professional distress (1) Difficulty replacing missing staff (12) interim staff (6) Time and speed imperatives (7) Hierarchy (4) including positive support (1) lack of trust and support (3) Staff concerned by patient safety (3) Problems with overtime (3) Difficulty taking time off (3) Difficulties/lack of time to participate in safety organizations or training (3)
Organization	41	(14)	Cooperation/coordination between departments (10) Transmission/communication (10) Teamwork (9): lack (6) or appropriate (3) Beds in the corridor (3) Patient transfer (2) from emergency unit (1) or to diagnostic imaging department (1) Redeployment of staff (2) Lack of communication on hospital's actions (2)

Adverse event reporting and risk management	25	(9)	Reporting (13) including report not taken into consideration (7) no feedback on reporting (4) lack of time (3) fear of reporting (3) Involvement in patient safety management system (9), including Morbidity and Mortality Conference (4) and Experience Feedback Committee (5) Blame culture (3)
Premises / equipment	16	(6)	Premises (5) including Patient comfort (5) risks for the patient (3) patient privacy (2) hygiene (1) Equipment (13) including computer system (5) lack of equipment (4) inadequate or defective equipment (3)
Staff safety	5	(2)	Violent patient (3) Risks related to equipment (1) Risks related to night work (1)

Results of analysis of comments and interviews

Questionnaire and survey

One hundred one comments raised issues about the questionnaire or the survey. Most of these comments declared the questionnaire to be unsuitable. Some of them explained the questionnaire was unsuitable for their occupational group (particularly for secretaries). Other comments considered it to be unsuitable to evaluate patient safety. Seven respondents had difficulties answering one or several items. Seventeen respondents used the comments to give complementary information about an answer. Eight respondents had questions about the objective of the survey and nine had worries about anonymity. We did not investigate this topic further in the interviews.

Staffing and hospital management support: "Chronic understaffing in the department seriously affects patient safety" (Comment)

Staffing was a frequently raised topic in the comments: all comments in this topic complained about understaffing and heavy workload. Comments indicated that this understaffing was responsible for fatigue, stress, and a decrease in patient safety. Comments also reported problems stemming from missing staff (e.g., sick leave). The interviews confirmed the problem of understaffing, which was nearly always mentioned. Interviewees from all occupational groups reported understaffing and the resulting problems (stress, fatigue, absenteeism), linking it with a decrease in quality and safety in patient care.

Several professionals interviewed clearly and spontaneously talked about professional distress and even discontent stemming from a desire to resigning from their job or from seeing colleagues resign, which led to a problem of skill preservation within teams. Some professionals declared feeling that their work no longer had meaning for them. Even more interviewees declared feeling dissatisfied with their work: they considered that the work conditions made it impossible to give patients proper care.

"I don't say it's systematic but sometimes the working conditions affect the quality of the care we provide. I know that some colleagues have quit their job because of this, because they couldn't offer an acceptable quality of care and they preferred to leave and transfer to another hospital that allowed them to provide a better quality of care" (Physician 3).

Some interviewees declared the problem of skill preservation and staff departure was significant in their department, with nearly 20% of a team leaving the hospital per year: "in a team of 58 people, 10 to 12 people quit every year" (Head nurse 5). The difficulty of replacing missing staff worsened the understaffing problem.

The staff members leaving comments on the questionnaires reported the risks associated with lack of adequate time. They also explained that management seemed more concerned with budgetary objectives than quality and safety. One comment explained that "The hospital managers have some good ideas, such as the implementation of training courses. But once again, understaffing shows that safety is not really a priority" (Comment). These results prompted us to investigate this issue further in the interviews.

During the interviews, staff members mentioned problems with management, explaining that management had contradictory objectives and preferred to take budgetary and economic objectives into consideration rather than quality and safety. Moreover, several professionals reported that there was a lack of consideration, respect (especially in emails), and listening on the part of management. More generally, the interviewees underlined a lack of resources in the hospital. The management was also said to be cut off from the realities of the workplace.

"It is a daily struggle against management: they want to reduce our resources while asking us to increase activity" (Physician 2).

This increased activity without increased resources resulted in problematic situations. Some staff members explained that the availability of beds did not always allow transferring a patient to the appropriate department. The interviewees explained that in these situations they did their best for the patients but admitted they did not have the skills required to give them appropriate care. The lack of available beds led to installing additional beds in the corridors. Corridor beds were reported to be a

significant problem and to create significant risks since this practice increased the workload of all staff members and equipment using electrical power could not be used for these beds.

Organization: "Better communication between departments would significantly reduce adverse events" (Comment)

Communication and organizational issues were often reported in the comments, mainly concerning transmission of information and transferring patients. The interviews confirmed that handoffs and transitions in care were sometimes difficult. They pointed out that information was sometimes incomplete. There was also a problem of communication within departments. Interviewees explained that information transmission and communication was well organized by protocols, but they underlined that this was not always the case.

"We organized and protocoled communication and patient transfer because things are very difficult if this is not done" (Physician 5).

Transfers from the emergency department were sometimes made without reevaluation of the patient's condition. Beds in the corridors, partly due to lack of bed availability and resources, were also the consequence of a lack of cooperation between the receiving department and the emergency department. Interviewees conceded that time constraints did not always allow correctly transmitting information about the patient, especially in a department such as the emergency department where patient influx was not predictable.

"When patients come from the emergency department, they are not always reevaluated before their transfer… And sometimes we are surprised because the patients are not like we have been told because their condition has evolved…" (Physician 4).

Adverse event reporting and risk management: "we stopped reporting adverse events because we feel that there is no feedback." (Comment)

Several comments mentioned issues related to adverse event reporting. The two most frequently cited problems were the lack of feedback following self-reporting and the lack of time to report. Other healthcare professionals declared a fear of punishment related to medical error, resulting in underreporting serious adverse events. Interestingly, adverse event reporting was considered as a source of interprofessional conflict:

"In reporting incidents, we could be more tactful, respectful and discreet towards one another and make decisions within nonviolent communication" (Comment).

"Adverse event reporting is too time-consuming and mistakes don't receive the attention they deserve" (Comment).

Some professionals indicated they were involved in risk management activities and related this kind of activity to a good patient safety culture. They cited two programs designed to involve the medical team in patient safety management: experience feedback committees (EFC) and morbidity and mortality conferences (MMC):

"In this department, many risk management programs have been set up: EFCs and MMCs". (Comment).

"Patient safety, EFCs, and quality procedures have been developed and implemented satisfactorily in my department" (Comment).

On the other hand, a nurse regretted that nurses were not allowed the time to participate in EFC or MMC meetings:

"For a nurse, it is unfortunately impossible to take the time to participate in EFC or MMC activities" (Nurse 2).

Equipment: "The equipment is increasingly fragile and ill-suited." (Comment)

Several comments reported different equipment problems.

"The equipment is increasingly fragile and ill-suited. A recent example in the department is a perfusion line that spontaneously broke, fortunately before it was set up on the patient" (Comment).

These issues were also raised in the interviews. They concerned the quality of the equipment, the quantity of equipment, and ill-suited equipment. Staff members also complained about repair delays. Interviewees felt the choice of the equipment was made based on economic criteria and not quality or suitability.

"It's a problem when the equipment you are working with and are accustomed to is replaced with something worse, because most of the time when equipment is changed it is replaced with something of lower quality. This is really a disgrace" (Nurse 1).

The employees interviewed declared that these problems involve a real risk for the patient. It was also explained that even when it does not put the patient at risk it deteriorates the quality of the care and comfort for both patient and staff.

Staff Safety

Staff safety issues were raised by nine comments and in some of the interviews. The main risks were musculoskeletal injuries, needle stick and blood exposure, and finally a risk of violence in some departments, notably the emergency department.

Discussion

Through the analysis of the comments and interviews, we identified several structural failures that could explain the very low scores obtained for certain of the HSOPS dimensions. The lack of resources was identified by all categories of staff members. This could be strongly related to the low dimension scores for "hospital management support" and "staffing." Moreover, the vast majority of comments reported important problems of communication between units and departments,

related to the penultimate dimension score for “hospital handoffs & transitions.” Highlighting concrete failures in the organization and management, the open comments analysis provides a better understanding of the dimension scores computed from the 42 closed questions of the HSOPS. Too rarely done in this type of survey, consideration of the open comments could also help complete the closed questionnaire and target concrete failures in the hospital organization and management.

First, the comments on the survey and the staff members interviewed reported a global lack of resources. Insufficient resources concerned all aspects of their work: staffing, equipment, resources devoted to training, and bed availability. This resulted in an increased level of fatigue, stress, and mental load. Understaffing favored task interruptions, which were considered a significant risk for patients. Insufficient training was said to increase the risk of error but also the level of stress. Bed availability was raised by the interviewees, which degraded patient safety by creating situations in which departments had to handle patients they were not trained or equipped to handle or by making it necessary to resort to corridor beds. The negative effects of lack of training, staffing and equipment issues, as well as the related fatigue, stress, and mental load are widely known and these factors tend to decrease patient safety and the quality of care.[12-15] The deterioration of working conditions (and the subsequent quality of working life) as well as the quality and safety of care, as perceived by the staff members interviewed, encourage medical and paramedical staff to quit their jobs. Several interviewees admitted their desire to resign. High turnover has an impact on cost and skill preservation and creates disorganization. Indeed, departing staff members need to be replaced and new recruits need to be trained so they can work well within the department. Both replacement and training have significant costs.[12,16] The impact of a high turnover rate and staff dissatisfaction over patient satisfaction is not to be neglected and is associated with a decrease in patient safety and quality of care.[17,18]

Second, a problem of communication and cooperation between units and departments was identified. This problem resulted in loss of information during transmission or an absence of transmission. It also participated in patient transfer problems. Although several comments argued that patient transfers were guided by defined written procedures, it was obvious that these procedures did not produce solutions to limit the problem occurring during transitions between hospital departments. Therefore, the comments analysis made it possible to focus on concrete problems caused by these procedures, whether related to a lack of dissemination to the wards or to the impossibility of dealing with complex transfer situations. Finally, this tended to underline that other enhancements can be explored.

Third, several open comments reported difficulties related to adverse event reporting, which is of utmost importance because it allows the organization to learn from its errors. Nevertheless, several studies demonstrated that many incidents were not actually reported [19,20]. The majority of reporting systems are based on healthcare professionals’ self-reports, and it is difficult to properly identify the obstacles to self-reporting. Our qualitative data confirmed that lack of feedback and time constraints were the two main barriers. Time constraints could be related to insufficient resources; the lack of feedback was closely related to the organization of risk management. Indeed, the vast majority of French hospitals have a centralized incident-reporting system. Interestingly, this perception is not collected with the closed questions of the HSOPS. The use of open comments could thus help identify specific organizational failures and elaborate tailored solutions to enhance the management of patient safety. Improving feedback after incident reporting does not require major resources and could be accomplished through organizational changes. For example, the staff interviewed raised the possibility of decentralizing risk management in hospital departments by promoting programs directly involving the medical teams in patient safety, such as experience feedback committees and the morbidity and mortality conferences. [21,22]

Besides targeting concrete ways to improve organization of the hospital site, the open comments and interviews provided a number of elements critical of hospital management’s support for patient safety. The very low score for the “hospital handoffs and transitions” dimension is observed in the majority of HSOPS surveys in both Europe and the United States, showing a weakness shared by all organizations.[5,7,23-27] Interestingly, this does not extend to the “hospital management support” dimension, which is nearly always one of the lowest dimension scores in the HSOPS survey conducted in Europe, whereas it is one of the highest dimensions in the United States.[5,7,23-27] Although true contextual differences between the US and Europe might explain inconsistency, the results reported herein have provided food for thought on the priority policy of a French university hospital, which partly explains this difference observed for the “hospital management support” dimension.[28] According to the professionals interviewed, the lack of resources was partly due to management policy, which demands an increase in hospital activity without increasing the available resources, including bed capacity. However, we cannot exclude that the open comments section was used to express complaints that went far beyond the concerns of safety culture. Indeed, the surprising level of anger and resentment of certain comments might suggest that the HSOPS-based survey was used as a sounding board for some of the respondents. This hypothesis raises questions about the reliability of the HSOPS results. Whether the HSOPS score differences observed

between the respondents who wrote a comment or did not reflected true patient safety culture differences or information bias remains unclear. Although speculative, these questions show that the qualitative approach adds richness to the classical quantitative approach used for the HSOPS.

This study was conducted in a single hospital, its main limitation. Therefore, not all results can be applied to other sites. However, some of the findings may be applied to sites with a comparable context. A second limitation is the small number of interviews. However, these interviews were conducted in addition to the comments analysis. In that sense, we had rich data and were able to reach data saturation. Finally, we did not explore the different perceptions depending on professional categories. Yet it cannot be excluded that true differences in education, activity, and experience specific to different professions could strongly influence the respondent's perception. Despite this limitation, the study design followed the AHRQ recommendations, which developed the HSOPS questionnaire to conduct hospital-wide surveys, assessing staff awareness of patient safety independently of professional categories.

Conclusion

Through a hospital-wide survey, the analysis of open comments provided a better understanding of the results of the HSOPS closed questions. The lack of resources was perceived as the major barrier to improving patient safety and was strongly related to the support provided by hospital management, which was the worst dimension computed by the closed questions. The open comments also identified a number of specific failures related to patient transfers and risk coordination. Mixing qualitative and quantitative approaches in the HSOPS-based surveys could help not only to achieve a better understanding of the HSOPS scores, but also to develop corrective actions to enhance the patient safety culture of healthcare professionals.

Funding

This work was supported by a grant from the French Ministry of Health (Programme de recherche sur la performance du système de soins-PREPS 12-026-0001).

Acknowledgments

This work was developed within the framework of the Grenoble Alpes Data Institute, which is supported by the French National Research Agency under the "Investissements d'avenir" program (ANR-15-IDEX-02).

We acknowledge Linda Northrup (English Solutions, Voiron, France) for her assistance in editing the manuscript in English.

References

1. Makary MA, Daniel M (2016) Medical error-the third leading cause of death in the US. BMJ 353: i2139.
2. Donaldson LJ, Fletcher MG (2006) The WHO World Alliance for Patient Safety: towards the years of living less dangerously. Med J Aust 184: S69-72.
3. Leotsakos A, Caisley L, Karga M, et al. (2009) High 5s: addressing excellence in patient safety. World Hosp Health Serv 45: 19-22.
4. Banihashemi S, Hatam N, Zand F, et al. (2015) Assessment of Three "WHO" Patient Safety Solutions: Where Do We Stand and What Can We Do? Int J Prev Med 6: 120.
5. Sorra JS, Dyer N (2010) Multilevel psychometric properties of the AHRQ hospital survey on patient safety culture. BMC Health Serv Res 10: 199.
6. Sorra J NV (2004) Hospital Survey on Patient Safety Culture (Prepared by Westat, under Contract No. 290-96-0004). AHRQ Publication No. 04-0041. Rockville, MD: Agency for Healthcare Research and Quality.
7. Occelli P, Quenon JL, Kret M, et al. (2013) Validation of the French version of the Hospital Survey on Patient Safety Culture questionnaire. Int J Qual Health Care 25: 459-468.
8. Maxwell JA (2008) Designing a qualitative study. The SAGE handbook of applied social research methods 2: 214-253.
9. Steckler A, McLeroy KR, Goodman RM, et al. (1992) Toward integrating qualitative and quantitative methods: an introduction. Health Educ Q 19: 1-8.
10. Alleaume S, Boisguérin B, Brilhault G, et al. (2014) Le panorama des établissements de santé édition 2014. DREES, French Ministry of Health.
11. Giai J, Boussat B, Occelli P, et al. (2017) Hospital survey on patient safety culture (HSOPS): variability of scoring strategies. Int J Qual Health Care 29: 685-692.
12. Aiken LH, Clarke SP, Sloane DM, et al. (2002) Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. JAMA 288: 1987-1993.
13. Donaldson L (2002) An organisation with a memory. Clin Med (Lond) 2: 452-457.
14. Reason J (2000) Human error: models and management. BMJ 320: 768-770.
15. Rogers AE, Hwang WT, Scott LD, et al. (2004) The working hours of hospital staff nurses and patient safety. Health Aff (Millwood) 23: 202-212.
16. Waldman JD, Kelly F, Arora S, et al. (2004) The shocking cost of turnover in health care. Health Care Manage Rev 29: 2-7.
17. McHugh MD, Kutney-Lee A, Cimiotti JP, et al. (2011) Nurses' widespread job dissatisfaction, burnout, and frustration with health benefits signal problems for patient care. Health Aff (Millwood) 30: 202-210.
18. Vahey DC, Aiken LH, Sloane DM, et al. (2004) Nurse burnout and patient satisfaction. Med Care 42: II57-66.
19. Blendon RJ, DesRoches CM, Brodie M, et al. (2002) Views of practicing physicians and the public on medical errors. N Engl J Med 347: 1933-1940.
20. Sarvadikar A, Prescott G, Williams D (2010) Attitudes to reporting medication error among differing healthcare professionals. Eur J Clin Pharmacol 66: 843-853.
21. Boussat B, Seigneurin A, Giai J, et al. (2017) Involvement in Root Cause Analysis and Patient Safety Culture Among Hospital Care Providers. J Patient Saf.
22. Francois P, Prate F, Vidal-Trecan G, et al. (2016) Characteristics of morbidity and mortality conferences associated with the implementation of patient safety improvement initiatives, an observational study. BMC Health Serv Res 16: 35.
23. Agnew C, Flin R, Mearns K (2013) Patient safety climate and worker safety behaviours in acute hospitals in Scotland. J Safety Res 45: 95-101.
24. Ballangrud R, Hedelin B, Hall-Lord ML (2012) Nurses' perceptions of patient safety climate in intensive care units: a cross-sectional study. Intensive Crit Care Nurs 28: 344-354.
25. Bagnasco A, Tibaldi L, Chirone P, et al. (2011) Patient safety culture: an Italian experience. J Clin Nurs 20: 1188-1195.
26. Olsen E (2007) [Workers' perceptions of safety culture at a hospital]. Tidsskr Nor Laegeforen 127: 2656-2660.
27. van Noord I, de Bruijne MC, Twisk JW (2010) The relationship between patient safety culture and the implementation of organizational patient safety defences at emergency departments. Int J Qual Health Care 22: 162-169.
28. Fielding JE, Lancry PJ (1993) Lessons from France--'vive la difference'. The French health care system and US health system reform. JAMA 270: 748-756.

4. CHAPITRE III: Association entre les Comités de Retour d'Expérience et la culture de sécurité des professionnels de santé

Le fonctionnement des CREX et les propriétés de l'outil de mesure de la culture de sécurité ayant été analysés et décrits, il s'agissait désormais de répondre à l'objectif principal du projet ACREX, à savoir explorer les relations entre les CREX et la culture de sécurité des professionnels de santé. La première hypothèse était que les participants aux CREX avaient une culture de sécurité plus élevée que les non participants. La seconde hypothèse était que les professionnels de santé travaillant dans un service où était implanté un CREX avaient une culture de sécurité plus développée que les soignants travaillant dans un service sans CREX.

Cette étude a fait l'objet d'un article publié dans Journal of Patient Safety, intitulé « Involvement in root cause analysis and patient safety culture among hospital care providers » (Boussat, Seigneurin et al. 2017).

Objectif : L'objectif principal de cette étude était d'étudier les différences de culture de sécurité des professionnels de santé, selon la participation ou non à un CREX, et selon l'implantation ou non d'un CREX dans le service.

Méthode : Nous avons utilisé les données de l'enquête culture sécurité menée au CHU de Grenoble, à partir du questionnaire HSOPS. Trois variables ont été utilisées pour caractériser le lien individuel des 3888 répondants avec les CREX : 1/ la participation à un CREX (déclarative, recueillie à partir d'une question ajoutée à la fin du questionnaire HSOPS) ; 2/ la présence d'un CREX dans le service (variable générée à partir de la base de donnée administrative des programmes d'évaluation des pratiques professionnelles) ; 3/ la présence d'un CREX qualifié de productif, défini comme un CREX produisant au moins 5 actions par an (variable générée à partir des données de l'étude sur le fonctionnement global des CREX de l'établissement, présentée dans le chapitre I). Les non réponses à l'item questionnant la participation à un CREX étaient considérées comme négatives (3%). Les réponses manquantes

aux 42 items servant au calcul des scores dimensionnels étaient imputées en utilisant une méthode d'imputation multiple (predictive mean matching). Les scores des 12 dimensions de culture de sécurité étaient agrégés en utilisant la moyenne des réponses individuelles (Giai, Boussat et al. 2017). Nous avons comparé les valeurs des scores entre les sous-groupes de répondants, tels que définis par les 3 variables caractérisant le lien individuel avec les CREX. Les différences de scores étaient testées en utilisant le test *t* de Student, et quantifiées en utilisant la taille d'effet *d* de Cohen. Enfin, nous avons réalisé une analyse de variance multivariée (MANOVA) pour examiner les différences des scores des 12 dimensions selon chacune des 3 variables caractérisant le lien individuel avec les CREX, indépendamment du genre, de la profession, et de la spécialité.

Résultats : 440 répondants déclaraient avoir participé à un des vingt CREX de l'établissement (11.3%). Comparativement aux non participants, les participants aux CREX étaient plus souvent des médecins (22% vs 10%) et travaillaient plus souvent dans des spécialités médicales et pédiatriques (36% vs 29%). La participation à un CREX était associée à une culture de sécurité plus élevée pour 9 des 12 dimensions du HSOPS (taille d'effet moyen des 12 dimensions = 0.31, IC95% 0.21 à 0.41, P<0.001) (figure 2). La MANOVA indiquait que les 12 scores dimensionnels étaient supérieurs pour les participants aux CREX (P<0.0001), indépendamment du genre, de la profession et de la spécialité. Les coefficients du modèle de régression multivarié étaient consistants pour 8 des 9 scores dimensionnels significatifs en analyse univariée. Les différences de scores les plus importantes concernaient les dimensions « Retour d'information et communication des erreurs » (coefficient de régression = 0.25, CI95% = 0.17-0.25), « Organisation apprenante et amélioration continue » (coefficient de régression = 0.22, CI95% = 0.16-0.29), et « Réponse non punitive à l'erreur » (coefficient de régression = 0.18, CI95% = 0.11-0.26). La présence d'un CREX dans le service n'était quant à elle pas associée à une meilleure culture sécurité (taille d'effet moyen des 12 dimensions =

0.31, IC95% -0.04 à 0.00, P=0.32). Les répondants qui travaillaient dans un service comportant un CREX productif avaient une culture de sécurité significativement supérieure, comparativement aux répondants travaillant dans un service comportant un CREX produisant moins de 5 actions correctrices par an (taille d'effet moyen des 12 dimensions = 0.19, IC95% 0.10 à 0.27, P<0.001). Cette différence était confirmée en multivarié. Enfin le pourcentage de professionnels rapportant déclarer annuellement au moins un événement indésirable associé aux soins, était significativement plus élevé chez les participants aux CREX que chez les non participants (65% versus 51%, P>0.001).

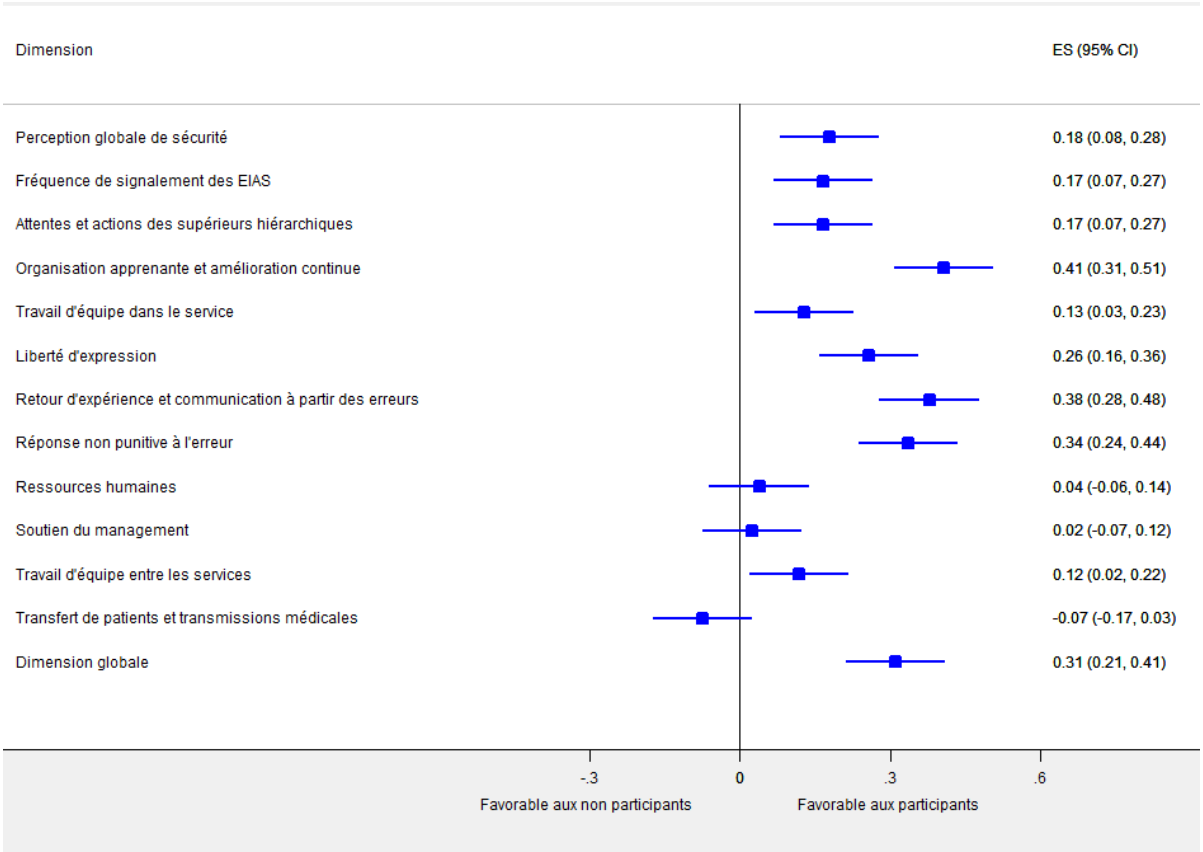


Figure 2 Comparaison des scores de culture de sécurité selon la participation à un CREX (Effect size)

Discussion et conclusions : A partir de cette étude réalisée auprès de l'ensemble des professionnels de santé d'un centre hospitalier universitaire, nous avons montré que la

participation aux activités d'un CREX était associée à une plus grande culture de sécurité. La participation à un CREX était significativement associée à 9 des 12 dimensions de la culture sécurité. Les 3 plus grandes différences concernaient des dimensions clefs de la culture de sécurité, l'« Organisation apprenante et amélioration continue », la « Réponse non punitive à l'erreur » et le « Retour d'information et communication des erreurs ». Malgré le design observationnel de cette étude, ces résultats semblent pertinents, étant donné qu'ils confirment l'hypothèse initiale ayant conduit au développement des CREX, à savoir que l'implication directe des professionnels de santé dans la gestion de la sécurité des soins a un impact positif sur leur culture de sécurité. De plus, ces résultats sont cohérents avec la littérature, qui rapporte une association en faveur d'une plus grande culture de sécurité pour des dispositifs comparables aux CREX, comme les Comprehensive Unit-based Safety Programm (CUSP) aux Etats-Unis (Weaver, Lubomksi et al. 2013, Weaver, Lofthus et al. 2015). Les professionnels participants aux CREX étaient plus conscients de l'intérêt du retour d'expérience et dépassaient la culture du blâme. Ces deux dimensions font échos au fonctionnement du CREX, qui met au cœur de son activité l'analyse des causes racines des signalements indésirables déclarés au niveau du service. Cette meilleure perception de l'erreur, comme étant utile pour aider l'organisation à s'améliorer, semblait également confirmée par le pourcentage plus élevé de déclarants d'EIAS chez les participants aux CREX. Trois dimensions ne différaient pas significativement selon la participation ou non à un CREX : « Ressources humaines », « Soutien du management pour la sécurité des soins » et « Transmissions et transferts ». L'absence de différences pour les dimensions concernant les ressources humaines et le soutien du mangement peuvent être expliquées en partie par des causes structurelles, compte-tenu des contraintes budgétaires auxquelles sont confrontés les hôpitaux français. Mais cela montre tout de même une certaine inefficacité des CREX, dont le périmètre d'action et les moyens sont limités. Les dimensions support du management et transferts sont considérées comme étant au niveau de l'hôpital

(hospital-level processes), comparativement aux autres dimensions situées au niveau du service (unit-level processes). On approche ici l'une des principales limites des CREX, qui peine à agir sur les EIAS dépassant le périmètre du service. Or, une grande partie des EIAS transcendent ce périmètre, et se situent au niveau des interfaces, lors des mutations de patients par exemple. Cette limite se rapproche également des critiques émises contre la pratique de l'analyse des causes racines effectuée par des professionnels de santé (Peerally, Carr et al. 2017). La seconde hypothèse que nous avions formulée initialement, à savoir une meilleure culture de sécurité pour les professionnels travaillant dans un service possédant un CREX n'était pas vérifiée. L'absence de diffusion de culture de sécurité à l'ensemble de l'équipe peut être révélatrice là aussi d'un cloisonnement des activités des CREX, et d'un manque de communication de la part des participants vers les autres membres des équipes médicales.

Involvement in Root Cause Analysis and Patient Safety Culture Among Hospital Care Providers

Bastien Boussat, MD,† Arnaud Seigneurin, MD, PhD,*† Joris Giai, MD,‡ Kevin Kamalanavin, MSc,§ José Labarère, MD, PhD,*†‖ and Patrice François, MD, PhD*†*

Background: The experience feedback committee (EFC) is a tool designed to involve medical teams in patient safety management, through root cause analysis within the team.

Objective: The aim of the study was to determine whether patient safety culture, as measured by the Hospital Survey on Patient Safety Culture (HSOPS), differed regarding care provider involvement in EFC activities.

Methods: Using the original data from a cross-sectional survey of 5064 employees at a single university hospital in France, we analyzed the differences in HSOPS dimension scores according involvement in EFC activities.

Results: Of 5064 eligible employees, 3888 (76.8%) participated in the study. Among the respondents, 440 (11.3%) participated in EFC activities. Experience feedback committee participants had a more developed patient safety culture, with 9 of the 12 HSOPS dimension scores significantly higher than EFC nonparticipants (overall effect size = 0.31, 95% confidence interval = 0.21 to 0.41, $P < 0.001$). A multivariate analysis of variance indicated that all 12 dimension scores, taken together, were significantly different between EFC participants and nonparticipants ($P < 0.0001$), independently of sex, hospital department, and healthcare profession category. The largest differences in scores related to the "feedback and communication about error," "organizational learning," and "Nonpunitive response to error" dimensions. The analysis of the subgroup of professionals who worked in a department with a productive EFC, defined as an EFC implementing at least five actions per year, showed a higher patient safety culture level for seven of the 12 HSOPS dimensions (overall effect size = 0.19, 95% confidence interval = 0.10 to 0.27, $P < 0.001$).

Discussion and Conclusions: Participation in EFC activities was associated with higher patient safety culture scores. The findings suggest that root cause analysis in the team's routine may improve patient safety culture.

Key Words: patient safety, root cause analysis, safety culture, adverse event, hospital departments

(*J Patient Saf* 2017;00: 00–00)

Despite considerable initiatives to enhance patient safety in healthcare, medical errors reached the third leading cause of death in the United States, with an estimated 251,454 deaths per year.[1] In 2008, a World Health Organization Patient Safety group (the Research Priority Setting Working Group) established the 20 research priority areas for developed countries.[2] "Poor safety culture and blame-oriented processes" was ranked third among these priorities. Because adverse events result from multiple organizational failures, it is imperative to move toward a collaborative teamwork culture, with all actors sharing the same vision and goals and considering themselves responsible for delivering safe care.[3] One of the ways to fight medical errors is therefore the development of a patient safety culture shared by all caregivers.[4]

In the last decade, root cause analysis (RCA) has become the main process to investigate medical errors and attempts to prevent future adverse events.[5] Root cause analysis provides undeniable benefits, including increased awareness of faulty processes, and leads to changing organizations despite cultural resistance. Therefore, organizations in charge of hospital accreditation, such as the Joint Commission in the United States and the French National Health Authority in France (Haute Autorité de Santé), have imposed RCA as the reference method to investigate adverse events.[6,7] However, its use is now controversial, because several studies have called its effectiveness into question.[8,9]

In France, a program called the experience feedback committee (EFC) was created in 2005, with the aim of enhancing patient safety within medical teams, involving the team in RCA.[10,11] The method was adapted to healthcare facilities with the assistance of Air France Consulting and was initially implemented in a radiotherapy unit. Briefly, the EFC is a team composed of professionals representing the diversity of the professions encountered in medical units, who examine adverse events on a monthly basis. Conducting RCA within the team, the EFC aims to implement corrective actions after multidisciplinary consensus. See the method section for further detail on the EFC framework.

Because several studies reported good compliance on the part of caregivers in terms of attending these meetings, the Haute Autorité de Santé promoted the wide implementation of this management system in all French hospitals and introduced its use as a criterion for the national accreditation process in 2010.[7,10,11] The early literature showed promising results for the EFC, in favor of key aspects of patient safety, such as enhancement of teamwork and the implementation of actions that were well adapted to specific contexts, such as emergency or psychiatry departments.[10,11] Despite its wide implementation, the EFC still lacks scientific evaluation.

The aim of this study was to determine whether patient safety culture, as measured by the Hospital Survey on Patient Safety Culture (HSOPS), differed regarding care provider involvement in EFC activities.[12] We hypothesized that 1/EFC participants had higher dimensions scores on patient safety culture than the nonparticipants and 2/caregivers working in a department involving an EFC had higher dimension scores than those working in departments without an EFC.

From the *Quality of Care Unit, Grenoble University Hospital; †TIMC UMR 5525 CNRS, Université Grenoble Alpes, Grenoble; ‡Service de Biostatistique, Hospices Civils de Lyon, Laboratoire de Biométrie et Biologie Evolutive, UMR 5558 CNRS, Lyon; §Department of Political Science, Institute of Political Studies; and ‖INSERM CIC 1406, Grenoble, France.

Correspondence: Bastien Boussat, MD, Service d' Evaluation Médicale, Pavillon Taillefer, CHU CS 10-217, F-38043, Grenoble Cedex, France (e-mail: bboussat@chu-grenoble.fr).

The authors disclose no conflict of interest.

This work was supported by a grant from the French Ministry of Health (Programme de recherche sur la performance du système de soins-PREPS 12-026-0001).

This work was developed within the framework of the Grenoble Alpes Data Institute, which is supported by the French National Research Agency under the "Investissements d'avenir" program (ANR-15-IDEX-02).

Supplemental digital contents are available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.journalpatientsafety.com).

Copyright © 2017 Wolters Kluwer Health, Inc. All rights reserved.

METHOD

Study Design

We conducted a cross-sectional survey of hospital staff, using the HSOPS. The project was submitted for advisory purposes to an independent protection committee (IRB00006705) and received its approval from the French National Commission for Information Technology and Civil Liberties.

Setting

The study was conducted at a single 1836-bed university-affiliated hospital serving a predominantly urban population of 675,000 inhabitants in France. Our study site reported 135,999 stays in 2016. The hospital staff comprised 4422 registered healthcare providers and 642 board-certified physicians, the vast majority of them specialty hospitalists.

Participants

Eligible participants were full- or part-time (half-time or more) employees with at least 6 months of employment in the clinical, laboratory/pathology, radiology, or pharmacy departments. Staff who worked in more than one department or moved to another department during the study period were included only once.

Administrative or extended sick leave was not an exclusion criterion. Because of 6-month rotations, medical residents were not eligible for this survey. Hospital administrators, including chief executive officers, food service workers, and security, maintenance, or housekeeping service employees were not within the scope of this study. As recommended,[13] HSOPS questionnaires were excluded from the analytical sample if respondents answered less than one entire section, fewer than half of the items, or every item with the same nonneutral response.

Hospital Survey on Patient Safety Culture Questionnaire

Patient safety culture was assessed using the adaptation of the HSOPS in French.[14] The present analysis included the 42 items covering the 12 dimensions that were included in the original HSOPS questionnaire. Eighteen items were negatively worded.

Each item was answered on a five-point Likert scale, from "strongly disagree" (1 point) to "strongly agree" (5 points), or from "never" to "always" when relevant. A global safety grade between "poor" and "excellent" and the numbers of reported incidents in the past 12 months were also assessed but were not used to compute any of the 12 dimension scores.

Information on respondent background characteristics was collected at the end of the questionnaire, including age group, sex, position, seniority in the department, working time (expressed as a percentage of full-time), and involvement in an EFC. Consistent with Perneger et al.,[15] we defined the following seven categories of employees: (1) physicians; (2) head nurses, nurses, and midwives; (3) nursing assistants; (4) other healthcare providers (physical therapists, psychologists, and speech therapists); (5) administrative staff (reception staff and medical secretaries); (6) technical staff (pharmacists, biologists, and laboratory technicians); and (7) others (escort staff, social workers, and unit clerks).

Data Collection

The survey was conducted between April 2013 and September 2014. This period was necessary to achieve a high response rate and a large sample of respondents. During this period, no specific patient safety interventions were reported. As recommended,[13] a prenotification letter signed by the hospital chief executive officer was sent to hospital staff, presenting the objectives and data collection process of the upcoming survey. In each department, a research assistant promoted the survey and distributed the questionnaire to staff in the workplace. To maximize response rates, employees were allowed to complete the survey during work time. Participants returned the questionnaire in a sealed ballot box and signed an eligible employee register, both located in the nurse manager's office in each department. To ensure confidentiality, we did not use any identifier and questionnaires were anonymous. Reminders were distributed to nonrespondents (i.e., eligible employees who did not sign the participant register) until targeted response rates were achieved within each department.

Sample Size

No formal sample size calculation was performed, and the study sample size was determined by the number of eligible participants and survey response rates. To ensure representativeness and consistently with previous surveys using a French version of the HSOPS,[14,15] we aimed at achieving response rates of at least 70% among both physician and nonphysician staff members within each department. We anticipated that approximately 5000 employees would be eligible and 3500 would participate in the survey.

Dimension Scores

We computed the individual means across the three or four items in a dimension (range = 1 5).[16] Multiple imputation for missing item values (2% over total item responses) was performed using the predictive mean matching imputation method. As recommended by Rubin et al[17] and consistently with previous HOSPS-related studies[18], we set the number of imputations at $m = 5$ and we replaced the missing values with the mean of the five imputed data set.

Experience Feedback Committee Framework and Related Variables

The EFC complies with a written procedure in accordance with the method proposed by Air France Consulting. They are composed of volunteer representatives of the various professions within the medical teams. Committee meetings are conducted monthly within a standardized framework: (1) reading the list of reported events, (2) choosing a priority event to investigate by consensus according to the criticality of each incident, (3) choosing the investigator, (4) reviewing the RCA made up of the event chosen the previous month, (5) choosing corrective actions, and (6) monitoring on-going actions. The RCA is carried out during the month after the EFC by a designated person using a method, called Orion (online supplementary appendix 1, http://links.lww.com/JPS/A134), developed from systemic analysis methods used in civil aviation and adapted to the healthcare domain by Air France Consulting. Previously trained investigators must follow the main steps of the Orion method to fill out a standardized report: collecting data, describing the chronological facts that occurred before, during and after the event, describing the failures, looking for causes of errors and latent factors that could have contributed to the failures, implementing corrective actions, and writing a report of the analysis.

Three variables were used to characterize the individual link of each respondent with the EFC:

1. Involvement in EFC activities:We added a question at the end of the HSOPS questionnaire to collect the individual participation in one of the hospital's EFCs (yes or no). The missing responses for this question were assumed to be negative (3% of total responses).

© 2017 Wolters Kluwer Health, Inc. All rights reserved.

2. EFC in the department:We generated this variable with the practice program's administrative database comprising the information of all EFCs implemented in the hospital (department, leader, and annual reports).
3. EFC producing more than five corrective actions per year in the department:To add an EFC qualitative criterion, we computed another individual variable for all caregivers working in a department with an EFC. We generated this variable accounting for the department administrative code of each respondent and the number of corrective actions reported for each EFC (the median number of actions, five, was chosen to dichotomize this variable). To obtain this cutoff, we used the data from another study[16] in which all written documents produced by the EFC during a 1-year period were analyzed independently by two investigators (data not shown).

Statistical Analysis

Background respondent characteristics and EFC-related variables were reported as numbers and percentages, and HSOPS composite scores as means along with standard deviations. We compared HSOPS mean composite score values across subgroups of respondents defined by the three EFC variables. The differences between score means were tested by Student *t* tests. In addition, we quantified the mean differences by using Cohen *d* effect size. Finally, we performed a multivariate analysis of variance (MANOVA) to examine differences in dimension scores after adjusting for background characteristics.

Differences in categorical variables were analyzed using χ^2 tests. Two-sided *P* values lower than 0.05 were considered statistically significant. All analyses were performed using Stata Version 14.0 (Stata Corporation, College Station, TX).

RESULTS

Of 5064 eligible employees, 3978 (78.6%) participated in the study. After the exclusion of 90 questionnaires due to the discovery of exclusion criteria (online supplementary appendix 2, http://links.lww.com/JPS/A135), our sample consisted of 3888 survey questionnaires (76.8%). Table 1 summarizes the main characteristics of the respondents by EFC participation. Compared with nonmembers, the EFC participants had higher percentages of physicians (22% versus 10%) and technical staff members (17% versus 9%), and they worked more in the medicine and pediatrics departments (36% versus 29%) than the nonmember group ($P < 0.001$). Overall, 20 departments had an EFC, including seven medicine departments, six medical-technical departments, five emergency or intensive care departments, and two surgical departments.

Figure 1 shows the association between EFC participation and patient safety culture dimensions scores. Compared with nonparticipation, EFC participation was associated with a statistically significant higher patient safety culture for 9 of the 12 HSOPS dimensions (overall effect size = 0.31, 95% confidence interval = 0.21 to 0.41, $P < 0.001$). The MANOVA indicated that all the 12 dimension scores, taken together, differed according to involvement in EFC activities ($P < 0.0001$), independent of sex, department, and professional category. The coefficients of the multivariate regression model were consistent for eight of the nine dimension scores that were significant in the univariate analysis (Table 2). The three largest differences in HSOPS score were related to the "feedback and communication about error" (regression coefficient = 0.25, 95% confidence interval = 0.17 to 0.32, $P < 0.001$), "organizational learning" (regression coefficient = 0.22, 95% confidence interval = 0.16 to 0.29, $P < 0.001$), and "nonpunitive response to error" (regression coefficient = 0.18, 95% confidence interval = 0.11 to 0.26, $P < 0.001$) dimensions.

Figure 2 shows that the presence of an EFC in a department was not associated with patient safety culture on the part of the caregivers (overall effect size = −0.02, 95% confidence interval = −0.04 to 0.00, $P = 0.32$). Figure 3 shows that the professionals who worked in a department that had an EFC that implemented five or more actions per year in their department had a significantly

TABLE 1. Respondent Characteristics

Characteristics, n (%)	EFC Nonparticipants (*n* = 3375)	EFC Participants (*n* = 440)	*P*
Female sex	2678 (81.0)	338 (77.9)	0.13
Age, y			0.16
<35	1262 (38.0)	144 (33.0)	
35–44	850 (25.6)	116 (26.6)	
45–54	839 (25.3)	128 (29.4)	
≥55	367 (11.1)	48 (11.0)	
Occupational group			<0.001
Nurse	1225 (36.3)	161 (36.8)	
Nursing assistant	657 (19.5)	51 (11.6)	
Physician	339 (10.0)	97 (22.2)	
Other healthcare	117 (3.5)	7 (1.6)	
Administrative	314 (9.3)	17 (3.9)	
Technical	304 (9.0)	74 (16.9)	
Other	419 (12.4)	31 (7.1)	
Hospital sector			<0.001
Medicine and pediatrics	1000 (29.0)	159 (36.1)	
Surgery and gynecology	769 (22.3)	64 (14.6)	
Technical (pharmacy, operating rooms, imaging, laboratories)	685 (19.9)	110 (25.0)	
Intensive care, emergency, and anesthesia	592 (17.2)	96 (21.8)	
Several or other	402 (11.7)	11 (2.5)	

*Values were missing for sex (*n* = 147), age (*n* = 134), occupational group (*n* = 75), and contact with patient (*n* = 55).

© 2017 Wolters Kluwer Health, Inc. All rights reserved.

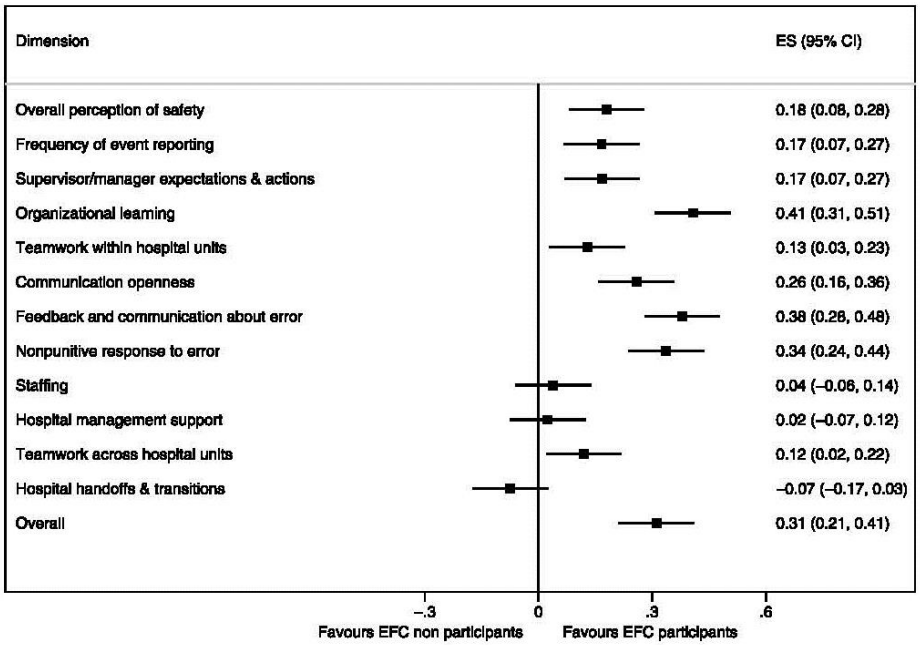


FIGURE 1. Comparison of HSOPS scores according to EFC participation (effect size).

higher patient safety culture than those who had an EFC in their department implementing fewer than five actions per year (overall effect size = 0.19, 95% confidence interval = 0.10 to 0.27, $P < 0.001$). This difference was confirmed by the MANOVA model.

Table 3 shows the comparison of the global safety grade and event reporting according the EFC-related variables. The percentages of professionals reporting at least one adverse event within the past year were significantly higher for EFC participants than nonparticipants (65% versus 51%, $P < 0.001$). Although all the other comparisons showed that the EFC participants had greater patient safety culture, the differences were not significant.

DISCUSSION

This study showed that being involved in EFC activities was associated with higher patient safety culture, in a large

TABLE 2. Multivariate Analysis of Variance Results

	EFC Participation		*P*
MANOVA*	Statistics	*F*	
Wilks' λ	0.9745	8.04	<0.001
Pillai's trace	0.0255	8.04	<0.001
Lawley-Hotelling trace	0.0262	8.04	<0.001
Roy's largest root	0.0262	8.04	<0.001
Multivariate regression underlying the MANOVA	Coefficient	95% CI	
Overall perception of safety	0.09	0.02 to 1.63	0.01
Frequency of event reporting	0.12	0.05 to 0.20	0.002
Supervisor/manager expectations and actions	0.13	0.05 to 0.21	0.002
Organizational learning	0.22	0.16 to 0.29	<0.001
Teamwork within hospital units	0.05	−0.02 to 0.13	0.17
Communication openness	0.14	0.07 to 0.21	<0.001
Feedback and communication about error	0.25	0.17 to 0.32	<0.001
Nonpunitive response to error	0.18	0.11 to 0.26	<0.001
Staffing	−0.01	−0.08 to 0.06	0.75
Hospital management support	0.02	−0.05 to 0.09	0.53
Teamwork across hospital units	0.06	0.01 to 0.12	0.03
Hospital handoffs and transitions	−0.03	−0.09 to 0.03	0.31

*The model included the 12 dimensions scores as dependant variables and the EFC participation, sex, hospital sector, and healthcare profession category as independant variables.

© 2017 Wolters Kluwer Health, Inc. All rights reserved.

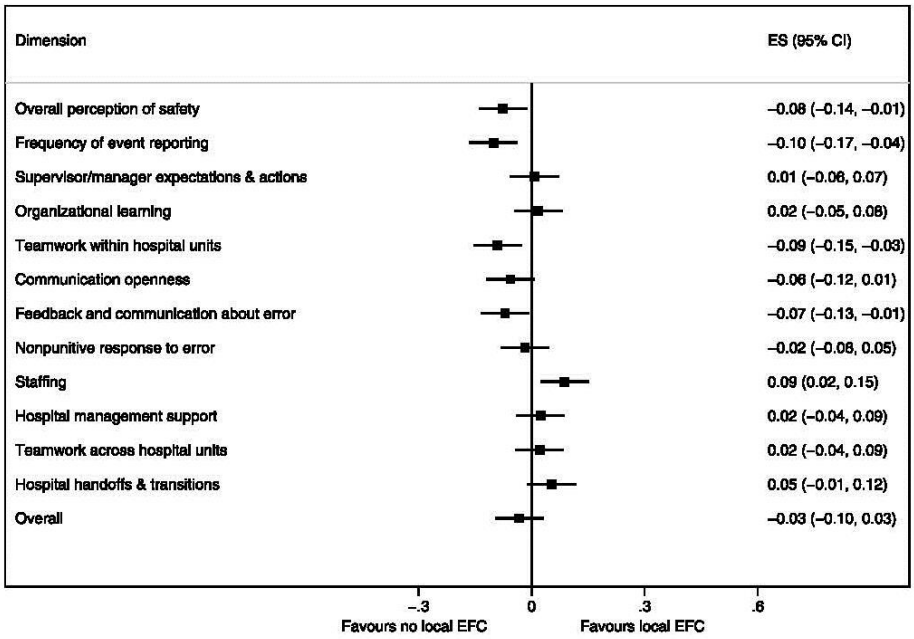


FIGURE 2. Comparison of HSOPS scores according to the presence of an EFC in the department (effect size).

university-affiliated hospital. At the dimensional level, EFC participation was significantly associated with nine of 12 dimensions of the HSOPS. The largest differences were found for three key dimensions of patient safety culture, i.e., "organizational learning," "feedback and communication about error," and "nonpunitive response to error." These associations remained significant after adjusting for sex, occupational group, and hospital sector. Moreover, the professionals involved in an EFC declared reporting adverse events more often than those who did not participate in an EFC. However, we did not find any association between the presence of an EFC in the department and professionals' patient safety culture. The analysis of the subgroup of professionals who worked in a department with a productive EFC, defined as an EFC implementing at least five actions per year, showed a higher patient safety culture level for the dimensions.

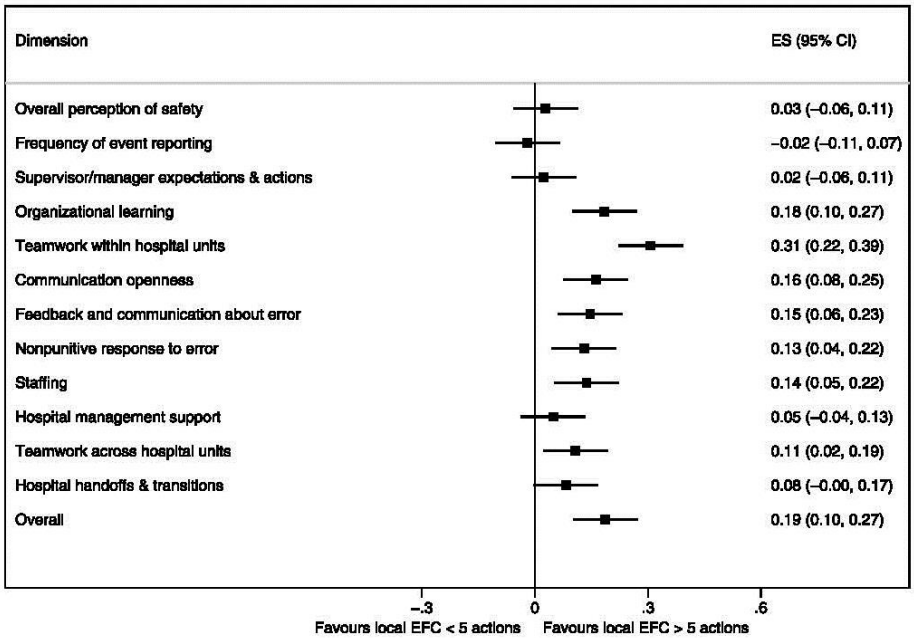


FIGURE 3. Comparison of HSOPS scores according to the presence of an EFC implementing at least five corrective actions per year in the department (effect size).

© 2017 Wolters Kluwer Health, Inc. All rights reserved.

TABLE 3. Comparison of Safety Grade and Event Reporting According the EFC-Related Variables

	Safety Grade			Reported an Event Within the Past 12 Months		
	Excellent or Very Good	Good, Fair, or Poor	*P*	≥1	None	*P*
EFC direct participation, n (%)			0.18			<0.001
Yes	170 (38.6)	270 (61.4)		286 (65.0)	154 (35.0)	
No	1220 (35.4)	2228 (64.2)		1755 (50.9)	1693 (49.1)	
EFC in the department, n (%)			0.79			0.08
Yes	817 (35.9)	1457 (64.1)		1221 (53.7)	1053 (46.3)	
No	573 (35.5)	1041 (64.5)		820 (50.8)	794 (49.2)	
EFC in the department, n (%)			0.18			0.13
Implementing at least 5 actions/y	535 (37.0)	913 (63.0)		795 (54.9)	653 (45.1)	
Implementing <5 actions/y	282 (34.1)	544 (65.7)		426 (51.6)	400 (48.4)	

Although widely implemented in French hospitals for more than 10 years, EFCs had practically never been assessed. Despite the observational design of this study, the results seem relevant, because they endorse the objectives, which led to the initial development of this management tool. Indeed, the EFC was initially developed to directly involve medical teams in patient safety management and to educate the professionals in the science of safety. The results reported herein are in line with several studies reporting a favorable association of similar tools on the patient safety climate, such as the Comprehensive Unit-based Safety Program in the United States.[19]

Interestingly, participation in EFC activities was strongly associated with the "organizational learning," "feedback and communication about error," and "nonpunitive response to error" dimension scores. The "organizational learning" dimension was compounded by three items related to team organization to improve patient safety (namely, "A6 active improvement of patient safety," "A9 mistakes led to positive changes," and "A13 evaluation of the implemented actions' effectiveness"). The association between this latter dimension and participation in EFC activities may reflect the appropriateness of the EFC framework for patient safety management, whether to implement corrective actions or to assess their effectiveness. Indeed, the standardized EFC framework includes two steps related to implementing corrective actions after multidisciplinary consensus and the monitoring of ongoing actions. These actions are prompted after reviewing the RCA presented by one of the participants. Therefore, the decision-making process is supplied by the adverse events encountered by the team after the identification of latent organizational failures. Experience feedback committee participants are likely more conscious of the importance of feedback and the nonpunitive response to error. Learning from errors is a fundamental aspect of safety science.[20] The routine practice of RCA specifically adapted to healthcare professionals may positively impact the participants' perception, resulting in their awareness that adverse events usually result from multiple organizational failures.

A parallel could be drawn between this knowledge about errors and adverse event reporting. The "frequency of event reporting" dimension was more developed for EFC participants who declared reporting significantly more adverse events than nonparticipants. The EFC framework establishes event reports as the raw material of the meeting. Indeed, the EFC requires reports of adverse events at each step of the meeting, which start by a reading of all the events to choose one of them to conduct a thorough analysis, and finish by elaborate corrective actions. The "nonpunitive response to error" dimension, more fully developed for the EFC participants, could also limit the fear of punishment, well-known as an obstacle to event reporting.[21,22] The proportion of professionals who reported at least one event per year was higher for departments with an EFC than for departments without an EFC (54% versus 51%, $P = 0.08$). Although not significant, this difference could stem from EFC participants encouraging event reporting, to supply the RCA in EFC meetings. Overall, these results showed that the EFC could promote event reporting, despite all the barriers hindering self-reporting, such as the time constraints, complex forms, and the lack of feedback.[20,23]

Three HSOPS dimensions were not associated with EFC participation, namely "staffing," "hospital management support," and "Hospital handoffs and transition," reflecting structural issues related to the context of French hospitals, but also specific RCA problems in healthcare. The "staffing" dimension relates to staff management, scheduling working hours, turnover, and replacements. The negative effects of staffing shortages are widely known as decreasing patient safety and quality of care.[24] The "hospital management support" dimension reflects the management's patient safety policy. These dimension scores must be analyzed within the European context of austerity policies and budget cuts after the global financial crisis.[25,26] It was also not surprising that EFCs, despite potential identification of staff management or lack of resources as root causes, fail to reverse this trend in hospital policy. However, these results confirmed the limited efficacy of RCA conducted at the department level, which cannot instigate deep organizational change at the hospital level.[5,8,9]

The absence of a difference for the "hospital handoffs and transitions" dimension also reflects some of the RCA limitations previously identified. The EFC targeted patient safety within medical teams, but it remained limited at the department level. This limitation was in line with the problem with RCA in healthcare.[27,28] Notably, Peerally et al.[8] highlighted that the multiplication of localized action plans that are not shared by all the departments resulted in failure to address patient safety concerns more fully. Therefore, a major limitation of the EFCs was their inability to act beyond the localized team organization. This limitation could also explain similar patient safety culture for professionals whether or not they worked in a department with an EFC.

Compared with similar patient safety management tools, EFCs involve the team in routine RCA.[29,30] Our results also support the main limit related to RCA in healthcare. Indeed, earlier studies reported that the RCA conducted by the EFC participants often lacked depth and rigor or that the RCA method was rarely well applied.[10,11] In these studies, some incidents analyzed that were supposed to be corrected by an action plan were recurrent, showing that the EFC was ineffective in really improving patient safety. Overall, our study cannot deny the limited efficacy of RCA,

© 2017 Wolters Kluwer Health, Inc. All rights reserved.

whether to properly identify the actual failures of the organization or to propose an appropriate action plan.

One of our hypotheses was that patient safety culture would be disseminated toward professionals working in a department with an EFC. The lack of this association at this level highlighted the weakness mentioned previously. Indeed, the association was limited to the direct participant, showing the narrow scope of the EFC. This may be related to the failure of this tool to implement in-depth corrective actions beyond the perimeter of the EFC. Beyond these issues specifically related to the RCA practice, another explanation for this limited association was a lack of communication. Through semidirected interviews of healthcare professionals, we identified the lack of communication of EFC activities as a weakness (data not shown). To extend the perimeter of the EFC, regular communication actions on their activities are necessary. Indeed, sharing the adverse events analysis and the related action plans to the nonparticipants could help them to face similar organizational issues and to encourage the involvement of new participants. Including communication plans in the EFC framework may improve its impact beyond the active members.

The limitations of this study deserve mention. First, this study had an observational design and was cross-sectional, which cannot exclude a reverse causality bias. Experience feedback committee participation is voluntary, and participants may be more aware of the patient safety concept than the nonparticipants, before their involvement in the EFC. However, we adjusted our analyses on baseline characteristics, showing the robustness of the association between EFC participation and patient safety culture. We believe that only a direct comparison through a randomized controlled trial with a longitudinal design can definitively settle this issue. Second, the present study was conducted at a single university hospital and the findings may not apply to other facilities. Third, we did not collect information on nonrespondent characteristics, and therefore, we could not examine the potential for selection bias.

CONCLUSIONS

Being involved in EFC activities was associated with higher patient safety culture scores. The strongest association was found for the following three key HSOPS dimensions: "organizational learning," "nonpunitive response to error," and "feedback and communication about error." However, the EFC had no impact on the patient safety culture of the nonparticipants, probably indicating a weakness related to the low dissemination and the lack of communication on the EFC activities. Moreover, another weakness of this tool was related to inherent issues of the RCA practice, indicating the difficulty effectively correcting the adverse events encountered by the hospital wards. Despite these limitations, the EFC remains a real opportunity to involve medical teams in patient safety management and seems to impact the culture of the caregivers positively. The findings, if causal, suggest that routine RCA improves the patient safety culture and is adapted to the idiosyncrasies of the healthcare setting.

ACKNOWLEDGMENTS

The authors thank Linda Northrup for her assistance in editing the article in English.

REFERENCES

1. Makary MA, Daniel M. Medical error-the third leading cause of death in the US. *BMJ*. 2016;353:i2139.
2. Bates DW, Larizgoitia I, Prasopa-Plaizier N, et al. Global priorities for patient safety research. *BMJ*. 2009;338:b1775.
3. Reason J. Human error: models and management. *BMJ*. 2000;320: 768 770.
4. Leape LL, Berwick DM, Bates DW. What practices will most improve safety? Evidence-based medicine meets patient safety. *JAMA*. 2002;288: 501 507.
5. Wu AW, Lipshutz AK, Pronovost PJ. Effectiveness and efficiency of root cause analysis in medicine. *JAMA*. 2008;299:685 687.
6. Commission TJ. Patient Safety Systems Chapter for the Hospital program. Secondary Patient Safety Systems Chapter for the Hospital program 2017. Available at: https://www.jointcommission.org/assets/1/18/CAMH_04a_PS.pdf. Accessed June 16, 2017.
7. Occelli P. La culture de sécurité des soins: du concept à la pratique. Secondary La culture de sécurité des soins: du concept à la pratique 2010. Available at: https://www.has-sante.fr/portail/upload/docs/application/pdf/2011-02/culture_de_securite_des_soins__du_concept_a_la_pratique.pdf. Accessed May 8, 2017.
8. Peerally MF, Carr S, Waring J, et al. The problem with root cause analysis. *BMJ Qual Saf*. 2017;26:417 422.
9. Kellogg KM, Hettinger Z, Shah M, et al. Our current approach to root cause analysis: is it contributing to our failure to improve patient safety?. *BMJ Qual Saf*. 2017;26:381 387.
10. Boussat B, Bougerol T, Detante O, et al. Experience Feedback Committee: a management tool to improve patient safety in mental health. *Ann Gen Psychiatry*. 2015;14:23.
11. Lecoanet A, Sellier E, Carpentier F, et al. Experience feedback committee in emergency medicine: a tool for security management. *Emerg Med J*. 2014;31:894 898.
12. Nieva VF, Sorra J. Safety culture assessment: a tool for improving patient safety in healthcare organizations. *Qual Saf Health Care*. 2003; 12(Suppl 2):ii17 ii23.
13. Sorra J, Nieva VF. *Hospital Survey on Patient Safety Culture. (Prepared by Westat, under Contract No. 290-96-0004). AHRQ Publication No. 04 0041*. Agency for Healthcare Research and Quality: Rockville, MD; 2004. Available at: http://www.ahrq.gov/sites/default/files/publications/files/hospcult.pdf. November 14, 2014.
14. Occelli P, Quenon JL, Kret M, et al. Validation of the French version of the Hospital Survey on Patient Safety Culture questionnaire. *Int J Qual Health Care*. 2013;25:459 468.
15. Perneger TV, Staines A, Kundig F. Internal consistency, factor structure and construct validity of the French version of the Hospital Survey on Patient Safety Culture. *BMJ Qual Saf*. 2014;23:389 397.
16. Giai J, Boussat B, Occelli P, et al. Hospital survey on patient safety culture (HSOPS): variability of scoring strategies. *Int J Qual Health Care*. 2017;29:685 692.
17. Vlayen A, Schrooten W, Wami W, et al. Variability of patient safety culture in Belgian acute hospitals. *J Patient Saf*. 2015;11:110 121.
18. Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. *Stat Med*. 1991;10:585 598.
19. Weaver SJ, Lubomksi LH, Wilson RF, et al. Promoting a culture of safety as a patient safety strategy: a systematic review. *Ann Intern Med*. 2013;158(5 Pt 2):369 374.
20. Evans SM, Berry JG, Smith BJ, et al. Attitudes and barriers to incident reporting: a collaborative hospital study. *Qual Saf Health Care*. 2006; 15:39 43.
21. Hobgood C, Weiner B, Tamayo-Sarver JH. Medical error identification, disclosure, and reporting: do emergency medicine provider groups differ? *Acad Emerg Med*. 2006;13:443 451.
22. Leape LL. Reporting of adverse events. *N Engl J Med*. 2002;347: 1633 1638.
23. Blendon RJ, DesRoches CM, Brodie M, et al. Views of practicing physicians and the public on medical errors. *N Engl J Med*. 2002;347: 1933 1940.

© 2017 Wolters Kluwer Health, Inc. All rights reserved.

24. Aiken LH, Clarke SP, Sloane DM, et al. Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. *JAMA*. 2002;288:1987 1993.

25. Wenzl M, Naci H, Mossialos E. Health policy in times of austerity-A conceptual framework for evaluating effects of policy on efficiency and equity illustrated with examples from Europe since 2008. *Health Policy*. 2017;121:947 954.

26. Morgan D, Astolfi R. Financial impact of the GFC: health care spending across the OECD. *Health Econ Policy Law*. 2015;10:7 19.

27. Bowie P, Skinner J, de Wet C. Training health care professionals in root cause analysis: a cross-sectional study of post-training experiences, benefits and attitudes. *BMC Health Serv Res*. 2013;13:50 doi: 10.1186/1472-6963-13-50.

28. Percarpio KB, Watts BV. A cross-sectional study on the relationship between utilization of root cause analysis and patient safety at 139 Department of Veterans Affairs medical centers. *Jt Comm J Qual Patient Saf*. 2013;39:32 37.

29. Weld LR, Stringer MT, Ebertowski JS, et al. TeamSTEPPS improves operating room efficiency and patient safety. *Am J Med Qual*. 2016;31:408 414.

30. Pronovost P, Needham D, Berenholtz S, et al. An intervention to decrease catheter-related bloodstream infections in the ICU. *N Engl J Med*. 2006; 355:2725 2732.

© 2017 Wolters Kluwer Health, Inc. All rights reserved.

5. Chapitre IV: Synthèse et perspectives

L'ensemble de ces travaux nous ont permis d'identifier les forces et les faiblesses du comité de retour d'expérience et d'appréhender sa contribution à la sécurité des patients au travers de la culture des professionnels de santé. Ces travaux ont également permis de dégager plusieurs pistes d'amélioration pour augmenter l'adhésion des professionnels de santé et l'efficacité globale de ce dispositif de gestion de la sécurité en équipe.

5.1. CREX et culture de l'erreur

La principale force des CREX est son impact positif sur la culture de l'erreur. Les 3 dimensions les plus fortement associées à la participation aux CREX concernaient l'« Organisation apprenante et amélioration continue », la « Réponse non punitive à l'erreur » et le « Retour d'information et communication des erreurs ».

Ces résultats indiquent que les professionnels de santé impliqués dans les CREX dépassent la perception habituelle de l'erreur en tant que faute individuelle et supplantent la culture de l'amélioration à la culture de la sanction (Reason 2000, Leape and Berwick 2005). Le CREX vise à corriger les défaillances de l'organisation à partir de l'analyse des causes racines d'événements indésirables auxquels sont confrontés les membres de l'équipe. Malgré les limites que nous avons constatées dans la pratique de l'analyse des causes racines, il parait probable que ce fonctionnement modifie positivement la culture de l'erreur des professionnels de santé engagés. Cette hypothèse était également confirmée par le pourcentage de professionnels de santé signalant des événements indésirables, qui était significativement plus élevé chez les participants aux CREX. Les participants, plus conscients de l'importance du retour d'expérience à partir des erreurs, sont donc naturellement plus enclins à déclarer les événements indésirables. Un autre marqueur de cette meilleure appréhension de l'erreur était la dimension attenante à la liberté d'expression, qui était également plus élevée chez les participants aux CREX. Les 3 items composant cette dimension concernaient la faculté à

remettre en cause les décisions des supérieurs, à s'exprimer librement, et sans peur, lorsqu'un dysfonctionnement pouvant avoir des conséquences sur la sécurité d'un patient était repéré. Les participants aux CREX déclaraient aussi plus souvent qu'ils n'avaient pas « l'impression que les erreurs leurs étaient reprochées », que « lorsqu'un événement était signalé c'est le problème et non la personne qui était pointée du doigt », et étaient moins « inquiets du fait que les erreurs soient notées dans les dossiers administratifs ». Cette association positive entre participation aux CREX et réponse non punitive est à souligner, car elle est un prérequis indispensable pour améliorer la transparence autour des erreurs, et pour diminuer la notion de culpabilité, qui entraine des conséquences néfastes aussi bien pour les patients que pour les soignants, avec l'exemple du concept de seconde victime (Wu 2000).

Cependant, nous avons également constaté un défaut de diffusion de la culture de sécurité au sein des services comportant un CREX. L'une de nos hypothèses était que s'opérait une diffusion de la culture de sécurité à l'ensemble des professionnels travaillant dans les services disposant d'un CREX, indépendamment de leur implication directe. Une explication possible de ce défaut de diffusion est un manque de communication au sein des équipes. En effet, seuls quelques CREX avaient systématisé la communication de leurs travaux aux autres membres du service, par l'intermédiaire par exemple d'affiches dans les offices de soins, ou de comptes rendus diffusés par voie électronique. Comme piste d'amélioration, on peut ainsi retenir l'organisation d'une communication régulière vers l'ensemble des professionnels des services. Un exemple intéressant concernait par un CREX qui avait mis en place une affiche intitulée « le CREX du mois », indiquant l'événement indésirable analysé en profondeur et les actions correctrices décidées. Ce type de communication pourrait être l'occasion de discussion et d'échanges au sein des équipes, et pourrait également valoriser l'engagement des professionnels participants aux CREX.

5.2. CREX et travail en équipe

Une autre dimension culturelle essentielle pour améliorer la qualité et la sécurité des soins est la communication et la synergie dans les équipes soignantes (Duclos, Peix et al. 2016). Il est en effet reconnu que le travail en équipe et la collaboration inter professionnelle ont un impact significatif sur la survenue d'événements indésirables : les approches multi disciplinaires permettent d'identifier plus facilement les défaillances du système et de mieux comprendre les causes racines des événements indésirables (Neily, Mills et al. 2010, Deering, Rosen et al. 2011, Jones, Skinner et al. 2013, Schmutz and Manser 2013). La multidisciplinarité est justement au cœur de la méthode CREX, le comité devant être composé d'un groupe de professionnels représentant les différents métiers du service. Cette volonté de promouvoir la multidisciplinarité distingue les CREX des Revues de mortalité et de morbidité (RMM), autre dispositif impliquant lui aussi les professionnels de santé dans la gestion des événements indésirables (Sellier, David-Tchouda et al. 2012). Si les RMM sont également reconnues pour favoriser le travail en équipe et la collaboration, elles sont historiquement centrées sur les médecins, bien que certaines études montrent qu'elles tendent vers plus de multidisciplinarité (Francois, Prate et al. 2016). Nos travaux confirment cette multidisciplinarité propre aux CREX, la diversité des professions étant effectivement représentée dans les CREX analysés. Dans l'étude du fonctionnement des 20 CREX, les listes d'émargement rapportaient 36% de personnels paramédicaux et 28% de personnels médicaux. Cet équilibre entre médecins et paramédicaux dans les CREX était également confirmé lors de l'enquête de culture sécurité (37% d'infirmières et 22% de médecins). L'enquête culture sécurité indiquait une association favorable entre la participation à un CREX et la dimension relative au travail d'équipe. Cette dimension regroupait 4 items questionnant sur le soutien mutuel et le respect des membres de l'équipe. Il était particulièrement intéressant de remarquer que les CREX considérés comme productifs (produisant au moins 5 actions correctrices par an) étaient associés à un meilleur score pour

cette dimension, comparativement aux CREX considérés comme non productifs (taille d'effet = 0.31). Même si la possibilité d'un biais de causalité inverse ne peut être écartée, ce résultat semble montrer que les CREX favorisent fortement le travail d'équipe, à condition de respecter la méthode telle qu'elle a été initialement développée. Cela est confirmé par le constat effectué lors de l'étude du fonctionnement des CREX. En effet, nous avons retrouvé un CREX ne respectant pas le quorum et la multidisciplinarité. Ce comité n'avait réuni au cours de l'année de fonctionnement analysé, qu'un seul médecin accompagné de deux cadres de santé. Même si ce comité avait examiné un nombre conséquent d'événements indésirables, aucun rapport d'analyse des causes racines n'avait été produit et aucune action correctrice n'avait été décidée. Le respect du caractère multidisciplinaire lors de la mise en place d'un CREX parait donc une condition nécessaire pour disposer d'un outil opérationnel, produisant des actions concrètes d'amélioration de la sécurité des patients, et favorisant l'entraide et le soutien des membres de l'équipe médicale.

5.3. Limites de l'analyse des causes racines

Si les CREX semblent favoriser plusieurs aspects essentiels de la culture de sécurité des professionnels de santé, les travaux réalisés ont également permis de dégager plusieurs faiblesses attenantes à cet outil. Premièrement, nous avons remarqué au travers de l'étude du fonctionnement des CREX certaines difficultés à suivre la méthode préconisée par Air France consulting. En effet, plusieurs CREX déviaient de la méthode originale, en ne systématisant pas la pratique de l'analyse des causes racines à chaque réunion. De plus, certaines analyses des causes racines effectuées ne suivaient pas la méthode Orion, aboutissant à des rapports souvent superficiels. On notait aussi un manque de planification et de suivi des actions correctrices. L'analyse du CREX de Neuropsychiatrie, montrait par exemple que si les rapports Orion étaient régulièrement présentés au cours de la première année de fonctionnement, cette pratique s'étiolait au cours des deux années suivantes, et que le suivi des actions correctrices diminuait.

Alors que certains événements, comme la présence de patients hospitalisés dans les couloirs de neurologie avaient nettement diminué suite à des analyses et actions mises en place au cours de la première année, on a vu réapparaitre des événements similaires au cours de la 3e année étudiée. Ces problèmes sont révélateurs d'un manque de suivi des actions, mais questionnent également sur l'efficacité de la méthode Orion, et plus généralement sur l'efficacité des méthodes d'analyses systémiques. Si l'analyse des causes racine s'est imposée comme la méthode de référence pour traiter les événements indésirables en santé, la littérature signale d'importantes limites dans sa mise en œuvre pratique (Percarpio, Watts et al. 2008, Wu, Lipshutz et al. 2008, Wallace, Spurgeon et al. 2009, Peerally, Carr et al. 2017). En effet, plusieurs études rapportent que les analyses manquent souvent de profondeur et de rigueur, et que les plans d'actions en découlant sont souvent trop simples et inadéquats pour prévenir efficacement les événements indésirables. Bien que proche de la méthode ALARM, la méthode Orion a pourtant été conçue pour être plus simple d'utilisation pour les soignants et moins chronophage (Vincent, Taylor-Adams et al. 2000, Debouck, Rieger et al. 2012). Malgré cela, les CREX qui réalisaient moins souvent d'analyses des causes imputaient cet écart à une activité considérée comme trop complexe et trop chronophage. En addition de ce manque de temps et de ressources, les difficultés pour réaliser des analyses des causes racines s'expliquent également par des interactions complexes avec le contexte socio-culturel. Afin de préserver de bonnes relations inter professionnelles et éviter les tensions avec la hiérarchie, il arrive fréquemment que les soignants en charge de l'analyse restent à la surface de l'événement et cachent certains problèmes organisationnels plus profonds (Braithwaite, Westbrook et al. 2006, Bowie, Skinner et al. 2013). Fort de ce constat, un groupe de travail se met actuellement en place au niveau national, avec pour objectif d'aboutir à une version simplifiée de la méthode Orion, afin d'améliorer son efficacité et sa praticité pour les professionnels de terrain, non spécialistes de la gestion des risques. Toutefois, l'adaptation de la méthode devrait

probablement être accompagnée de changements managériaux, avec une augmentation de moyens dédiés au CREX, afin de laisser du temps aux soignants pour réaliser les analyses sur leur temps de travail. Cela permettrait d'améliorer la qualité des analyses des causes racines, et également de mieux valoriser la participation à ces activités bénéfiques pour l'organisation et la sécurité des patients.

5.4. Périmètre d'action des CREX

Les difficultés des CREX à traiter en profondeur les événements indésirables sont également dues au périmètre dans lequel s'inscrivent les comités. Les CREX impliquent les professionnels de terrain dans l'analyse des événements indésirables en équipe, se situant donc au niveau de la première boucle de la sécurité des soins (Michel 2011). Si l'analyse des causes permet d'identifier et de traiter des défaillances liées à l'organisation de l'équipe, les événements indésirables dépassent régulièrement le périmètre strict de l'équipe. C'est notamment le cas des événements dans lesquels sont impliqués plusieurs équipes, lors de transfert de patients imposés, de dysfonctionnements du plateau technique ou des services logistiques, mettant en jeu des défauts de transmission et de coordination entre plusieurs services. L'étude des relations entre CREX et culture sécurité semble corroborer cette limite. En effet, la dimension « transfert de patients et transmission médicale » ne différait pas selon la participation aux CREX. Cette faiblesse devrait entrainer une réflexion sur le périmètre d'implantation des CREX. En effet, si la méthode a été initialement créée pour être implantée au niveau d'un service, deux CREX de notre établissement ont été mis en place avec succès sur plusieurs services. Ils s'agissaient des comités du pôle Hôpital Couple Enfant, regroupant les services de pédiatrie polyvalente, réanimation, chirurgie pédiatrique, gynécologie et obstétrique ou encore celui du pôle Neuropsychiatrie, regroupant les services de neurologie et psychiatrie. Cet élargissement de périmètre pourrait permettre de mieux agir sur les interfaces et la collaboration inter équipes, mais s'accompagne également de difficultés non négligeables, comme la répartition du

leadership entre les services et la quantité plus importante d'événements à examiner. Aussi se pose la question des moyens à mettre en œuvre pour garantir la pérennité de ces CREX réunissant plusieurs équipes.

5.5. CREX et management : un manque de soutien institutionnel

L'amélioration de l'efficacité des CREX passe également par une réflexion sur la place et le rôle des CREX dans l'institution. Nous avons précédemment indiqué que les CREX peinaient à dépasser le périmètre d'action du service. L'une des pistes d'amélioration à considérer est une meilleure communication entre les CREX et les managers. En effet, les managers, que l'on peut situer au second niveau de la triple boucle de la sécurité des soins, peuvent mettre en place les actions d'amélioration à l'échelle de l'établissement (Michel 2011). Ce défaut de coordination entre les CREX et le management hospitalier était visible lors de l'enquête de culture sécurité. En effet, le score de la dimension « soutien du management à la sécurité des soins » comportait le score le plus faible sur les 12 dimensions mesurées, et aucune différence n'était retrouvée selon la participation aux CREX. Cette dimension explore la perception des professionnels de santé au sujet du climat de travail instauré par la direction, des actions menées par celle-ci, et plus généralement de son intérêt pour la sécurité des patients. Ces scores faibles semblent indiquer que la sécurité des patients n'est pas perçue comme une priorité de la gouvernance dont le discours reste centré sur la nécessité de faire des économies, dans un contexte de difficultés budgétaires chroniques. Mais ces résultats indiquent également un défaut de communication entre les managers et les professionnels de terrain. Une meilleure prise en compte des activités des CREX pourrait fournir des pistes d'amélioration pour la sécurité des patients mais aussi pour le management, et favoriser les liens entre les professionnels de santé et la direction. Une meilleure coordination pourrait également permettre de contrer une autre limite de l'analyse des causes racines décrite dans la littérature. Peerally décrit ainsi que la multiplicité des plans d'actions localisés, non partagés par les autres services de l'hôpital,

empêche de lutter efficacement contre les événements indésirables à l'échelle de l'hôpital (Peerally, Carr et al. 2017).

5.6. Limites de l'étude

Ce travail comporte plusieurs limites qui méritent d'être mentionnées. Premièrement, l'étude était monocentrique. Or, il est établi que les résultats des travaux de recherche en qualité et sécurité des soins sont fortement influencés par la structure et le contexte (Pronovost and Marsteller 2011). Même si notre établissement est comparable sur bien des aspects aux autres centres hospitaliers universitaires français, il est probable que les CREX et la perception des professionnels de santé soient influencés par le contexte, les organisations et modes de management propres à notre établissement. Les problèmes du soutien du management et du manque de communication entre les CREX et les directions, pourraient par exemple être différents dans d'autres établissements et de ce fait modifier le fonctionnement des CREX et la perception qu'en ont les professionnels de santé impliqués. Des études similaires, conduites en multicentrique, pourraient permettre d'explorer cette question.

Deuxièmement, le design transversal de l'étude des relations entre participation aux CREX et culture de sécurité ne permet pas d'établir avec certitude des relations de causalité, et d'exclure la possibilité d'un biais protopathique. Une des hypothèses expliquant l'association favorable entre participation à un CREX et culture de sécurité pourrait être que les professionnels engagés dans les activités des comités avaient préalablement une conscience plus élevée de la question de la sécurité des soins. Même si cette hypothèse ne peut être exclue, une revue de la littérature rapportait des effets bénéfiques de programmes similaires aux CREX sur la culture de sécurité (Weaver, Lubomksi et al. 2013). De plus, nous avons réalisé des modèles multivariés ajustés sur les caractéristiques des professionnels, qui confirmaient les associations.

Une autre limite attenante au design transversal de l'étude était qu'il ne permettait pas d'évaluer la dynamique de la culture de sécurité au cours du temps, l'évaluation constituant une photographie de cette culture à un moment donné. Une étude de type quasi-expérimentale de type avant-après pourrait permettre d'augmenter le niveau de preuve et valider les hypothèses de causalité. Cependant, le développement des CREX a été très important ces dernières années, et ceux-ci sont déjà implantés dans la majorité des établissements de santé français.

5.7. Perspectives

Les travaux réalisés ouvrent des perspectives d'amélioration du dispositif CREX, ainsi que des pistes de recherche sur les services de santé.

5.7.1. Perspectives d'amélioration des CREX

Modification de la méthode d'analyse des causes racines. Il est nécessaire de développer une méthode d'analyse des causes racines mieux adaptée à la pratique des professionnels de santé que la méthode Orion, plus facile et rapide d'utilisation. Un atelier est prévu à cet objet, lors des prochaines journées de l'Institut National pour la Qualité et la Sécurité des soins

Mise en place de dispositifs pour favoriser la communication du CREX vers les autres professionnels du service. Les activités du CREX doivent être plus transparentes afin de garantir l'adhésion de l'équipe et favoriser le signalement d'événements indésirables, essentiels pour le fonctionnement du CREX.

Intégration forte des CREX dans le dispositif de gestion des risques des établissements. Il est nécessaire d'améliorer la communication et le partage d'information entre les CREX et la cellule de gestion des risques de l'établissement, avec des apports complémentaires, en particulier pour l'analyse d'événements indésirables impliquant plusieurs services et équipes.

Amélioration de la reconnaissance institutionnelle des CREX. Les gouvernants de l'établissement doivent être régulièrement informés des activités des CREX. Ils doivent être

prêts à soutenir les actions d'amélioration de la sécurité découlant de l'analyse approfondie des événements indésirables. Il est nécessaire d'affecter les ressources humaines suffisantes pour que les professionnels de santé aient du temps à consacrer aux activités de gestion des risques.

5.7.2. Perspectives de recherche sur les dispositifs de gestion de la sécurité des soins dans les équipes médicales

Cette thématique qui fait partie du domaine de la recherche sur les services de santé, comprend deux axes principaux :

Développement d'indicateurs de résultats pour la sécurité des patients. Outre les limites attenantes au design de notre étude, nos résultats concernent principalement des associations aux dimensions de la culture de sécurité, et donc aux perceptions des professionnels de santé. Si la littérature établit des liens directs entre culture de sécurité et résultats cliniques (Braithwaite, Herkes et al. 2017, Mannion and Smith 2017), il apparait indispensable de vérifier l'efficacité des dispositifs de gestion de la sécurité en équipe au travers de l'analyse des résultats cliniques sur les patients. Le développement d'indicateurs cliniques fiables, permettrait de satisfaire cet objectif et de permettre une évaluation régulière et dynamique de la sécurité des patients hospitalisés (Quan, Drosler et al. 2008, Januel, Chen et al. 2012, Southern, Burnand et al. 2017). L'amélioration progressive des systèmes d'informations hospitaliers, de la capacité d'analyse des données de grandes dimensions ou d'analyse textuelle libre laissent entrevoir des possibilités pour aider la recherche épidémiologique sur les événements indésirables. De plus, des travaux visant à inclure des codes relatifs aux événements indésirables dans la nouvelle version de la Classification Internationale des Maladie (CIM-11) sont actuellement menés par un groupe de chercheurs piloté par l'OMS (Forster, Bernard et al. 2017). Compte-tenu de l'ouverture du Système National de Données de Santé (SNDS) et de la prochaine mise en place de la CIM-11, l'analyse à grande échelle d'indicateurs de résultats en sécurité des soins pourrait progresser considérablement dans les prochaines années.

Elaboration et mise en œuvre de nouveaux dispositifs de gestion des risques. En restant dans le postulat que les équipes médicales doivent être impliquées directement dans la gestion des risques, il est nécessaire d'imaginer et d'explorer de nouveaux outils. Nous avons vu que le dispositif CREX répond au postulat, et nous avons montré certains avantages mais aussi certaines faiblesses. Il faut sans doute améliorer le dispositif CREX mais il pourrait également être intéressant d'en inventer des nouveaux, ou plutôt d'adapter des dispositifs développés dans d'autres pays. En effet, les CREX ne sont implantés qu'en France, alors que des dispositifs de gestion des risques en équipes ont été implantés avec succès à l'international. C'est notamment le cas pour les CUSP, qui bénéficient de nombreux travaux d'évaluation, et dont l'impact peut être mieux appréhendé (Smith and Flanders 2014, Champion, Sadek et al. 2017, Khan, Aljuaid et al. 2017, Pitts, Maruthur et al. 2017).

5.8. Conclusion

Le CREX est un dispositif inventé en France en 2005 et déployé dans les services de santé français depuis 2007. Son objectif est d'impliquer des professionnels de santé « de terrain », des équipes et services hospitaliers dans la gestion active de la sécurité des patients. Curieusement, après 10 ans de développement dans des milliers d'équipes, aucun travail scientifique publié ne s'était penché sur leur fonctionnement et leurs effets sur la sécurité des patients. Malgré ses limites, notre travail contribue à mieux connaitre les CREX et leur place dans les dispositifs de gestion de la sécurité du patient. Ce travail plaide pour le développement d'une recherche centrée sur l'implication des professionnels de santé dans la gestion de la sécurité des patients.

6. REFERENCES BIBLIOGRAPHIQUES

Facing up to medical error. BMJ 2000, 320:7237-A.

AHRQ Rockville, M.Hospital Survey On Patient Safety. http://www.ahrq.gov/qual/patientsafetyculture/hospscanform.pdf. Accessed 2004

AHRQ Rockville, M.Surveys on Patient Safety Culture Research Reference List http://www.ahrq.gov/professionals/quality-patient-safety/patientsafetyculture/resources/index.html. Accessed 2014

AHRQ Rockville, M.International Use of the Surveys on Patient Safety Culture. http://www.ahrq.gov/professionals/quality-patient-safety/patientsafetyculture/pscintusers.html. Accessed 2015

Bally, B. and P. Chevalier: Retour d'expérience en santé (REX) : Comprendre et mettre en œuvre Publications HAS 2014.

Bates, D. W., I. Larizgoitia, N. Prasopa-Plaizier, A. K. Jha and W. H. O. W. A. f. P. S. Research Priority Setting Working Group of the: Global priorities for patient safety research. BMJ 2009, 338:-b1775.

Boussat, B., T. Bougerol, O. Detante, A. Seigneurin and P. Francois: Experience Feedback Committee: a management tool to improve patient safety in mental health. Ann Gen Psychiatry 2015, 14:-23.

Boussat, B., A. Seigneurin, J. Giai, K. Kamalanavin, J. Labarere and P. Francois: Involvement in Root Cause Analysis and Patient Safety Culture Among Hospital Care Providers. J Patient Saf 2017.

Bowie, P., J. Skinner and C. de Wet: Training health care professionals in root cause analysis: a cross-sectional study of post-training experiences, benefits and attitudes. BMC Health Serv Res 2013, 13:-50.

Braithwaite, J., J. Herkes, K. Ludlow, L. Testa and G. Lamprell: Association between organisational and workplace cultures, and patient outcomes: systematic review. BMJ Open 2017, 7:11-e017708.

Braithwaite, J., M. T. Westbrook, N. A. Mallock, J. F. Travaglia and R. A. Iedema: Experiences of health professionals who conducted root cause analyses after undergoing a safety improvement programme. Qual Saf Health Care 2006, 15:6-393-399.

Caporossi, A., E. Brudieu, A. Lehmann, A. Seigneurin and P. François: An Experience Feedback Committee for Improving Medication Process Safety: An Observational Study in a Hospital Pharmacy Department. J Pharma Care Health Sys 2014, S1-010.

Champion, C., J. Sadek and H. Moloo: Strategic Change in Surgical Quality Improvement: The Ottawa Hospital (TOH) Comprehensive Unit-Based Safety Program (CUSP) Experience. Healthc Q 2017, 20:2-69-71.

Colla, J. B., A. C. Bracken, L. M. Kinney and W. B. Weeks: Measuring patient safety climate: a review of surveys. Qual Saf Health Care 2005, 14:5-364-366.

Darzi, A.: A time for revolutions--the role of clinicians in health care reform. N Engl J Med 2009, 361:6-e8.

Debouck, F., H. B. Petit and E. Lartigau: [From Crex mutualisation to clinical audit]. Cancer Radiother 2010, 14:6-7-571-575.

Debouck, F., E. Rieger, H. Petit, G. Noel and L. Ravinet: [ORION(R): a simple and effective method for systemic analysis of clinical events and precursors occurring in hospital practice]. Cancer Radiother 2012, 16:3-201-208.

Deering, S., M. A. Rosen, V. Ludi, M. Munroe, A. Pocrnich, C. Laky and P. G. Napolitano: On the front lines of patient safety: implementation and evaluation of team training in Iraq. Jt Comm J Qual Patient Saf 2011, 37:8-350-356.

Duclos, A., J. L. Peix, V. Piriou, P. Occelli, A. Denis, S. Bourdy, M. J. Carty, A. A. Gawande, F. Debouck, C. Vacca, J. C. Lifante, C. Colin and I. S. Group: Cluster randomized trial to evaluate the impact of team training on surgical outcomes. Br J Surg 2016, 103:13-1804-1814.

Faris, P. D., W. A. Ghali, R. Brant, C. M. Norris, P. D. Galbraith, M. L. Knudtson and A. I. A. P. P. f. O. A. i. C. H. Disease: Multiple imputation versus data enhancement for dealing with missing data in observational health care outcome analyses. J Clin Epidemiol 2002, 55:2-184-191.

Feng, X., K. Bobay, J. W. Krejci and B. L. McCormick: Factors associated with nurses' perceptions of patient safety culture in China: a cross-sectional survey study. J Evid Based Med 2012, 5:2-50-56.

Flin, R., C. Burns, K. Mearns, S. Yule and E. M. Robertson: Measuring safety climate in health care. Qual Saf Health Care 2006, 15:2-109-115.

Forster, A. J., B. Bernard, S. E. Drosler, Y. Gurevich, J. Harrison, J. M. Januel, P. S. Romano, D. A. Southern, V. Sundararajan, H. Quan, S. E. Vanderloo, H. A. Pincus and W. A. Ghali: A World Health Organization field trial assessing a proposed ICD-11 framework for classifying patient safety events. Int J Qual Health Care 2017, 29:4-548-556.

Francois, P., F. Prate, G. Vidal-Trecan, J. F. Quaranta, J. Labarere and E. Sellier: Characteristics of morbidity and mortality conferences associated with the implementation of patient safety improvement initiatives, an observational study. BMC Health Serv Res 2016, 16:-35.

Francois, P., E. Sellier, F. Imburchia and M. R. Mallaret: Experience feedback committee: a method for patient safety improvement. Rev Epidemiol Sante Publique 2013, 61:2-155-161.

Francois, P., E. Sellier, F. Imburchia and M. R. Mallaret: [Experience feedback committee: a method for patient safety improvement]. Rev Epidemiol Sante Publique 2013, 61:2-155-161.

Fujita, S., K. Seto, S. Ito, Y. Wu, C. C. Huang and T. Hasegawa: The characteristics of patient safety culture in Japan, Taiwan and the United States. BMC Health Serv Res 2013, 13:-20.

Giai, J., B. Boussat, P. Occelli, G. Gandon, A. Seigneurin, P. Michel and P. Francois: Hospital survey on patient safety culture (HSOPS): variability of scoring strategies. Int J Qual Health Care 2017, 29:5-685-692.

Hammer, A., N. Ernstmann, O. Ommen, M. Wirtz, T. Manser, Y. Pfeiffer and H. Pfaff: Psychometric properties of the Hospital Survey on Patient Safety Culture for hospital management (HSOPS_M). BMC Health Serv Res 2011, 11:-165.

Ito, S., K. Seto, M. Kigawa, S. Fujita, T. Hasegawa and T. Hasegawa: Development and applicability of Hospital Survey on Patient Safety Culture (HSOPS) in Japan. BMC Health Serv Res 2011, 11:-28.

Januel, J. M., G. Chen, C. Ruffieux, H. Quan, J. D. Douketis, M. A. Crowther, C. Colin, W. A. Ghali, B. Burnand and I. Group: Symptomatic in-hospital deep vein thrombosis and pulmonary embolism following hip and knee arthroplasty among patients receiving recommended prophylaxis: a systematic review. JAMA 2012, 307:3-294-303.

Jones, K. J., A. M. Skinner, R. High and R. Reiter-Palmon: A theory-driven, longitudinal evaluation of the impact of team training on safety culture in 24 hospitals. BMJ Qual Saf 2013, 22:5-394-404.

Khan, R. M., M. Aljuaid, H. Aqeel, M. M. Aboudeif, S. Elatwey, R. Shehab, Y. Mandourah, K. Maghrabi, H. Hawa, I. Khalid, I. Qushmaq, A. Latif, B. Chang, S. M. Berenholtz, S. Tayar, K. Al-Harbi, A. Yousef, A. A. Amr and Y. M. Arabi: Introducing the

Comprehensive Unit-based Safety Program for mechanically ventilated patients in Saudi Arabian Intensive Care Units. Ann Thorac Med 2017, 12:1-11-16.

Lartigau, E., B. Coche-Dequeant, V. Dumortier, S. Giscard, T. Lacornerie, A. Lasue, V. Cheval, V. Martel, B. Malfait, A. Fuchs, M. Pestel, M. Damman and M. Forrest: [Experience feed back committee in radiotherapy (CREx): a compulsory tool for security management evaluation at two years]. Cancer Radiother 2008, 12:6-7-610-613.

Lartigau, E., A. Vitoux and F. Debouck: [Crex and Orion analysis in radiation oncology: towards a mutualisation of corrective actions]. Cancer Radiother 2009, 13:6-7-458-460.

Leape, L. L. and D. M. Berwick: Five years after To Err Is Human: what have we learned? JAMA 2005, 293:19-2384-2390.

Lecoanet, A., E. Sellier, F. Carpentier, M. Maignan, A. Seigneurin and P. Francois: Experience feedback committee in emergency medicine: a tool for security management. Emerg Med J 2013.

Makary, M. A. and M. Daniel: Medical error-the third leading cause of death in the US. BMJ 2016, 353:-i2139.

Mallouli, M., M. A. Tlili, W. Aouicha, M. Ben Rejeb, C. Zedini, A. Salwa, A. Mtiraoui, M. Ben Dhiab and T. Ajmi: Assessing patient safety culture in Tunisian operating rooms: A multicenter study. Int J Qual Health Care 2017.

Mannion, R. and J. Smith: Hospital culture and clinical performance: where next? BMJ Qual Saf 2017.

MeaH: Améliorer la sécurité des organisations de soins – Exploiter les retours d'expériences. Publications ANAP 2008.

Michel, P.: Pour une politique globale et intégrée de sécurités des patients: principes et propositions. In.: Haut Conseil de la Santé Publique; 2011.

Michel, P., C. Minodier and M. Lathelize: Les événements indésirables graves associés aux soins observés dans les établissements de santé. Résultats des enquêtes nationales menées en 2009 et 2004. Dossiers solidarité et santé 2010, 17:1-17.

Michel, P., J. L. Quenon, A. Djihoud, S. Tricaud-Vialle and A. M. de Sarasqueta: French national survey of inpatient adverse events prospectively assessed with ward staff. Qual Saf Health Care 2007, 16:5-369-377.

Nath, S. B. and S. C. Marcus: Medical errors in psychiatry. Harv Rev Psychiatry 2006, 14:4-204-211.

Neily, J., P. D. Mills, Y. Young-Xu, B. T. Carney, P. West, D. H. Berger, L. M. Mazzia, D. E. Paull and J. P. Bagian: Association between implementation of a medical team training program and surgical mortality. JAMA 2010, 304:15-1693-1700.

Occelli, P.La culture de sécurité des soins: du concept à la pratique. http://www.has-sante.fr/portail/jcms/c_1025226/le-developpement-d-une-culture-de-securite. Accessed 2010

Occelli, P., J. L. Quenon, B. Hubert, T. Kosciolek, H. Hoarau, M. L. Pouchadon, R. Amalberti, Y. Auroy, L. R. Salmi, M. Sibe, P. Parneix and P. Michel: Development of a safety culture: initial measurements at six hospitals in France. J Healthc Risk Manag 2011, 30:4-42-47.

Occelli, P., J. L. Quenon, M. Kret, S. Domecq, F. Delaperche, O. Claverie, B. Castets-Fontaine, R. Amalberti, Y. Auroy, P. Parneix and P. Michel: Validation of the French version of the Hospital Survey on Patient Safety Culture questionnaire. Int J Qual Health Care 2013, 25:4-459-468.

OMS: Sécurité du patient. www.who.int/topics/patient_safety/fr/ 2004.

Peerally, M. F., S. Carr, J. Waring and M. Dixon-Woods: The problem with root cause analysis. BMJ Qual Saf 2017, 26:5-417-422.

Percarpio, K. B., B. V. Watts and W. B. Weeks: The effectiveness of root cause analysis: what does the literature tell us? Jt Comm J Qual Patient Saf 2008, 34:7-391-398.

Perneger, T. V., A. Staines and F. Kundig: Internal consistency, factor structure and construct validity of the French version of the Hospital Survey on Patient Safety Culture. BMJ Qual Saf 2014, 23:5-389-397.

Pitts, S. I., N. M. Maruthur, N. P. Luu, K. Curreri, R. Grimes, C. Nigrin, H. F. Sateia, M. D. Sawyer, P. J. Pronovost, J. M. Clark and K. S. Peairs: Implementing the Comprehensive Unit-Based Safety Program (CUSP) to Improve Patient Safety in an Academic Primary Care Practice. Jt Comm J Qual Patient Saf 2017, 43:11-591-597.

Pronovost, P. and B. Sexton: Assessing safety culture: guidelines and recommendations. Qual Saf Health Care. 2005, 14:4-231-233.

Pronovost, P. J. and J. A. Marsteller: A physician management infrastructure. JAMA 2011, 305:5-500-501.

Quan, H., S. Drosler, V. Sundararajan, E. Wen, B. Burnand, C. M. Couris, P. Halfon, J. M. Januel, E. Kelley, N. Klazinga, J. C. Luthi, L. Moskal, E. Pradat, P. S. Romano, J. Shepheard, L. So, L. Sundaresan, L. Tournay-Lewis, B. Trombert-Paviot, G. Webster and W. A. Ghali (2008). Adaptation of AHRQ Patient Safety Indicators for Use in ICD-10 Administrative Data by an International Consortium. Advances in Patient Safety: New Directions and Alternative Approaches (Vol. 1: Assessment). K. Henriksen, J. B. Battles, M. A. Keyes and M. L. Grady. Rockville (MD).

Reason, J.: Human error: models and management. BMJ 2000, 320:7237-768-770.

Schmutz, J. and T. Manser: Do team processes really have an effect on clinical performance? A systematic literature review. Br J Anaesth 2013, 110:4-529-544.

Scott, T., R. Mannion, H. T. Davies and M. N. Marshall: Implementing culture change in health care: theory and practice. Int J Qual Health Care. 2003, 15:2-111-118.

Sellier, E., S. David-Tchouda, G. Bal and P. Francois: Morbidity and mortality conferences: their place in quality assessments. Int J Health Care Qual Assur 2012, 25:3-189-196.

Smith, L. E. and S. A. Flanders: Application of a Comprehensive Unit-based Safety Program in critical care: the royal exchange. Crit Care Nurs Clin North Am 2014, 26:4-447-460.

Smits, M., I. Christiaans-Dingelhoff, C. Wagner, G. Wal and P. P. Groenewegen: The psychometric properties of the 'Hospital Survey on Patient Safety Culture' in Dutch hospitals. BMC Health Serv Res 2008, 8:-230.

Sorra, J., T. Famolaro, N. Dyer, S. A. Smith, S. Wilson and H. Liu.Hospital Survey on Patient Safety Culture: 2014 User Comparative Database Report http://www.ahrq.gov/sites/default/files/wysiwyg/professionals/quality-patient-safety/patientsafetyculture/hospital/2014/hsops14pt1.pdf. Accessed 2014

Southern, D. A., B. Burnand, S. E. Droesler, W. Flemons, A. J. Forster, Y. Gurevich, J. Harrison, H. Quan, H. A. Pincus, P. S. Romano, V. Sundararajan, N. Kostanjsek and W. A. Ghali: Deriving ICD-10 Codes for Patient Safety Indicators for Large-scale Surveillance Using Administrative Hospital Data. Med Care 2017, 55:3-252-260.

Thellier, S.: Les professionnels de la radiothérapie face à l'obligation d'améliorer la sécurité des traitements. 2011.

van Everdingen JJ, S. S., Schellekens W, Cucic S. : SimPaTIE Project. Houten: Bohn Stafleu van Loghum. Patient safety toolbox: instruments for improving safety in health care organisations; p. 19. Available from:. http://www.simpatie.org/Main/files/fsf1175778411/S.WP6.final.original.pdf. 2006.

Vincent, C., S. Taylor-Adams, E. J. Chapman, D. Hewett, S. Prior, P. Strange and A. Tizzard: How to investigate and analyse clinical incidents: clinical risk unit and association of litigation and risk management protocol. Bmj 2000, 320:7237-777-781.

Vlayen, A., J. Hellings, N. Claes, H. Peleman and W. Schrooten: A nationwide hospital survey on patient safety culture in Belgian hospitals: setting priorities at the launch of a 5-year patient safety plan. BMJ Qual Saf 2012, 21:9-760-767.

Vlayen, A., W. Schrooten, W. Wami, M. Aerts, L. G. Barrado, N. Claes and J. Hellings: Variability of patient safety culture in Belgian acute hospitals. J Patient Saf 2015, 11:2-110-121.

Wallace, L. M., P. Spurgeon, S. Adams, L. Earll and J. Bayley: Survey evaluation of the National Patient Safety Agency's Root Cause Analysis training programme in England and Wales: knowledge, beliefs and reported practices. Qual Saf Health Care 2009, 18:4-288-291.

Ward, M. M., X. Zhu, M. Lampman and G. L. Stewart: TeamSTEPPS implementation in community hospitals: adherence to recommended training approaches. Int J Health Care Qual Assur 2015, 28:3-234-244.

Weaver, S. J., J. Lofthus, M. Sawyer, L. Greer, K. Opett, C. Reynolds, R. Wyskiel, S. Peditto and P. J. Pronovost: A Collaborative Learning Network Approach to Improvement: The CUSP Learning Network. Jt Comm J Qual Patient Saf 2015, 41:4-147-159.

Weaver, S. J., L. H. Lubomksi, R. F. Wilson, E. R. Pfoh, K. A. Martinez and S. M. Dy: Promoting a culture of safety as a patient safety strategy: a systematic review. Ann Intern Med 2013, 158:5 Pt 2-369-374.

William "Bill" Frist, S. f. T. R.: Text of the Healthcare Research and Quality Act of 1999. https://www.govtrack.us/congress/bills/106/s580/text 1999.

World Alliance For Patient Safety Drafting, G., H. Sherman, G. Castro, M. Fletcher, S. World Alliance for Patient, M. Hatlie, P. Hibbert, R. Jakob, R. Koss, P. Lewalle, J. Loeb, T. Perneger, W. Runciman, R. Thomson, T. Van Der Schaaf and M. Virtanen: Towards an

International Classification for Patient Safety: the conceptual framework. Int J Qual Health Care 2009, 21:1-2-8.

Woynar, S., F. Debouck, P. Cellier, J. Bourhis, M. Cauterman and E. Lartigau: [Towards global security in radiation oncology. Using event notification feed-back]. Cancer Radiother 2007, 11:6-7-320-328.

Wu, A. W.: Medical error: the second victim. The doctor who makes the mistake needs help too. BMJ 2000, 320:7237-726-727.

Wu, A. W., A. K. Lipshutz and P. J. Pronovost: Effectiveness and efficiency of root cause analysis in medicine. JAMA 2008, 299:6-685-687.

Zohar, D., Y. Livne, O. Tenne-Gazit, H. Admi and Y. Donchin: Healthcare climate: a framework for measuring and improving patient safety. Crit Care Med 2007, 35:5-1312-1317.

7. VALORISATION DES TRAVAUX DE THESE

7.1. Articles publiés

Boussat B, Bougerol T, Detante O, Seigneurin A, Francois P. Experience Feedback Committee: a management tool to improve patient safety in mental health. *Annals of general psychiatry*. 2015;14:23. Doi: 10.1186/s12991-015-0062-2.

Giai J, Boussat B, Occelli P, Gandon G, Seigneurin A, Michel P, François P. Hospital Survey on Patient Safety Culture: variability of scoring strategies, *International Journal For Quality in Healthcare.* 29:5-685-692. doi:10.1093/intqhc/mzx086.

Boussat B, François P, Gandon G, Seigneurin A, Giai J, Perneger T, Labarère J. Inconsistencies Between Two Cross-Cultural Adaptations of the Hospital Survey on Patient Safety Culture into French, *Journal of Patient Safety*, 2017, doi: 10.1097/PTS.0000000000000443.

Boussat B, Seigneurin A, Giai J, Kamalanavin K, Labarère J, François P. Involvement in root cause analysis and patient safety culture among hospital care providers, *Journal of Patient Safety 2017,* doi: 10.1097/PTS.0000000000000456.

7.2. Articles soumis

François P, Lecoanet A, Caporossi A, Dols AM, Seigneurin A, Boussat B. Experience feedback committees: implementation and functioning in hospital medical departments.

Bastien Boussat, Julien Viotti, Olivier François, Joris Giai, Arnaud Seigneurin, José Labarère, Patrice François. Dealing with missing data in the Hospital Survey on Patient Safety Culture: a simulation study.

Boussat B, Kamalanavin K, François P. The contribution of open comments to understanding the results from the Hospital Survey on Patient Safety Culture (HSOPS): a qualitative study (article en révision mineur dans PLoS One).

Kamalanavin K, François P, Dols AM, Seigneurin A, Boussat B. Revues de Mortalité et de Morbidité et Comités de retour d'Expérience : la perception des soignants.

8. ANNEXES

8.1. Questionnaire de l'enquête culture sécurité

MESURE DE LA CULTURE DE SECURITE DES SOINS EN MILIEU HOSPITALIER

OUTIL DE MESURE

Nous souhaitons connaître votre opinion sur les questions de sécurité des soins et de signalement des évènements indésirables et dans votre service et dans l'établissement.
Le temps estimé pour remplir ce questionnaire est de 10 à 15 minutes. Il est confidentiel et anonyme aucune réponse individuelle ne sera communiquée.
Pour toute question relative à ce questionnaire, contactez **Mme Claudette Bois** au 06 08 80 66 34 ou CBois1@chu-grenoble.fr

Rappel de certaines définitions :

Un **évènement indésirable** est un dommage survenant chez le patient au cours de sa prise en charge, lié aux activités de soins.
Un **évènement indésirable lié aux soins** est consécutif aux stratégies et actes de prévention, de diagnostic, de traitement, ou de réhabilitation.
La **sécurité des soins** est l'absence d'évènement indésirable lié aux soins.
Une **erreur médicale** est une erreur commise au cours de la délivrance des soins par un professionnel de santé. Une erreur peut être à l'origine d'un événement indésirable.

Consigne de remplissage exemple

Entourez la réponse la plus appropriée… **1** **(2)** **3** **4** **5**

Identification du service ou de l'unité fonctionnelle : ……

A : Votre **service** ou votre **unité fonctionnelle**

Indiquez votre **accord** ou votre **désaccord** avec les affirmations suivantes concernant votre **service** ou votre **unité fonctionnelle**…		Pas du tout d'accord	Pas d'accord	Neutre	D'accord	Tout à fait d'accord
1.	Les personnes se soutiennent mutuellement dans le service	1	2	3	4	5
2.	Nous avons suffisamment de personnel pour faire face à la charge de travail	1	2	3	4	5
3.	Quand une importante charge de travail doit être effectuée rapidement, nous conjuguons nos efforts en équipe	1	2	3	4	5
4.	Dans le service, chacun considère les autres avec respect	1	2	3	4	5
5.	Le nombre d'heures de travail des professionnels de l'équipe est trop important pour assurer les meilleurs soins	1	2	3	4	5
6.	Nous menons des actions afin d'améliorer la sécurité des soins	1	2	3	4	5
7.	Nous faisons trop appel à du personnel intérimaire pour une meilleure qualité des soins	1	2	3	4	5
8.	Le personnel a l'impression que ses erreurs lui sont reprochées	1	2	3	4	5
9.	Dans notre service, les erreurs ont conduit à des changements positifs	1	2	3	4	5
10.	C'est uniquement par hasard s'il n'y a pas eu des erreurs plus graves dans le service jusqu'ici	1	2	3	4	5
11.	Quand l'activité d'un secteur du service est très dense, les autres secteurs lui viennent en aide	1	2	3	4	5
12.	Au contact des collègues du service, nous améliorons nos pratiques en termes de sécurité des soins	1	2	3	4	5
13.	Lorsqu'un évènement est signalé, on a l'impression que c'est la personne qui est pointée du doigt et non le problème	1	2	3	4	5

1/4

Indiquez votre **accord** ou votre **désaccord** avec les affirmations suivantes concernant votre **service** ou votre **unité fonctionnelle**…		Pas du tout d'accord	Pas d'accord	Neutre	D'accord	Tout à fait d'accord
14.	Après avoir mis en place des actions d'amélioration de la sécurité des soins, nous évaluons leur efficacité	**1**	**2**	**3**	**4**	**5**
15	Nous travaillons en mode de crise, en essayant de faire trop de choses, trop rapidement	**1**	**2**	**3**	**4**	**5**
16.	La sécurité des soins n'est jamais négligée au profit d'un rendement plus important	**1**	**2**	**3**	**4**	**5**
17	Le personnel s'inquiète du fait que les erreurs soient notées dans les dossiers administratifs du personnel	**1**	**2**	**3**	**4**	**5**
18.	Nous avons des problèmes de sécurité des soins dans ce service	**1**	**2**	**3**	**4**	**5**
19.	Notre fonctionnement et nos procédures sont efficaces pour prévenir la survenue d'erreurs	**1**	**2**	**3**	**4**	**5**

B : Votre supérieur hiérarchique immédiat (cadre de santé, chef de service, référent…)

*Pour les **médecins libéraux** ou les **chefs de service,** le supérieur hiérarchique est la Direction de l'établissement, la Direction médicale ou la Direction de pôle*

Indiquez votre **accord** ou votre **désaccord** avec les affirmations suivantes concernant votre **service** ou votre **unité fonctionnelle**…		Pas du tout d'accord	Pas d'accord	Neutre	D'accord	Tout à fait d'accord
1.	Mon supérieur hiérarchique immédiat exprime sa satisfaction quand il/elle voit un travail réalisé dans le respect des règles de sécurité des soins	**1**	**2**	**3**	**4**	**5**
2.	Mon supérieur hiérarchique immédiat tient vraiment compte des suggestions du personnel pour améliorer la sécurité des soins	**1**	**2**	**3**	**4**	**5**
3.	Chaque fois que la pression augmente, mon supérieur hiérarchique immédiat veut nous faire travailler plus rapidement, même si c'est au détriment de la sécurité	**1**	**2**	**3**	**4**	**5**
4.	Mon supérieur hiérarchique immédiat néglige les problèmes récurrents de sécurité des soins	**1**	**2**	**3**	**4**	**5**

C : Communication

Avec quelle fréquence les situations suivantes surviennent-elles dans votre **service** ou votre **unité fonctionnelle** ?		Jamais	Rarement	De temps en temps	La plupart du temps	Toujours
1.	Nous recevons un retour d'information sur les actions mises en place suite au signalement d'un évènement	**1**	**2**	**3**	**4**	**5**
2.	Le personnel s'exprime librement s'il voit quelque chose dans les soins qui peut avoir des conséquences négatives sur les patients	**1**	**2**	**3**	**4**	**5**
3.	Nous sommes informés des erreurs qui se produisent dans ce service	**1**	**2**	**3**	**4**	**5**
4.	Le personnel se sent libre de remettre en cause les décisions ou les actions de ses supérieurs	**1**	**2**	**3**	**4**	**5**
5.	Dans ce service, nous discutons des moyens à mettre en place afin que les erreurs ne se reproduisent pas	**1**	**2**	**3**	**4**	**5**
6.	Le personnel a peur de poser des questions quand quelque chose ne semble pas être correct	**1**	**2**	**3**	**4**	**5**

D : Fréquence de signalement des évènements indésirables

Lorsque les erreurs suivantes surviennent, avec quelle fréquence sont-elles signalées dans votre **service** ou **unité fonctionnelle ?**	Jamais	Rarement	De temps en temps	La plupart du temps	Toujours
1. Quand une erreur est faite, mais est détectée et corrigée avant d'avoir affecté le patient, elle est signalée…	**1**	**2**	**3**	**4**	**5**
2. Quand une erreur est faite, mais n'a pas le potentiel de nuire au patient, elle est signalée…	**1**	**2**	**3**	**4**	**5**
3. Quand une erreur est faite et qu'elle pourrait nuire au patient mais qu'elle n'a finalement pas d'effet, elle est signalée…	**1**	**2**	**3**	**4**	**5**

E : Niveau de sécurité des soins

Globalement, à quel niveau situez-vous la sécurité des soins dans votre **service** ou votre **unité fonctionnelle** ?
Entourer **UNE** seule réponse.

A. Excellent B. Très Bon C. Acceptable D. Faible E. Défaillant

F : Votre établissement de santé

Indiquez votre **accord** ou votre **désaccord** avec les affirmations suivantes concernant votre **service** ou votre **unité fonctionnelle**…	Pas du tout d'accord	Pas d'accord	Neutre	D'accord	Tout à fait d'accord
1. La Direction de l'établissement instaure un climat de travail qui favorise la sécurité des soins	**1**	**2**	**3**	**4**	**5**
2. Les services de l'établissement ne se coordonnent pas bien les uns avec les autres	**1**	**2**	**3**	**4**	**5**
3. Des dysfonctionnements surviennent quand les patients sont transférés d'une unité à l'autre	**1**	**2**	**3**	**4**	**5**
4. Il y a une bonne coopération entre les services qui doivent travailler ensemble	**1**	**2**	**3**	**4**	**5**
5. D'importantes informations concernant les soins des patients sont souvent perdues lors des changements d'équipes	**1**	**2**	**3**	**4**	**5**
6. Il est souvent déplaisant de travailler avec le personnel des autres services de l'établissement	**1**	**2**	**3**	**4**	**5**
7. Des problèmes surviennent souvent dans les échanges d'information entre les services de l'établissement	**1**	**2**	**3**	**4**	**5**
8. Les actions menées par la direction de l'établissement montrent que la sécurité des soins est une des premières priorités	**1**	**2**	**3**	**4**	**5**
9. La direction de l'établissement semble s'intéresser à la sécurité des soins uniquement après qu'un évènement indésirable se soit produit	**1**	**2**	**3**	**4**	**5**
10. Les services de l'établissement travaillent ensemble pour fournir aux patients les meilleurs soins	**1**	**2**	**3**	**4**	**5**
11. Les changements d'équipes sont problématiques pour les patients dans l'établissement	**1**	**2**	**3**	**4**	**5**

G : Nombre d'évènements indésirables signalés

Au cours des 12 derniers mois, combien de fiches de signalement des événements indésirables avez-vous remplies et transmises ?
Entourer **UNE** seule réponse.

A. Aucune B. 1 à 2 fiches C. 3 à 5 fiches D. 6 à 10 fiches E. 11 à 20 fiches F. Plus de 20 fiches

H : Informations générales

Entourer **UNE** seule réponse.

1. Vous êtes…

A. Médecin **B.** Infirmier(e),Cadre infirmier **C.** Aide soignant(e) **D.** Sage-Femme **E.** ASH

F. Secrétaire **G.** Autre métier, préciser : .....

2. Vous êtes…

A. Une femme **B.** Un homme

3. Vous avez…

A. moins de 35ans **B.** Entre 35 à 45 ans **C.** Entre 45 et 55 ans **D**. Plus de 55 ans

4. Depuis combien d'années exercez-vous votre spécialité ou profession actuelle ?

A. Moins de 1 an **B.** 1 à 2 ans **C.** 3 à 5 ans **D**. 6 à 10 ans **E.** 11 ans ou plus

5. Depuis combien de temps travaillez-vous dans cet établissement de santé ?

A. Moins de 1 an **B.** 1 à 2 ans **C.** 3 à 5 ans **D.** 6 ans ou plus

6. Depuis combien de temps travaillez-vous dans ce service ?

A. Moins de 1 an **B.** 1 à 2 ans **C.** 3 à 5 ans **D.** 6 ans ou plus

7. Vous travaillez dans ce service …

A. Moins de 50 % de votre temps de travail **B.** Plus de 50 % de votre temps de travail

8. Participez-vous ou avez-vous participé dans cet établissement à un Comité de Retour d'Expérience (CREX) ?

A. OUI **B.** NON

9. Participez-vous ou avez-vous participé dans cet établissement à d'autres structures impliquées dans la gestion des risques (Comité de Lutte contre les Infections Nosocomiales, Comité de LUtte contre la Douleur, Vigilances, Commission du MEdicament et des DIspositifs Médicaux Stériles, etc.) ?

A. OUI **B.** NON

I : Vos commentaires

NOUS VOUS REMERCIONS D'AVOIR PARTICIPE A CETTE ENQUETE

8.2. Annexes des articles

8.2.1. Annexes article: Experience Feedback Committee: a management tool to improve patient safety in mental health

Appendix 1. ORION standardized report

ORION analysis report

Experience Feedback Committee:

Investigator:

Date:

I – <u>Incident identification: Data collection</u>

Event title:

Event date:

Event site:

Brief description of the event:

Event consequences:

Modalities for data collection:

☐ Individual interview ☐ Debriefing ☐ Consultation of medical files ☐ Site visit

II – <u>Schedule of facts</u>

Present facts in the order that they occurred and identify points of failure

Chronological description	**Guidelines and operational procedures**	**Points of failure, deviations from guidelines**
1) **Before the event**	1) **Before the event**	1) **Before the event**
2) **During the event**	2) **During the event**	2) **During the event**
3) **After the event**	3) **After the event**	3) **After the event**

III – <u>Root-cause analysis</u>

Report the causes and the latent factor for each domain

1- *Hospital policy*

2- *Organization*

3- *Working conditions*

4- *Team functioning*

5- *Operational procedures*

6- *Healthcare professionals*

7- *Patient*

8- *Summary of failures and causes*

IV – <u>Proposal for corrective actions</u>

Action plan proposed to the EFC to rectify the root causes.

Present the possible corrective actions, if possible by indicating efficiency criteria, the professional in charge of implementing the proposal, the planned deadline, an estimate of the cost, the opportunity for scaling up the action).

Corrective actions	Efficiency criteria	Deadline	Professional in charge

Analysis report validation during the EFC:

EFC members' comments:

8.2.2. Annexes article: Inconsistencies between two cross-cultural adaptations of the Hospital Survey on Patient Safety Culture into French

Online supplementary appendix 1. Item Wording Across the Original (US), Vlayen, and Occelli Versions of the Hospital Survey on Patient Safety Culture Questionnaire.

Online supplementary appendix 2: Survey participation

Online supplementary appendix 3. Respondent Characteristics

Online supplementary appendix 4. Item Characteristics for the Occelli Version of the Hospital Survey on Patient Safety Culture Questionnaire (Present Study).

Online supplementary appendix 5. Number (Percentages) of Missing Values for the Occelli Version of the Hospital Survey on Patient Safety Culture Questionnaire. Online supplementary appendix 6: Factor Loadings for the 42 Items of the Occelli Version of the Hospital Survey on Patient Safety Culture questionnaire in Exploratory Factor Analysis (Present Study) (n=3888)

Online supplementary appendix 7: Exploratory Factor Analysis for the Original (US), Vlayen, and Occelli Versions of the Hospital Survey on Patient Safety Culture Questionnaire

Online Appendix 1. Item Wording Across the Original (US), Vlayen, and Occelli Versions of the Hospital Survey on Patient Safety Culture Questionnaire.

Item	Dimension†	Original (US) version [10]	Vlayen version [15]	Occelli version [14]
A15	1	Patient safety is never sacrificed to get more work done.	La sécurité des patients n'est jamais délaissée même s'il y a beaucoup de travail.	La sécurité des soins n'est jamais négligée au profit d'un rendement plus important.
A18	1	Our procedures and systems are good at preventing errors from happening.	Nos procédures et nos systèmes de sécurité sont bons et préviennent des incidents.	Notre fonctionnement et nos procédures sont efficaces pour prévenir la survenue d'erreurs.
A10*	1	It is just by chance that more serious mistakes don't happen around here.	C'est une chance que des incidents plus sérieux n'aient pas encore eu lieu.	C'est uniquement par hasard s'il n'y a pas eu des erreurs plus graves dans ce service jusqu'ici.
A17*	1	We have patient safety problems in this unit.	Nous avons des problèmes de sécurité des patients dans notre environnement de travail.	Nous avons des problèmes de sécurité des soins dans ce service.
D1	2	When a mistake is made, but is caught and corrected before affecting the patient, how often is this reported?	Si cet incident a pu être détecté avant d'avoir affecté le patient.	Quand une erreur est faite, mais est détectée et corrigée avant d'avoir affecté le patient, elle est signalée.
D2	2	When a mistake is made, but has no potential to harm the patient, how often is this reported?	Si cet incident ne comportait aucun risque de nuire au patient.	Quand une erreur est faite, mais n'a pas le potentiel de nuire au patient, elle est signalée.
D3	2	When a mistake is made that could harm the patient, but does not, how often is this reported?	Si cet incident n'a eu aucune conséquence pour le patient même s'il avait pu nuire à sa santé.	Quand une erreur est faite et qu'elle pourrait nuire au patient mais qu'elle n'a finalement pas d'effet, elle est signalée.
B1	3	My supervisor/manager says a good word when he/she sees a job done according to established patient safety procedures.	Lorsque nous travaillons selon les procédures établies en matière de sécurité des patients, notre supérieur nous félicite.	Mon supérieur hiérarchique immédiat exprime sa satisfaction quand il/elle voit un travail réalisé dans le respect des règles de sécurité des soins.

Item	Dimension†	Original (US) version [10]	Vlayen version [15]	Occelli version [14]
B2	3	My supervisor/manager seriously considers staff suggestions for improving patient safety.	Notre supérieur prend sérieusement en considération les suggestions de l'équipe pour améliorer la sécurité des patients	Mon supérieur hiérarchique immédiat tient vraiment compte des suggestions du personnel pour améliorer la sécurité des soins.
B3*	3	Whenever pressure builds up, my supervisor/manager wants us to work faster, even if it means taking shortcuts.	. Lorsqu'il y a surcharge de travail, notre supérieur exige que nous travaillions plus vite, même si cela implique des entorses aux procédures.	Chaque fois que la pression augmente, mon supérieur hiérarchique immédiat veut nous faire travailler plus rapidement, même au détriment de la sécurité.
B4*	3	My supervisor/manager overlooks patient safety problems that happen over and over.	Notre supérieur néglige les problèmes récurrents de sécurité des patients.	Mon supérieur hiérarchique immédiat néglige les problèmes récurrents de sécurité des soins.
A6	4	We are actively doing things to improve patient safety.	Nous nous attachons activement à améliorer la sécurité des patients.	Nous menons des actions afin d'améliorer la sécurité des soins.
A9	4	Mistakes have led to positive changes here.	Des incidents ont entrainé des changements positifs dans notre manière de travailler.	Dans notre service, les erreurs ont conduit à des changements positifs.
A13	4	After we make changes to improve patient safety, we evaluate their effectiveness.	Si des actions sont mises en place en vue de l'amélioration de la sécurité des patients, nous vérifions l'efficacité.	Après avoir mis en place des actions d'amélioration de la sécurité des soins, nous évaluons leur efficacité.
A1	5	People support one another in this unit.	Les membres de l'équipe se soutiennent mutuellement.	Les personnes se soutiennent mutuellement dans le service.
A3	5	When a lot of work needs to be done quickly, we work together as a team to get the work done.	Lorsqu'il y a un pic important d'activité, nous nous entraidons pour réaliser l'ensemble des tâches.	Quand une importante charge de travail doit être effectuée rapidement, nous conjuguons nos efforts en équipe
A4	5	In this unit, people treat each other with respect.	Les membres de l'équipe se respectent.	Dans le service, chacun considère les autres avec respect.
A11	5	When one area in this unit gets really busy, others help out.	Quand une équipe est surchargée de travail, les autres lui viennent en aide.	Quand l'activité d'un secteur du service est très dense, les autres secteurs lui viennent en aide.

Item	Dimension†	Original (US) version [10]	Vlayen version [15]	Occelli version [14]
C2	6	Staff will freely speak up if they see something that may negatively affect patient care.	L'équipe peut s'exprimer librement si elle constate qu'une situation pourrait altérer la qualité des soins aux patients.	Le personnel s'exprime librement s'il voit quelque chose dans les soins qui peut avoir des conséquences négatives sur les patients.
C4	6	Staff feel free to question the decisions or actions of those with more authority.	Les membres de l'équipe se sentent libres de débattre des décisions ou actions prises par la hiérarchie.	Le personnel se sent libre de remettre en cause les décisions ou les actions de ses supérieurs.
C6*	6	Staff are afraid to ask questions when something does not seem right.	L'équipe a peur de poser des questions lorsqu'une situation ne semble pas correcte.	Le personnel a peur de poser des questions quand quelque chose ne semble pas être correct.
C1	7	We are given feedback about changes put into place based on event reports.	Nous sommes informés des actions correctives mises en place suite aux incidents enregistrés.	Nous recevons un retour d'information sur les actions mises en place suite au signalement d'un évènement.
C3	7	We are informed about errors that happen in this unit.	Nous sommes informés des incidents qui surviennent dans notre environnement de travail.	Nous sommes informés des erreurs qui se produisent dans ce service.
C5	7	In this unit, we discuss ways to prevent errors from happening again.	Dans notre équipe, nous discutons des moyens à mettre en œuvre pour éviter la répétition d'incidents.	Dans ce service, nous discutons des moyens à mettre en place afin que les erreurs ne se reproduisent pas.
A8*	8	Staff feel like their mistakes are held against them.	Les membres de l'équipe ont l'impression que leurs erreurs sont utilisées contre eux.	Le personnel a l'impression que ses erreurs lui sont reprochées.
A12*	8	When an event is reported, it feels like the person is being written up, not the problem.	Quand un incident est signalé, le coupable est recherché mais pas la cause du problème.	Lorsqu'un événement est signalé, on a l'impression que c'est la personne qui est pointée du doigt et non le problème.
A16*	8	Staff worry that mistakes they make are kept in their personnel file.	Les membres de l'équipe s'inquiètent du fait que leurs erreurs puissent figurer dans leur dossier personnel.	Le personnel s'inquiète du fait que les erreurs soient notées dans les dossiers administratifs du personnel.
A2	9	We have enough staff to handle the workload.	La dotation en personnel est suffisante pour supporter la charge de travail.	Nous avons suffisamment de personnel pour faire face à la charge de travail.

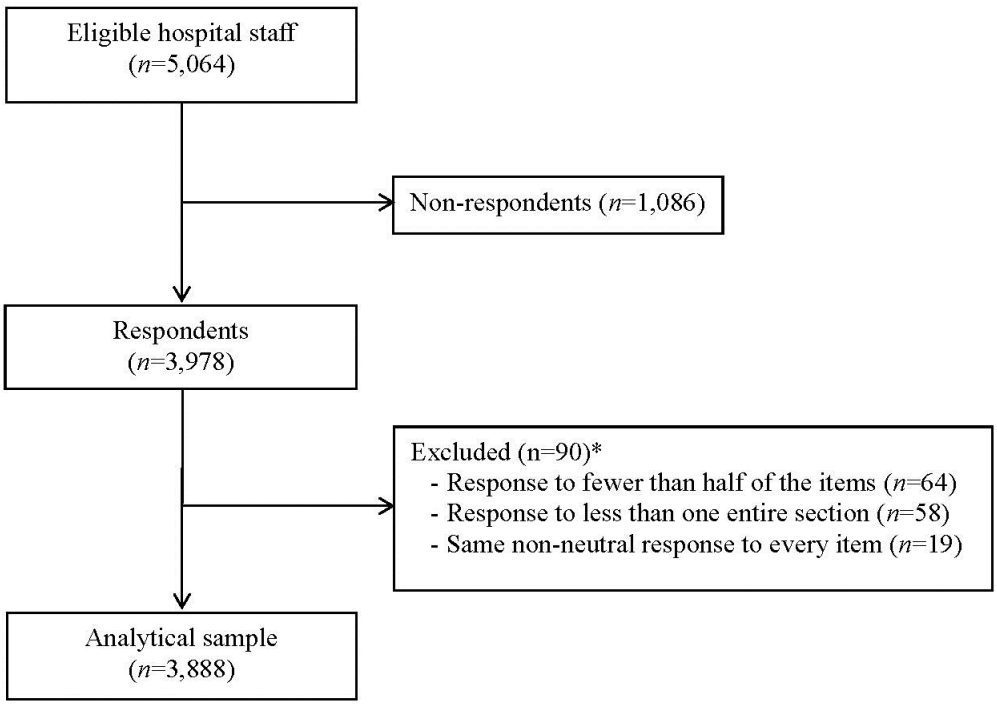
Item	Dimension†	Original (US) version [10]	Vlayen version [15]	Occelli version [14]
A5*	9	Staff in this unit work longer hours than is best for patient care.	Les membres de notre équipe prestent des heures supplémentaires qui pourraient mettre en péril la sécurité des patients.	Le nombre d'heures de travail des professionnels de l'équipe est trop important pour assurer les meilleurs soins.
A7*	9	We use more agency/temporary staff than is best for patient care.	Nous employons plus d'intérimaires que ce qui est idéal pour les soins.	Nous faisons trop appel à du personnel intérimaire pour une meilleure qualité des soins.
A14*	9	We work in "crisis mode," trying to do too much, too quickly.	Nous travaillons la plupart du temps dans l'urgence en voulant faire trop, trop vite	Nous travaillons en mode de crise, en essayant de faire trop de choses, trop rapidement.
F1	10	Hospital management provides a work climate that promotes patient safety.	La direction de l'hôpital engendre un climat de travail favorable pour la sécurité des patients.	La direction de l'établissement instaure un climat de travail qui favorise la sécurité des soins.
F8	10	The actions of hospital management show that patient safety is a top priority.	Les actions prises par la direction de l'hôpital montrent que la sécurité du patient est une priorité absolue.	Les actions menées par la direction de l'établissement montrent que la sécurité des soins est une des premières priorités.
F9*	10	Hospital management seems interested in patient safety only after an adverse event happens.	La direction de l'hôpital ne s'intéresse à la sécurité des patients qu'après un incident.	La direction de l'établissement semble s'intéresser à la sécurité des soins uniquement après qu'un évènement indésirable se soit produit.
F4	11	There is good cooperation among hospital units that need to work together.	La collaboration est bonne entre les unités/secteurs hospitaliers.	Il y a une bonne coopération entre les services qui doivent travailler ensemble.
F10	11	Hospital units work well together to provide the best care for patients.	Les différent(e)s unités/secteurs hospitaliers collaborent bien pour procurer les meilleurs soins aux patients.	Les services de l'établissement travaillent ensemble pour fournir aux patients les meilleurs soins.
F2*	11	Hospital units do not coordinate well with each other.	Les différent(e)s unités/secteurs hospitaliers ne sont pas bien coordonné(e)s entre (elles) eux.	Les services de l'établissement ne se coordonnent pas bien les uns avec les autres.

Item	Dimension†	Original (US) version [10]	Vlayen version [15]	Occelli version [14]
F6*	11	It is often unpleasant to work with staff from other hospital units.	C'est souvent désagréable de travailler avec les équipes des autres unités/secteurs hospitaliers.	Il est souvent déplaisant de travailler avec le personnel des autres services de l'établissement.
F3*	12	Things "fall between the cracks" when transferring patients from one unit to another.	Lorsque des patients sont transférés d'une unité à une autre des informations ne sont pas transmises.	Des dysfonctionnements surviennent quand les patients sont transférés d'une unité à l'autre.
F5*	12	Important patient care information is often lost during shift changes.	Des informations importantes pour les soins des patients se perdent régulièrement lors des changements d'équipe.	D'importantes informations concernant les soins des patients sont souvent perdues lors des changements d'équipes.
F7*	12	Problems often occur in the exchange of information across hospital units.	Il y a souvent des problèmes de communication entre les unités/secteurs hospitaliers.	Des problèmes surviennent souvent dans les échanges d'information entre les services de l'établissement
F11*	12	Shift changes are problematic for patients in this hospital.	Les changements d'équipe sont une source de problèmes pour les patients dans votre hôpital.	Les changements d'équipes sont problématiques pour les patients dans l'établissement.

* Negatively worded item.

† The dimensions in the original (US) version of the Hospital Survey on Patient Safety Culture questionnaire were 1) Overall perceptions of safety, 2) frequency of event reporting, 3) supervisor/manager expectations & actions promoting safety, 4) organizational learning—continuous improvement, 5) teamwork within hospital units, 6) communication openness, 7) feedback and communication about error, 8) nonpunitive response to error, 9) staffing, 10) hospital management support for patient safety, 11) teamwork across hospital units, 12) hospital handoffs & transitions.

Online Appendix 2: Survey participation



* A respondent may have more than one exclusion criterion

Online Appendix 3. Respondent Characteristics

Characteristics, *n* (%)	Present study* (*n*=3,888)		External validation study of the Vlayen version [9] † (*n*=1,171)		*P*
Female	3,016	(80.6)	910	(79.1)	0.30
Age, *year*					<0.001
< 35	1,406	(37.5)	329	(28.5)	
35–44	966	(25.7)	360	(31.1)	
45–54	967	(25.8)	341	(29.5)	
≥ 55	415	(11.1)	126	(10.9)	
Occupational group					<0.001
Nurse	1,386	(36.3)	493	(42.8)	
Nursing assistant	708	(18.6)	124	(10.8)	
Physician	436	(11.4)	118	(10.2)	
Other healthcare	124	(3.3)	94	(8.2)	
Administrative	331	(8.7)	91	(7.9)	
Technical	378	(9.7)	97	(8.4)	
Other	450	(11.8)	136	(11.8)	
Hospital sector					<0.001
Medicine and pediatrics	1,159	(29.8)	321	(27.9)	
Surgery and gynecology	833	(21.4)	211	(18.3)	
Technical (pharmacy, operating rooms, imaging, laboratories)	795	(20.4)	196	(17.0)	
Intensive care, emergency, and anesthesia	688	(17.7)	139	(12.1)	
Several or other	413	(10.6)	285	(24.7)	
Contact with patient	3,357	(88.0)	1,038	(89.9)	0.07

* The present study used the Occelli version of the Hospital Survey on Patient Safety Culture. Values were missing for gender (*n*=147), age (*n*=134), occupational group (*n*=75), and contact with patient (*n*=55).

† Values were missing for gender (*n*=21), age (*n*=15), occupational group (*n*=18), hospital sector (*n*=19), and contact with patient (*n*=17). Original (US) data were not available.

Online Appendix 4. Item Characteristics for the Occelli Version of the Hospital Survey on Patient Safety Culture Questionnaire (Present Study).

Dimension - item	Missing value*, *n* (%)		Floor effect, *n* (%)		Ceiling effect, *n* (%)		Mean (SD)		Item-rest correlation†	Average inter-item correlation‡	Cronbach's alpha#
Overall perceptions of safety	6	(0.15)	15	(0.40)	17	(0.46)	3.21	(0.71)		0.33	0.65
A10r	64	(1.65)	214	(5.60)	589	(15.40)	3.36	(1.09)	0.39	0.36	0.62
A15	57	(1.47)	264	(6.89)	302	(7.88)	3.05	(1.08)	0.39	0.36	0.62
A17r	73	(1.88)	215	(5.64)	297	(7.79)	3.15	(1.01)	0.52	0.28	0.53
A18	53	(1.36)	105	(2.74)	141	(3.68)	3.29	(0.85)	0.47	0.31	0.57
Frequency of event reporting	94	(2.42)	50	(1.33)	152	(4.05)	3.39	(0.75)		0.64	0.84
D1	106	(2.73)	106	(2.80)	222	(5.87)	3.31	(0.91)	0.67	0.67	0.80
D2	109	(2.80)	90	(2.38)	212	(5.61)	3.34	(0.88)	0.75	0.57	0.72
D3	120	(3.09	69	(1.83)	256	(6.79)	3.51	(0.82)	0.68	0.66	0.80
Supervisor expectations & actions	33	(0.85)	39	(1.02)	124	(3.25)	3.49	(0.81)		0.52	0.81
B1	45	(1.16)	293	(7.62)	421	(10.95)	3.29	(1.10)	0.59	0.55	0.78
B2	48	(1.23)	199	(5.18)	414	(10.78)	3.44	(1.01)	0.72	0.47	0.72
B3r	47	(1.21)	110	(2.86)	565	(14.71)	3.56	(0.97)	0.56	0.57	0.80
B4r	47	(1.21)	116	(3.02)	757	(19.71)	3.69	(0.97)	0.66	0.50	0.75
Organizational learning	7	(0.18)	8	(0.21)	33	(0.89)	3.4	(0.64)		0.35	0.62
A6	49	(1.26)	79	(2.06)	346	(9.01)	3.53	(0.87)	0.46	0.30	0.46
A9	63	(1.62)	120	(3.14)	191	(4.99)	3.35	(0.86)	0.42	0.35	0.51
A13	90	(2.31)	94	(2.42)	139	(3.58)	3.32	(0.82)	0.38	0.40	0.57
Teamwork within hospital units	0	(0.00)	14	(0.37)	82	(2.17)	3.54	(0.75)		0.40	0.71
A1	40	(1.03)	102	(2.65)	732	(19.02)	3.71	(0.99)	0.62	0.33	0.57
A3	21	(0.54)	61	(1.58)	701	(18.13)	3.84	(0.86)	0.54	0.38	0.63
A4	18	(0.46)	116	(3.00)	599	(15.48)	3.55	(1.01)	0.56	0.36	0.60
A11	46	(1.18)	555	(14.45)	358	(9.32)	3.05	(1.23)	0.32	0.53	0.77
Communication openness	14	(0.36)	14	(0.29)	72	(1.88)	3.5	(0.70)		0.37	0.63
C2	36	(0.93)	66	(1.71)	588	(15.26)	3.82	(0.83)	0.49	0.31	0.47
C4	41	(1.05)	311	(8.08)	130	(3.34)	2.98	(1.00)	0.43	0.39	0.56

Dimension - item	Missing value*, ***n*** (%)		Floor effect, ***n*** (%)		Ceiling effect, ***n*** (%)		Mean (SD)		Item-rest correlation†	Average inter-item correlation‡	Cronbach's alpha#
C6r	25	(0.64)	87	(2.25)	642	(16.62)	3.69	(0.94)	0.41	0.41	0.57
Feedback and communication about error	16	(0.41)	25	(0.66)	54	(1.42)	3.38	(0.73)		0.40	0.66
C1	62	(1.59)	289	(7.55)	115	(3.01)	3.03	(1.01)	0.46	0.41	0.58
C3	40	(1.03)	125	(3.25)	304	(7.90)	3.44	(0.96)	0.49	0.37	0.54
C5	31	(0.80)	100	(2.59)	410	(10.63)	3.66	(0.88)	0.47	0.41	0.58
Nonpunitive response to error	7	(0.18)	41	(1.10)	30	(0.80)	2.94	(0.74)		0.33	0.60
A8r	57	(1.47)	273	(7.13)	215	(5.61)	2.91	(1.01)	0.48	0.26	0.40
A12r	40	(1.03)	373	(9.69)	254	(6.60)	2.91	(1.10)	0.46	0.27	0.42
A16r	82	(2.11)	166	(4.36)	192	(5.04)	3.01	(0.87)	0.30	0.47	0.64
Staffing	0	(0.00)	21	(0.58)	7	(0.19)	2.89	(0.73)		0.27	0.60
A2	36	(0.93)	899	(23.34)	119	(3.09)	2.38	(1.12)	0.46	0.22	0.45
A5r	122	(3.14)	291	(7.73)	216	(5.74)	3.01	(1.00)	0.37	0.28	0.53
A7r	140	(3.60)	175	(4.67)	1238	(33.03)	3.72	(1.15)	0.24	0.36	0.63
A14r	38	(0.98)	730	(18.96)	111	(2.88)	2.45	(1.07)	0.45	0.23	0.47
Hospital management support	48	(1.23)	98	(2.62)	4	(0.11)	2.67	(0.73)		0.39	0.66
F1	86	(2.21)	562	(14.78)	31	(0.82)	2.51	(0.94)	0.49	0.37	0.54
F8	104	(2.67)	310	(8.19)	117	(3.09)	2.93	(0.96)	0.53	0.31	0.48
F9r	91	(2.34)	450	(11.85)	79	(2.08)	2.57	(0.95)	0.39	0.50	0.67
Teamwork across hospital units	47	(1.21)	5	(0.13)	3	(0.08)	3.04	(0.58)		0.24	0.56
F2r	77	(1.98)	486	(12.75)	52	(1.36)	2.45	(0.91)	0.33	0.25	0.50
F4	89	(2.29)	188	(4.95)	63	(1.66)	2.98	(0.90)	0.44	0.19	0.41
F6r	84	(2.16)	73	(1.92)	320	(8.41)	3.49	(0.85)	0.25	0.30	0.56
F10	101	(2.60)	116	(3.06)	164	(4.33)	3.27	(0.87)	0.36	0.23	0.48
Hospitals handoffs & transitions	67	(1.72)	17	(0.46)	2	(0.05)	2.88	(0.6)		0.28	0.60
F3r	118	(3.03)	259	(6.87)	45	(1.19)	2.63	(0.85)	0.35	0.29	0.55
F5r	129	(3.32)	205	(5.45)	112	(2.98)	2.96	(0.95)	0.43	0.25	0.49

Dimension - item	Missing value*, ***n*** (%)		Floor effect, ***n*** (%)		Ceiling effect, ***n*** (%)		Mean (SD)		Item-rest correlation†	Average inter-item correlation‡	Cronbach's alpha#
F7r	113	(2.91)	151	(4.00)	52	(1.38)	2.76	(0.83)	0.47	0.23	0.47
F11r	126	(3.24)	149	(3.96)	208	(5.53)	3.16	(0.92)	0.29	0.34	0.61

Abbreviations: SD = standard deviation

* Composite scores were coded as missing if all corresponding items were unanswered (see Methods).

† Item-rest correlation was computed as the correlation between an item and the composite score that was formed by all other items in the dimension.

‡ Inter-item correlation was computed by the average inter-item correlations for composite score that was formed by all other items in the dimension.

Cronbach's alpha was computed for composite score that was formed by all other items in the dimension.

Online Appendix 5. Number (Percentages) of Missing Values* for the Occelli Version of the Hospital Survey on Patient Safety Culture Questionnaire.

Composite score	Number of items	Present study (*n*=3,888)		Preliminary validation study of the Occelli version [14] (*n*=401)		*P*
Outcomes						
Overall perception of safety	4	168	(4.3)	18	(4.5)	0.87
Frequency of event reporting	3	135	(3.5)	21	(5.2)	0.07
Unit-level processes						
Supervisor/manager expectations & actions	4	72	(1.8)	9	(2.2)	0.58
Organizational learning	3	161	(4.1)	11	(2.7)	0.17
Teamwork within hospital units	4	110	(2.8)	7	(1.7)	0.20
Communication openness	3	65	(1.7)	13	(3.2)	0.02
Feedback and communication about error	3	87	(2.2)	18	(4.5)	0.01
Nonpunitive response to error	3	145	(3.7)	10	(2.5)	0.21
Staffing	4	287	(7.4)	15	(3.7)	0.01
Hospital-level processes						
Hospital management support	3	152	(3.9)	18	(4.5)	0.57
Teamwork across hospital units	4	160	(4.1)	13	(3.2)	0.40
Hospital handoffs & transitions	4	203	(5.2)	18	(4.5)	0.53

* Dimension scores were coded as missing if one or more items were unanswered (see Methods)

Online Appendix 6: Factor Loadings for the 42 Items of the Occelli Version of the Hospital Survey on Patient Safety Culture questionnaire in Exploratory Factor Analysis (Present Study) (*n*=3888)

Factor loadings (>0.4)	1	2	3	4	5	6	7	8	9	10
A6. Doing this to improve safety	0.42									
A13. Evaluate effectiveness of changes	0.40									
B2. Supervisor considers staff suggestions		0.53								
B1. Supervisor says good word		0.53								
B4r. Supervisor overlooks problems		0.46								
B3r. Supervisor wants faster work, shortcuts		0.43								
A1. People support one another			0.54							
A4. People treat each other with respect			0.49							
A3. Work together as a team			0.47							
D2. Mistake has no potential to harm				0.59						
D1. Mistake caught and corrected				0.57						
D3. Mistake could harm				0.54						
C2. Freely speak up					0.50					
C3. Informed about errors that occurred					0.46					
C4. Feel free to question decisions and actions					0.44					
F8. Safety is a top priority						0.55				
F9r. Interest only after adverse events						0.50				
F1. Management provides good work climate						0.47				
F4. Good cooperation among units							0.48			
F3r. Things "fall between the cracks"							0.43			
F2r. Units do not coordinate well							0.40			
A8r. Mistakes held against them								0.50		
A12r. Person written up when event reported								0.49		
A16r. Mistakes kept in their personnel file								0.44		
A5r. Work longer hours									0.55	
A2. Enough staff to handle workload									0.51	
A14r. Work in "crisis mode"									0.43	
F11r. Shift changes are problematic										0.53
F5r. Lost information during shift changes										0.49
A7r. Use more agency/temporary staff										
A9. Mistakes led to positive changes										
A10r. Just by chance no more serious mistakes										
A11.Others help out when busy										
A15. Safety never sacrificed										
A17r. We have safety problems										
A18. Good at preventing errors										
C1. Feedback about changes										
C5. Discuss ways of preventing errors										
C6r. Staff afraid to ask questions										
F6r. Unpleasant with other units										
F7r. Problems in the exchange of information										
F10. Units work well together										

Online Appendix 7: Exploratory Factor Analysis for the Original (US), Vlayen, and Occelli Versions of the Hospital Survey on Patient Safety Culture Questionnaire.

Original (US) [10]		External validation study of the Vlayen version [18]		Present study*	
Dimensions	Items	Dimensions	Items	Dimensions	Items
1. Overall perceptions of safety	A10r, A15, A17r, A18	1. Overall perceptions of safety	A10r, A15, A17r, A18		
2. Frequency of event reporting	D1, D2, D3	2. Frequency of event reporting	D1, D2, D3	1. Frequency of event reporting	D1, D2, D3
3. Supervisor expectations & actions	B1, B2, B3r, B4r	3. Supervisor expectations & actions	B1, B2, B3r, B4r	2. Supervisor expectations & actions	B1, B2, B3r, B4r
4. Organizational learning	A6, A9, A13	4. Organizational learning	A6, A9, A13	3. Organizational learning	A6, A13
5. Teamwork within hospital units	A1, A3, A4, A11	5. Teamwork within hospital units	A1, A3, A4, A11	4. Teamwork within hospital units	A1, A3, A4
6. Communication openness	C2, C4, C6r	6. Communication openness + Feedback and communication about error	C1, C2, C3, C4, C5	5. Communication openness	C2, C3, C4
7. Feedback and communication about error	C1, C3, C5				
8. Nonpunitive response to error	A8r, A12r, A16r	7. Nonpunitive response to error	A8r, A12r, A16r, C6r	6. Nonpunitive response to error	A8r, A12r, A16r
9. Staffing	A2, A5r, A7r, A14r	8. Staffing	A2, A5r, A7r, A14r	7. Staffing	A2, A5r, A14r
10. Hospital management support for patient safety	F1, F8, F9r	9. Hospital management support for patient safety	F1, F8, F9r	8. Hospital management support for patient safety	F1, F8, F9r
11. Teamwork across hospital units	F2r, F4, F6r, F10	10. Teamwork across hospital units + Hospital handoffs and transitions	F2r, F3r, F4, F5r, F6r, F7r, F10, F11r	9. Teamwork across hospital units	F2r, F4, F3r
12. Hospital handoffs and transitions	F3r, F5r, F7r, F11r			10. Hospital handoffs and transitions	F5r, F11r
Items with primary loading <0.40					A7r, A9, A10r, A11, A15, A17r, A18, C1, C5, C6r, F6r, F7r, F10

* The present study used the Occelli version of the Hospital Survey on Patient Safety Culture

8.2.3. Annexes article: Hospital survey on patient safety culture (HSOPS): variability of scoring strategies

List of Supplemental Digital Content

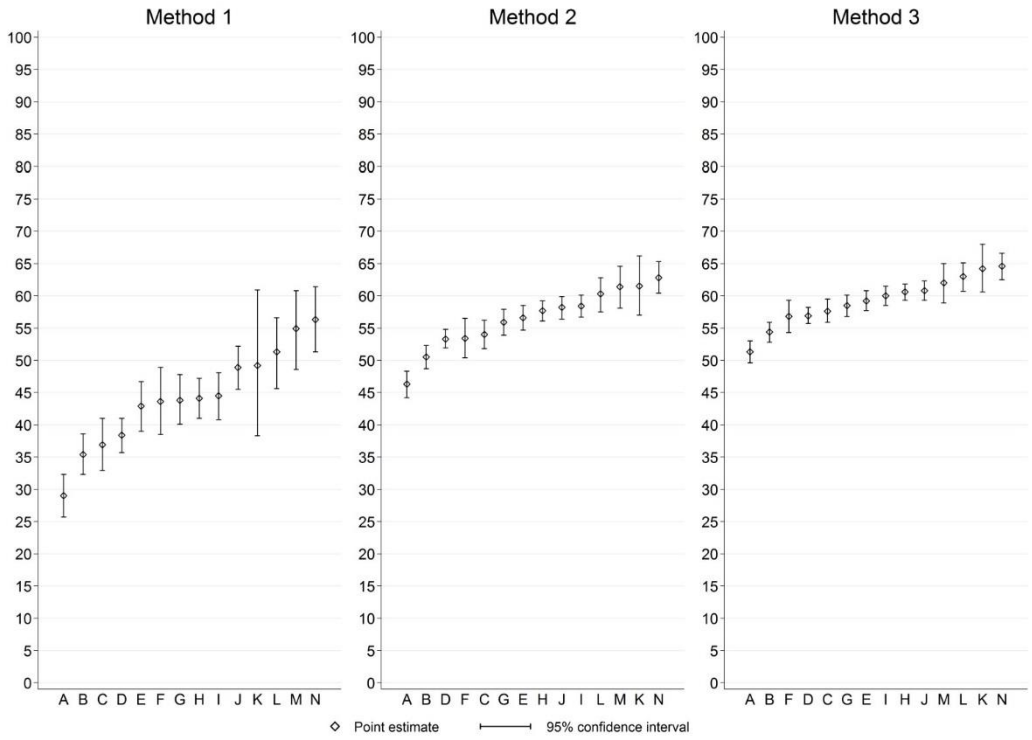
Supplemental Digital Content A. 12 panel figures with caterpillar-type plots depicting scores with bootstrap 95% confidence intervals and ranking variability for departments in the three methods compared, one figure per dimension.

Supplemental Digital Content B. 12 tables illustrating scoring and ranking variability for age classes, occupation, gender, specialty and departments in the three methods compared, one table per dimension.

Supplemental Digital Content C: Unweighted means of M1*, M2† and M3 dimension scores (top) and summary statistics (bottom) at a continental level for 38 international studies on the Hospital Survey On Patient Safety Culture.

Supplemental Digital Content A-Figure 1: "Overall perception of safety" dimension scores for each department.

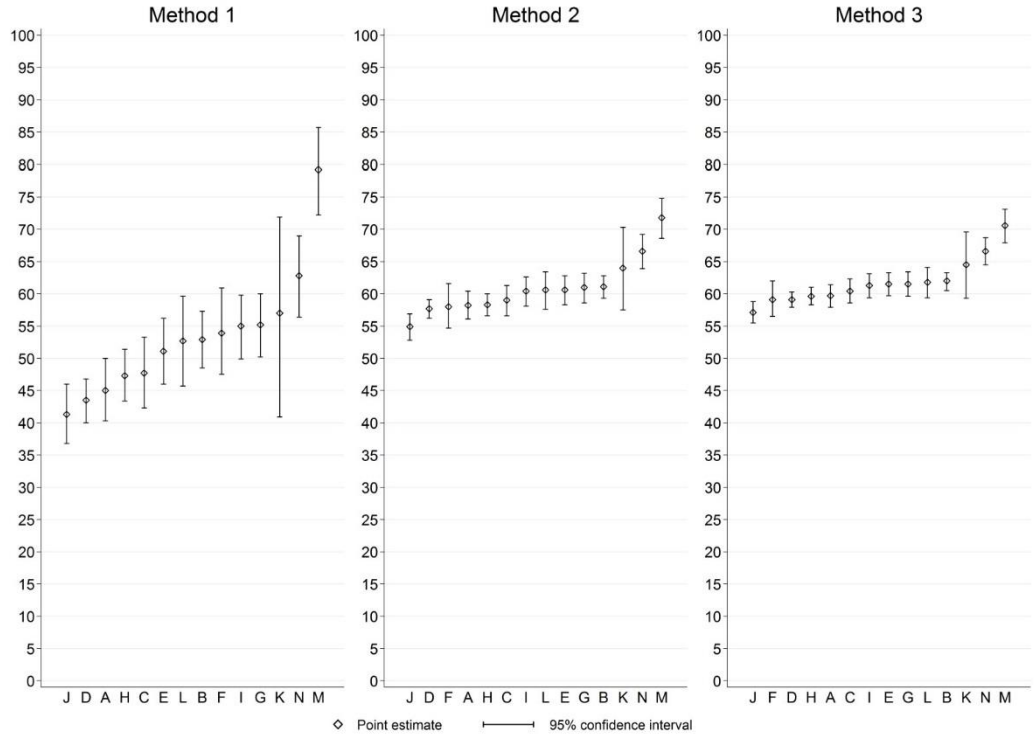
X-axis: ordered de-identified departments for each method. Y-axis: rescaled score.



2

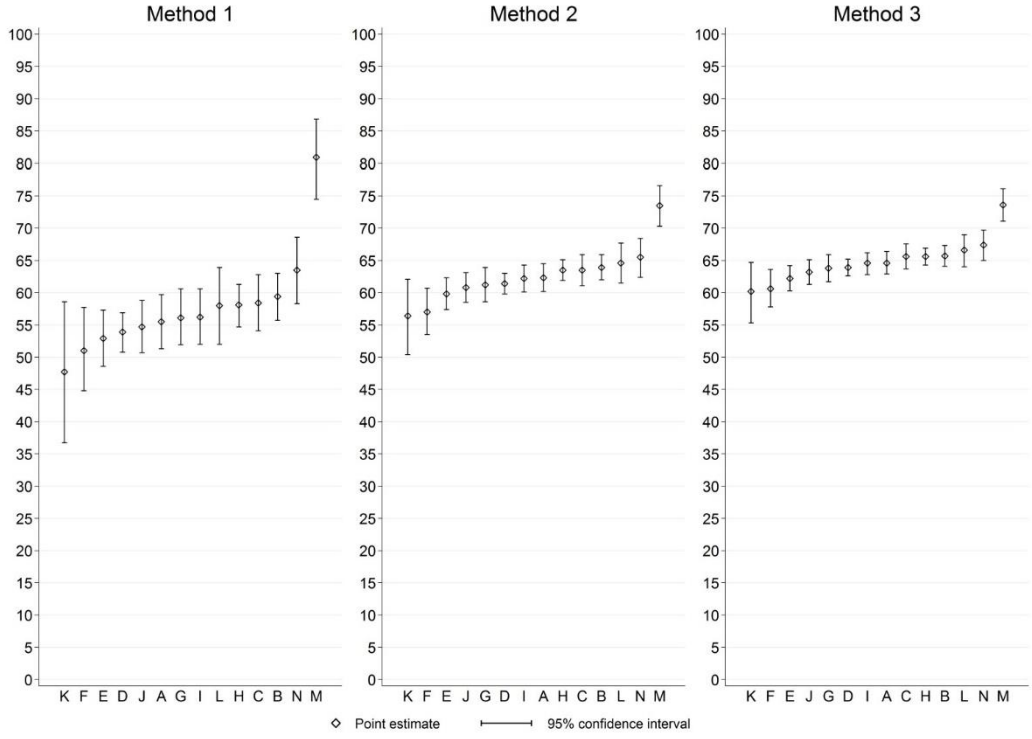
Supplemental Digital Content A-Figure 2: "Frequency of event reporting" dimension scores for each department.

X-axis: ordered de-identified departments for each method. Y-axis: rescaled score.



Supplemental Digital Content A-Figure 3: "Supervisor expectations and actions" dimension scores for each department.

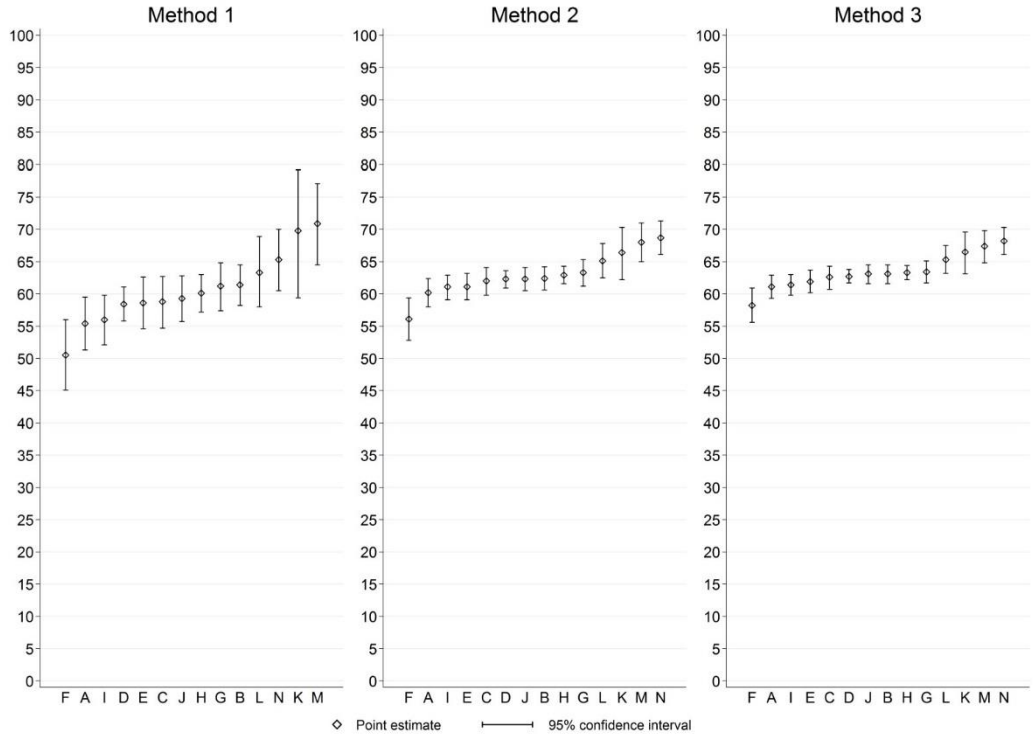
X-axis: ordered de-identified departments for each method. Y-axis: rescaled score.



4

Supplemental Digital Content A-Figure 4: "Communication openness" dimension scores for each department.

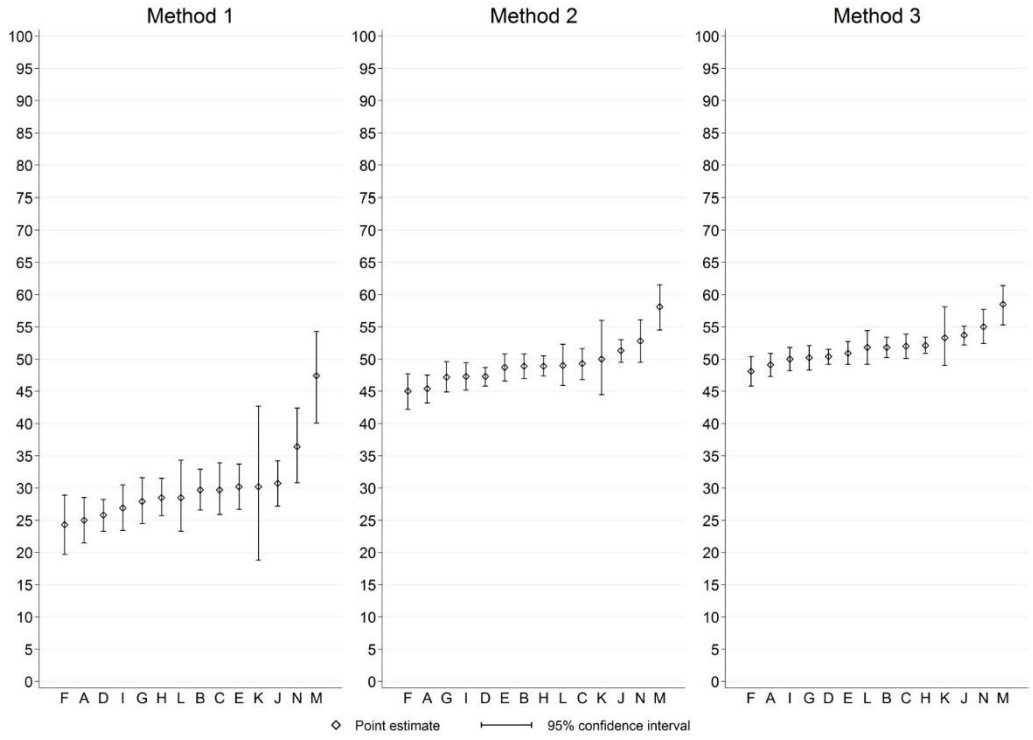
X-axis: ordered de-identified departments for each method. Y-axis: rescaled score.



5

Supplemental Digital Content A-Figure 5: "Non-punitive response to error" dimension scores for each department.

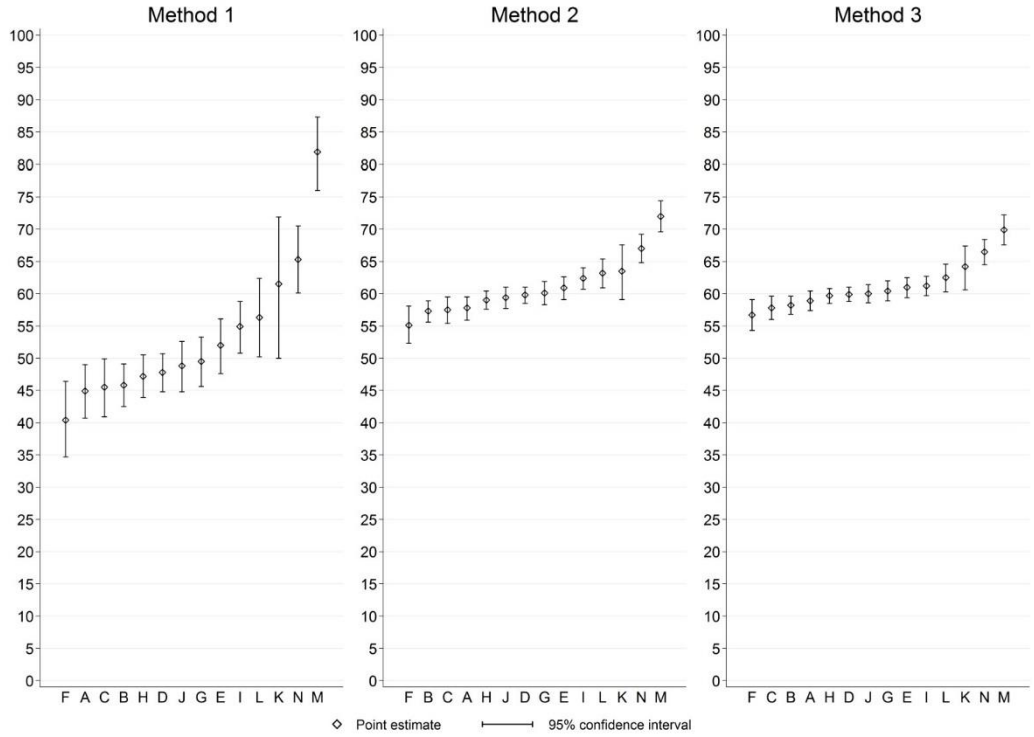
X-axis: ordered de-identified departments for each method. Y-axis: rescaled score.



6

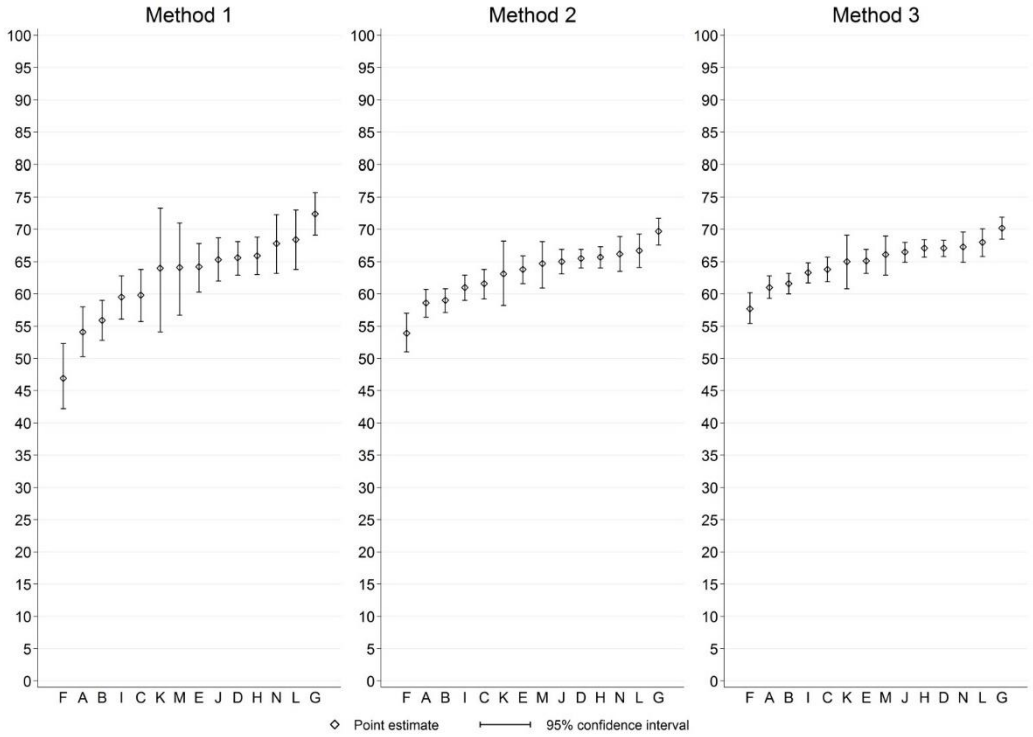
Supplemental Digital Content A-Figure 6: "Organizational learning" dimension scores for each department.

X-axis: ordered de-identified departments for each method. Y-axis: rescaled score.



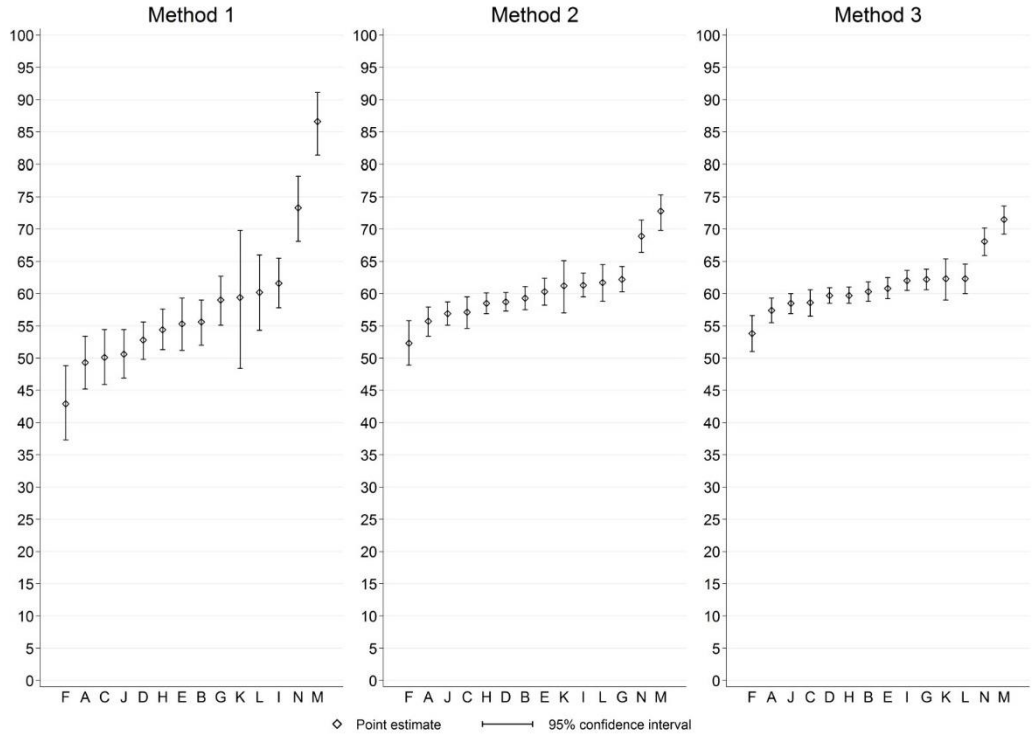
Supplemental Digital Content A-Figure 7: "Teamwork within hospital units" dimension scores for each department.

X-axis: ordered de-identified departments for each method. Y-axis: rescaled score.



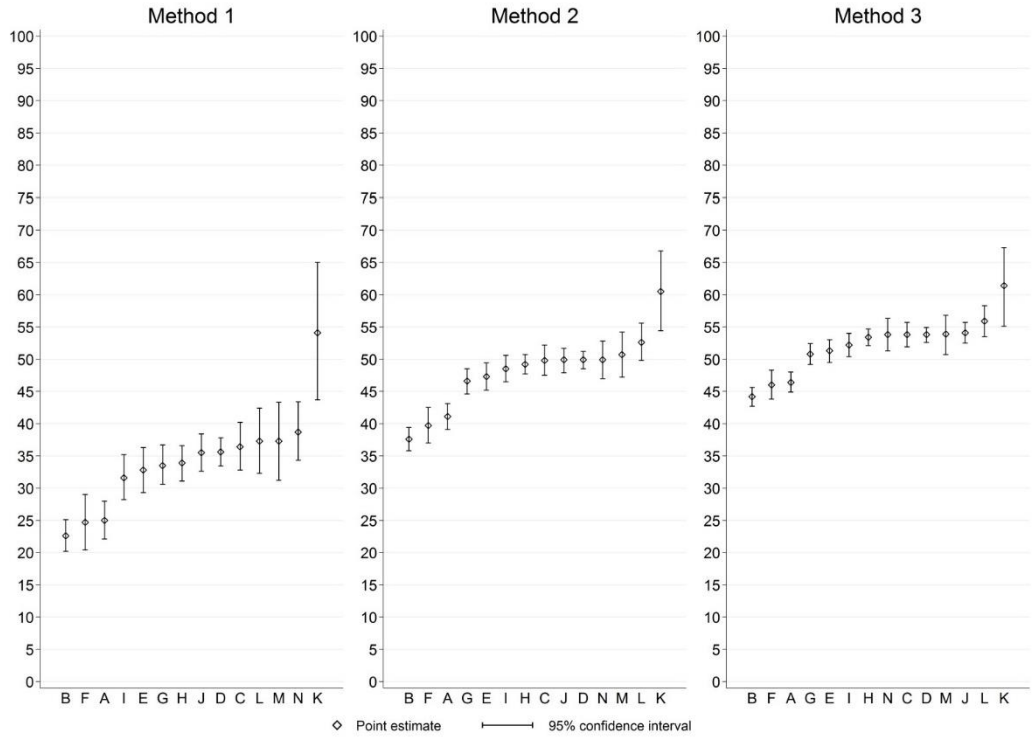
Supplemental Digital Content A-Figure 8: "Feedback and communication about errors" dimension scores for each department.

X-axis: ordered de-identified departments for each method. Y-axis: rescaled score.



Supplemental Digital Content A-Figure 9: "Staffing" dimension scores for each department.

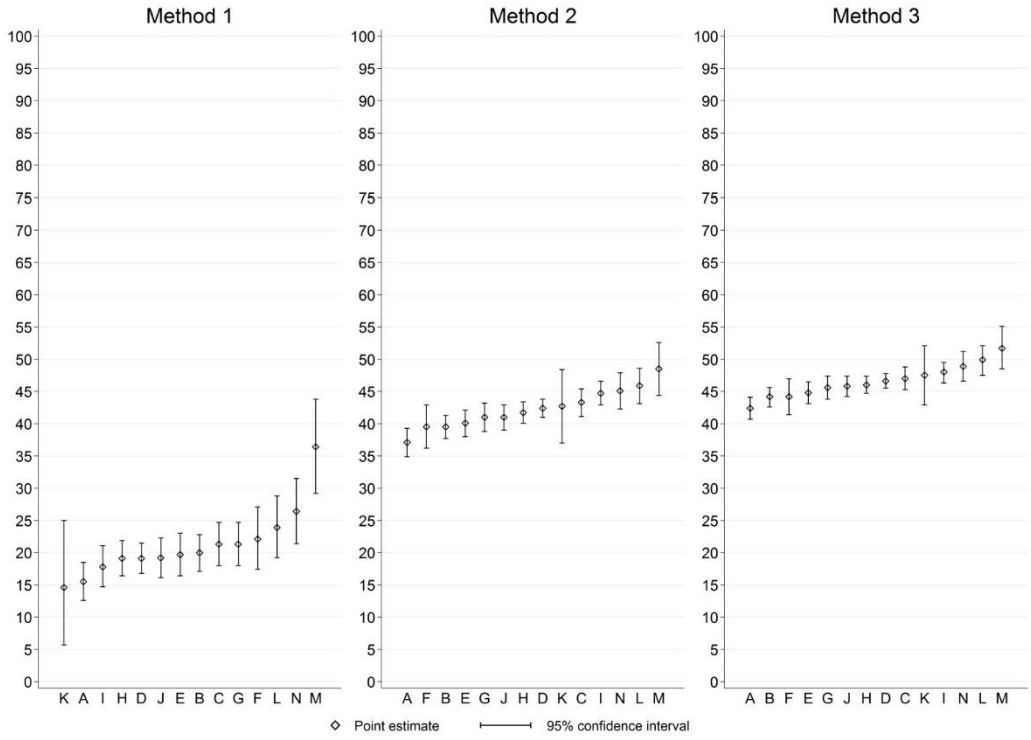
X-axis: ordered de-identified departments for each method. Y-axis: rescaled score.



10

Supplemental Digital Content A-Figure 10: "Hospital management support" dimension scores for each department.

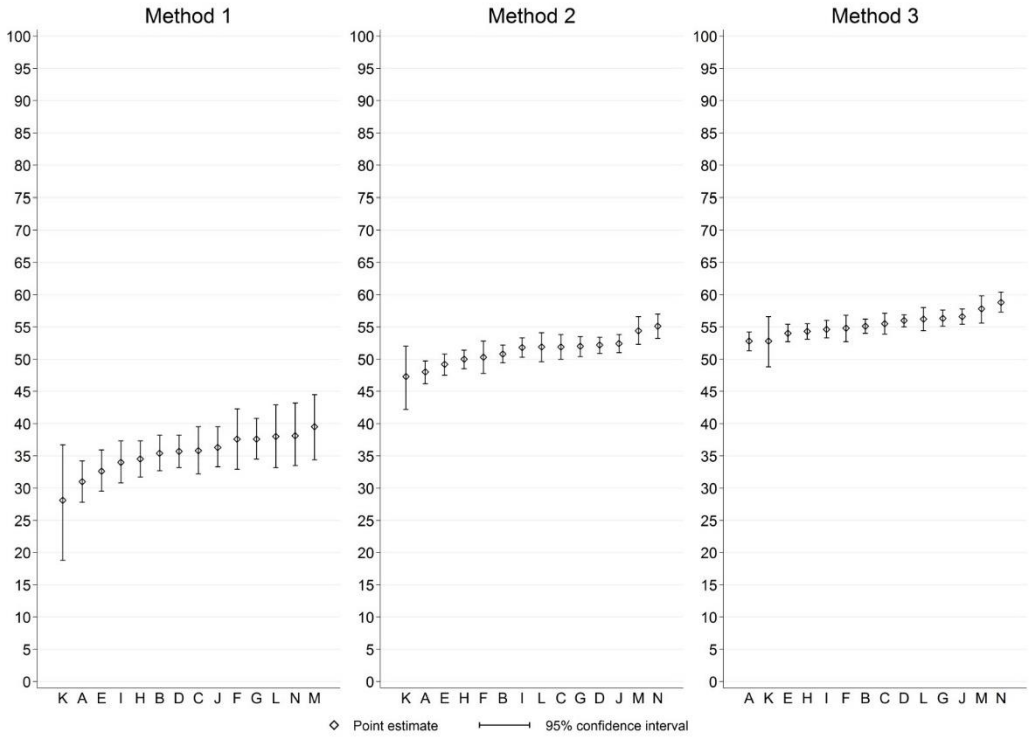
X-axis: ordered de-identified departments for each method. Y-axis: rescaled score.



11

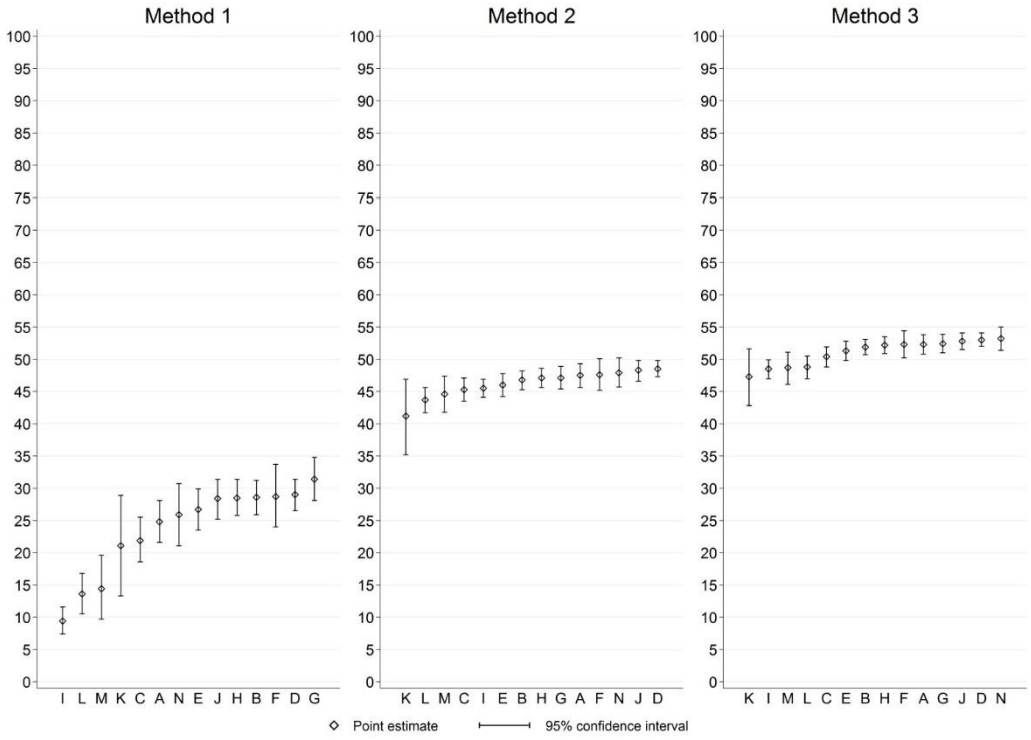
Supplemental Digital Content A-Figure 11: "Teamwork across hospital units" dimension scores for each department.

X-axis: ordered de-identified departments for each method. Y-axis: rescaled score.



Supplemental Digital Content A-Figure 12: "Hospital handoffs and transitions" dimension scores for each department.

X-axis: ordered de-identified departments for each method. Y-axis: rescaled score.



Supplemental Digital Content B-Table 1: "Overall perception of safety" subgroup scores using M1, M2 and M3 with their associated ranks. Ordered according to the AHRQ M1 score, *n*=3888

Variable	*N*	(%)	Scores			Ranks			Sums of absolute rank differences (SARD)		
			M1	M2	M3	M1	M2	M3	$SARD_{1,2}$	$SARD_{1,3}$	$SARD_{2,3}$
Age classes (years)									0	0	0
≥55	415	(11.1)	47.2	58.2	60.1	1	1	1			
45–54	967	(25.8)	43.3	55.7	58.3	2	2	2			
<35	1406	(37.5)	40.9	54.7	58.2	3	3	3			
35–44	966	(25.7)	40.1	54.5	58.0	4	4	4			
Occupation									12	10	4
Other non-healthcare	26	(0.7)	55.8	63.2	64.0	1	1	1			
Physician	457	(12.0)	49.9	58.6	61.0	2	2	2			
Nursing assistant	706	(18.5)	43.4	53.9	57.2	3	7	7			
Nurse	1323	(34.7)	42.9	54.1	57.8	4	6	5			
Technical	775	(20.3)	41.6	56.7	59.1	5	3	3			
Other healthcare	179	(4.7)	35.7	54.4	58.0	6	5	4			
Administrative	347	(9.1)	28.8	54.8	57.5	7	4	6			
Department									14	14	8
N	158	(4.1)	56.3	62.8	64.6	1	1	1			
M	97	(2.5)	54.9	61.4	62.0	2	3	4			
L	150	(3.9)	51.3	60.3	63.0	3	4	3			
K	32	(0.8)	49.2	61.5	64.2	4	2	2			
J	337	(8.7)	48.9	58.2	60.8	5	6	5			
I	303	(7.8)	44.5	58.4	60.0	6	5	7			
H	466	(12.0)	44.1	57.7	60.6	7	7	6			
G	310	(8.0)	43.8	55.9	58.5	8	9	9			
F	153	(3.9)	43.6	53.4	56.8	9	11	12			
E	300	(7.7)	42.9	56.6	59.2	10	8	8			
D	591	(15.2)	38.4	53.3	56.9	11	12	11			
C	266	(6.9)	36.9	54.0	57.6	12	10	10			
B	415	(10.7)	35.4	50.5	54.4	13	13	13			
A	302	(7.8)	29.0	46.3	51.3	14	14	14			
Gender									0	0	0
Male	725	(19.4)	44.3	55.6	59.0	1	1	1			
Female	3016	(80.6)	41.5	55.2	58.2	2	2	2			
Speciality									0	0	0
Biology, imaging, pharmacy	586	(15.1)	48.1	59.5	61.2	1	1	1			
Medicine, psychiatry	1163	(29.9)	42.9	56.3	59.2	2	2	2			
Surgery, gynecology, obstetrics	990	(25.5)	41.0	55.0	58.3	3	3	3			
Anesthesia, reanimation, emergency	771	(19.8)	39.1	52.5	56.3	4	4	4			
Home care, nursing home, follow-up care	378	(9.7)	38.4	51.7	55.5	5	5	5			

Supplemental Digital Content B-Table 2: "Frequency of event reporting" subgroup scores using M1, M2 and M3 with their associated ranks. Ordered according to the AHRQ M1 score, *n*=3888

Variable	*N*	(%)	Scores			Ranks			Sums of absolute rank differences (SARD)		
			M1	M2	M3	M1	M2	M3	$SARD_{1,2}$	$SARD_{1,3}$	$SARD_{2,3}$
Age classes (years)									0	0	0
≥55	415	(11.1)	55.3	62.3	62.6	1	1	1			
45–54	967	(25.8)	52.0	60.5	61.4	2	2	2			
35–44	966	(25.7)	48.8	59.0	60.3	3	3	3			
<35	1406	(37.5)	47.7	58.2	59.8	4	4	4			
Occupation									4	4	2
Other non-healthcare	26	(0.7)	63.9	63.2	63.9	1	1	1			
Nursing assistant	706	(18.5)	57.0	62.6	63.1	2	2	2			
Nurse	1323	(34.7)	52.8	58.9	60.2	3	5	4			
Physician	457	(12.0)	51.7	59.0	60.1	4	4	5			
Technical	775	(20.3)	51.5	61.1	61.9	5	3	3			
Administrative	347	(9.1)	31.1	56.5	58.3	6	6	6			
Other healthcare	179	(4.7)	29.1	52.8	55.2	7	7	7			
Department									20	24	8
M	97	(2.5)	79.2	71.8	70.6	1	1	1			
N	158	(4.1)	62.8	66.6	66.6	2	2	2			
K	32	(0.8)	57.0	64.0	64.5	3	3	3			
G	310	(8.0)	55.2	61.0	61.5	4	5	6			
I	303	(7.8)	55.0	60.4	61.3	5	8	8			
F	153	(3.9)	53.9	58.0	59.1	6	12	13			
B	415	(10.7)	52.9	61.1	62.0	7	4	4			
L	150	(3.9)	52.7	60.6	61.8	8	7	5			
E	300	(7.7)	51.1	60.6	61.5	9	6	7			
C	266	(6.9)	47.7	59.0	60.4	10	9	9			
H	466	(12.0)	47.3	58.3	59.6	11	10	11			
A	302	(7.8)	45.0	58.2	59.7	12	11	10			
D	591	(15.2)	43.6	57.7	59.2	13	13	12			
J	337	(8.7)	41.3	54.9	57.1	14	14	14			
Gender									2	2	0
Male	725	(19.4)	50.5	59.2	60.4	1	2	2			
Female	3016	(80.6)	49.6	59.5	60.7	2	1	1			
Speciality									2	2	0
Biology, imaging, pharmacy	586	(15.1)	57.3	62.0	62.7	1	1	1			
Home care, nursing home, follow-up care	378	(9.7)	55.4	61.4	62.0	2	3	3			
Medicine, psychiatry	1163	(29.9)	52.6	61.5	62.4	3	2	2			
Surgery, gynecology, obstetrics	990	(25.5)	45.7	57.5	58.8	4	4	4			
Anesthesia, reanimation, emergency	771	(19.8)	44.1	56.7	58.5	5	5	5			

Supplemental Digital Content B-Table 3: "Supervisor expectations and actions" subgroup scores using M1, M2 and M3 with their associated ranks. Ordered according to the AHRQ M1 score, *n*=3888

Variable	*N* (%)	Scores			Ranks			Sums of absolute rank differences (SARD)		
		M1	M2	M3	M1	M2	M3	$SARD_{1,2}$	$SARD_{1,3}$	$SARD_{2,3}$
Age classes (years)								0	0	0
45–54	967 (25.8)	58.1	63.2	65.2	1	1	1			
≥55	415 (11.1)	56.5	62.9	65.1	2	2	2			
35–44	966 (25.7)	56.5	62.3	64.7	3	3	3			
<35	1406 (37.5)	56.4	61.9	64.3	4	4	4			
Occupation								8	10	2
Nursing assistant	706 (18.5)	61.2	63.3	65.2	1	3	3			
Technical	775 (20.3)	59.3	64.8	66.5	2	2	2			
Physician	457 (12.0)	57.5	62.6	65.0	3	4	4			
Nurse	1323 (34.7)	57.0	61.2	63.9	4	5	6			
Other non-healthcare	26 (0.7)	57.0	66.3	68.0	5	1	1			
Other healthcare	179 (4.7)	50.6	61.1	63.9	6	6	5			
Administrative	347 (9.1)	44.4	60.0	62.5	7	7	7			
Department								14	14	2
M	97 (2.5)	80.9	73.5	73.6	1	1	1			
N	158 (4.1)	63.5	65.5	67.4	2	2	2			
B	415 (10.7)	59.4	63.9	65.7	3	4	4			
C	266 (6.9)	58.4	63.5	65.6	4	5	6			
H	466 (12.0)	58.1	63.5	65.6	5	6	5			
L	150 (3.9)	58.0	64.6	66.6	6	3	3			
I	303 (7.8)	56.2	62.2	64.6	7	8	8			
G	310 (8.0)	56.1	61.2	63.8	8	10	10			
A	302 (7.8)	55.5	62.3	64.6	9	7	7			
J	337 (8.7)	54.7	60.8	63.2	10	11	11			
D	591 (15.2)	53.9	61.4	63.9	11	9	9			
E	300 (7.7)	52.9	59.8	62.2	12	12	12			
F	153 (3.9)	51.0	57.0	60.6	13	13	13			
K	32 (0.8)	47.7	56.4	60.2	14	14	14			
Gender								0	0	0
Male	725 (19.4)	60.1	64.1	66.2	1	1	1			
Female	3016 (80.6)	56.2	62.1	64.4	2	2	2			
Speciality								0	0	0
Home care, nursing home, follow-up care	378 (9.7)	60.6	64.7	66.5	1	1	1			
Medicine, psychiatry	1163 (29.9)	60.6	64.5	66.3	2	2	2			
Biology, imaging, pharmacy	586 (15.1)	60.1	64.4	66.3	3	3	3			
Anesthesia, reanimation, emergency	771 (19.8)	53.1	60.1	62.9	4	4	4			
Surgery, gynecology, obstetrics	990 (25.5)	51.7	59.5	62.4	5	5	5			

Supplemental Digital Content B-Table 4: "Communication openness" subgroup scores using M1, M2 and M3 with their associated ranks. Ordered according to the AHRQ M1 score, *n*=3888

Variable	*N* (%)	Scores			Ranks			Sums of absolute rank differences (SARD)		
		M1	M2	M3	M1	M2	M3	$SARD_{1,2}$	$SARD_{1,3}$	$SARD_{2,3}$
Age classes (years)								0	0	0
≥55	415 (11.1)	61.9	63.9	64.1	1	1	1			
45–54	967 (25.8)	61.0	62.9	63.2	2	2	2			
35–44	966 (25.7)	58.8	61.9	62.6	3	3	3			
<35	1406 (37.5)	58.0	61.7	62.4	4	4	4			
Occupation								2	2	0
Other non-healthcare	26 (0.7)	78.2	74.4	72.8	1	1	1			
Physician	457 (12.0)	66.6	66.5	66.0	2	2	2			
Nurse	1323 (34.7)	64.2	63.6	64.1	3	3	3			
Nursing assistant	706 (18.5)	61.5	62.5	63.0	4	5	5			
Other healthcare	179 (4.7)	58.4	63.2	63.6	5	4	4			
Technical	775 (20.3)	53.0	59.7	60.5	6	6	6			
Administrative	347 (9.1)	41.3	57.1	58.6	7	7	7			
Department								12	12	0
M	97 (2.5)	70.9	68.0	67.4	1	2	2			
K	32 (0.8)	69.8	66.4	66.5	2	3	3			
N	158 (4.1)	65.3	68.7	68.2	3	1	1			
L	150 (3.9)	63.3	65.1	65.3	4	4	4			
B	415 (10.7)	61.4	62.4	63.1	5	7	7			
G	310 (8.0)	61.2	63.3	63.4	6	5	5			
H	466 (12.0)	60.1	62.9	63.3	7	6	6			
J	337 (8.7)	59.3	62.3	63.1	8	8	8			
C	266 (6.9)	58.8	62.0	62.6	9	10	10			
E	300 (7.7)	58.6	61.1	61.9	10	11	11			
D	591 (15.2)	58.5	62.3	62.8	11	9	9			
I	303 (7.8)	56.0	61.1	61.4	12	12	12			
A	302 (7.8)	55.4	60.2	61.1	13	13	13			
F	153 (3.9)	50.5	56.1	58.2	14	14	14			
Gender								2	2	0
Male	725 (19.4)	59.6	62.2	62.8	1	2	2			
Female	3016 (80.6)	59.3	62.3	62.9	2	1	1			
Speciality								2	2	0
Home care, nursing home, follow-up care	378 (9.7)	62.7	63.5	63.9	1	2	2			
Medicine, psychiatry	1163 (29.9)	62.2	64.4	64.5	2	1	1			
Biology, imaging, pharmacy	586 (15.1)	60.3	63.2	63.4	3	3	3			
Anesthesia, reanimation, emergency	771 (19.8)	57.4	61.1	61.9	4	4	4			
Surgery, gynecology, obstetrics	990 (25.5)	55.6	60.1	61.0	5	5	5			

Supplemental Digital Content B-Table 5: "Nonpunitive response to error" subgroup scores using M1, M2 and M3 with their associated ranks. Ordered according to the AHRQ M1 score, *n*=3888

Variable	*N* (%)	Scores			Ranks			Sums of absolute rank differences (SARD)		
		M1	M2	M3	M1	M2	M3	$SARD_{1,2}$	$SARD_{1,3}$	$SARD_{2,3}$
Age classes (years)								0	2	2
<35	1406 (37.5)	30.3	50.0	52.9	1	1	1			
≥55	415 (11.1)	29.4	48.3	51.0	2	2	2			
45–54	967 (25.8)	27.7	47.7	50.2	3	3	4			
35–44	966 (25.7)	27.4	47.7	50.9	4	4	3			
Occupation								4	4	0
Physician	457 (12.0)	42.0	57.1	58.2	1	1	1			
Other non-healthcare	26 (0.7)	38.5	52.6	55.4	2	2	2			
Nurse	1323 (34.7)	33.3	50.3	53.0	3	4	4			
Other healthcare	179 (4.7)	27.8	50.9	53.8	4	3	3			
Nursing assistant	706 (18.5)	24.3	45.1	48.5	5	6	6			
Technical	775 (20.3)	23.4	45.6	48.7	6	5	5			
Administrative	347 (9.1)	15.1	43.5	47.5	7	7	7			
Department								14	12	8
M	97 (2.5)	47.4	58.1	58.5	1	1	1			
N	158 (4.1)	36.4	52.8	55.0	2	2	2			
J	337 (8.7)	30.7	51.3	53.7	3	3	3			
K	32 (0.8)	30.2	50.0	53.3	4	4	4			
E	300 (7.7)	30.2	48.7	50.9	5	9	9			
C	266 (6.9)	29.7	49.3	52.0	6	5	6			
B	415 (10.7)	29.7	48.9	51.8	7	8	7			
L	150 (3.9)	28.5	49.0	51.8	8	6	8			
H	466 (12.0)	28.5	48.9	52.1	9	7	5			
G	310 (8.0)	27.9	47.2	50.2	10	12	11			
I	303 (7.8)	26.9	47.3	50.0	11	11	12			
D	591 (15.2)	25.7	47.3	50.4	12	10	10			
A	302 (7.8)	25.0	45.4	49.1	13	13	13			
F	153 (3.9)	24.3	45.0	48.1	14	14	14			
Gender								0	0	0
Male	725 (19.4)	30.6	49.5	52.5	1	1	1			
Female	3016 (80.6)	28.4	48.4	51.2	2	2	2			
Speciality								4	0	4
Home care, nursing home, follow-up care	378 (9.7)	31.4	49.8	52.6	1	3	1			
Medicine, psychiatry	1163 (29.9)	31.2	50.0	52.4	2	1	2			
Biology, imaging, pharmacy	586 (15.1)	30.9	49.8	52.2	3	2	3			
Anesthesia, reanimation, emergency	771 (19.8)	27.5	47.7	50.8	4	4	4			
Surgery, gynecology, obstetrics	990 (25.5)	24.7	46.4	49.8	5	5	5			

Supplemental Digital Content B-Table 6: "Organizational learning" subgroup scores using M1, M2 and M3 with their associated ranks. Ordered according to the AHRQ M1 score, *n*=3888

Variable	*N* (%)	Scores			Ranks			Sums of absolute rank differences (SARD)		
		M1	M2	M3	M1	M2	M3	$SARD_{1,2}$	$SARD_{1,3}$	$SARD_{2,3}$
Age classes (years)								0	0	0
≥55	415 (11.1)	59.7	63.9	63.0	1	1	1			
45–54	967 (25.8)	54.5	61.9	61.4	2	2	2			
35–44	966 (25.7)	48.1	58.9	59.5	3	3	3			
<35	1406 (37.5)	45.0	58.1	59.3	4	4	4			
Occupation								4	2	2
Physician	457 (12.0)	65.1	65.5	65.3	1	1	1			
Other non-healthcare	26 (0.7)	53.8	61.2	62.3	2	2	2			
Nursing assistant	706 (18.5)	52.2	60.3	60.5	3	3	3			
Nurse	1323 (34.7)	49.9	59.1	59.9	4	5	4			
Technical	775 (20.3)	47.5	59.8	59.7	5	4	5			
Other healthcare	179 (4.7)	38.1	55.6	57.3	6	7	7			
Administrative	347 (9.1)	37.5	58.3	58.5	7	6	6			
Department								6	4	4
M	97 (2.5)	81.9	72.0	69.9	1	1	1			
N	158 (4.1)	65.3	67.0	66.5	2	2	2			
K	32 (0.8)	61.5	63.5	64.2	3	3	3			
L	150 (3.9)	56.3	63.2	62.5	4	4	4			
I	303 (7.8)	54.9	62.4	61.2	5	5	5			
E	300 (7.7)	52.0	60.9	61.0	6	6	6			
G	310 (8.0)	49.5	60.1	60.4	7	7	7			
J	337 (8.7)	48.8	59.4	60.0	8	9	8			
D	591 (15.2)	47.8	59.8	59.9	9	8	9			
H	466 (12.0)	47.2	59.0	59.7	10	10	10			
B	415 (10.7)	45.8	57.3	58.2	11	13	12			
C	266 (6.9)	45.5	57.5	57.8	12	12	13			
A	302 (7.8)	44.9	57.8	58.9	13	11	11			
F	153 (3.9)	40.4	55.1	56.7	14	14	14			
Gender								0	0	0
Male	725 (19.4)	52.5	60.7	61.5	1	1	1			
Female	3016 (80.6)	49.1	59.7	59.9	2	2	2			
Speciality								0	2	2
Biology, imaging, pharmacy	586 (15.1)	59.6	64.1	63.0	1	1	1			
Medicine, psychiatry	1163 (29.9)	52.1	61.0	61.2	2	2	2			
Home care, nursing home, follow-up care	378 (9.7)	48.9	58.7	58.9	3	3	4			
Anesthesia, reanimation, emergency	771 (19.8)	47.3	58.4	59.3	4	4	3			
Surgery, gynecology, obstetrics	990 (25.5)	44.1	58.0	58.8	5	5	5			

Supplemental Digital Content B-Table 7: "Teamwork within hospital units" subgroup scores using M1, M2 and M3 with their associated ranks. Ordered according to the AHRQ M1 score, *n*=3888

Variable	*N* (%)	Scores			Ranks			Sums of absolute rank differences (SARD)		
		M1	M2	M3	M1	M2	M3	$SARD_{1,2}$	$SARD_{1,3}$	$SARD_{2,3}$
Age classes (years)								0	0	0
<35	1406 (37.5)	66.1	65.6	67.2	1	1	1			
≥55	415 (11.1)	62.7	64.0	65.6	2	2	2			
45–54	967 (25.8)	61.3	62.5	64.2	3	3	3			
35–44	966 (25.7)	59.2	61.2	63.3	4	4	4			
Occupation								0	2	2
Nurse	1323 (34.7)	68.2	66.8	68.0	1	1	1			
Physician	457 (12.0)	65.7	65.3	66.9	2	2	2			
Other non-healthcare	26 (0.7)	64.1	65.2	66.3	3	3	4			
Other healthcare	179 (4.7)	63.5	65.0	66.5	4	4	3			
Nursing assistant	706 (18.5)	61.1	62.1	64.1	5	5	5			
Administrative	347 (9.1)	57.4	61.1	62.9	6	6	6			
Technical	775 (20.3)	56.6	59.3	61.7	7	7	7			
Department								2	4	2
G	310 (8.0)	72.4	69.7	70.2	1	1	1			
L	150 (3.9)	68.4	66.7	68.0	2	2	2			
N	158 (4.1)	67.8	66.2	67.3	3	3	3			
H	466 (12.0)	65.9	65.7	67.1	4	4	5			
D	591 (15.2)	65.6	65.5	67.1	5	5	4			
J	337 (8.7)	65.3	65.0	66.5	6	6	6			
E	300 (7.7)	64.2	63.8	65.1	7	8	8			
M	97 (2.5)	64.1	64.7	66.1	8	7	7			
K	32 (0.8)	64.0	63.1	65.0	9	9	9			
C	266 (6.9)	59.8	61.6	63.8	10	10	10			
I	303 (7.8)	59.5	61.0	63.3	11	11	11			
B	415 (10.7)	55.9	59.0	61.6	12	12	12			
A	302 (7.8)	54.1	58.6	61.0	13	13	13			
F	153 (3.9)	46.9	53.9	57.7	14	14	14			
Gender								0	0	0
Female	3016 (80.6)	63.6	63.9	65.6	1	1	1			
Male	725 (19.4)	59.5	61.6	63.8	2	2	2			
Speciality								0	0	0
Medicine, psychiatry	1163 (29.9)	65.7	65.6	67.0	1	1	1			
Biology, imaging, pharmacy	586 (15.1)	62.5	63.1	65.0	2	2	2			
Surgery, gynecology, obstetrics	990 (25.5)	62.4	62.9	64.8	3	3	3			
Anesthesia, reanimation, emergency	771 (19.8)	61.3	62.7	64.6	4	4	4			
Home care, nursing home, follow-up care	378 (9.7)	57.2	60.1	62.1	5	5	5			

Supplemental Digital Content B-Table 8: "Feedback and communication about errors" subgroup scores using M1, M2 and M3 with their associated ranks. Ordered according to the AHRQ M1 score, *n*=3888

Variable	*N* (%)	Scores			Ranks			Sums of absolute rank differences (SARD)		
		M1	M2	M3	M1	M2	M3	$SARD_{1,2}$	$SARD_{1,3}$	$SARD_{2,3}$
Age classes (years)								0	0	0
≥55	415 (11.1)	63.0	63.1	63.2	1	1	1			
45–54	967 (25.8)	58.2	60.8	61.4	2	2	2			
35–44	966 (25.7)	55.6	59.4	60.6	3	3	3			
<35	1406 (37.5)	51.9	57.6	58.8	4	4	4			
Occupation								4	6	2
Physician	457 (12.0)	63.1	62.4	62.6	1	1	2			
Technical	775 (20.3)	58.8	61.2	61.9	2	3	3			
Nursing assistant	706 (18.5)	58.5	61.0	61.6	3	4	4			
Other non-healthcare	26 (0.7)	55.1	62.2	63.1	4	2	1			
Nurse	1323 (34.7)	54.7	58.2	59.4	5	5	5			
Administrative	347 (9.1)	44.0	56.3	57.6	6	6	6			
Other healthcare	179 (4.7)	40.5	53.5	56.0	7	7	7			
Department								12	10	8
M	97 (2.5)	86.6	72.8	71.5	1	1	1			
N	158 (4.1)	73.3	68.9	68.1	2	2	2			
I	303 (7.8)	61.6	61.3	62.0	3	5	6			
L	150 (3.9)	60.2	61.7	62.3	4	4	3			
K	32 (0.8)	59.4	61.2	62.3	5	6	4			
G	310 (8.0)	59.0	62.2	62.2	6	3	5			
B	415 (10.7)	55.6	59.3	60.3	7	8	8			
E	300 (7.7)	55.3	60.3	60.8	8	7	7			
H	466 (12.0)	54.4	58.5	59.7	9	10	9			
D	591 (15.2)	52.8	58.7	59.7	10	9	10			
J	337 (8.7)	50.6	56.9	58.5	11	12	12			
C	266 (6.9)	50.1	57.1	58.6	12	11	11			
A	302 (7.8)	49.3	55.7	57.4	13	13	13			
F	153 (3.9)	42.9	52.3	53.8	14	14	14			
Gender								0	0	0
Male	725 (19.4)	57.2	59.5	60.8	1	1	1			
Female	3016 (80.6)	55.3	59.4	60.3	2	2	2			
Speciality								0	0	0
Biology, imaging, pharmacy	586 (15.1)	64.6	62.9	63.4	1	1	1			
Medicine, psychiatry	1163 (29.9)	58.8	62.2	62.5	2	2	2			
Home care, nursing home, follow-up care	378 (9.7)	56.7	59.8	60.8	3	3	3			
Anesthesia, reanimation, emergency	771 (19.8)	51.2	56.5	58.1	4	4	4			
Surgery, gynecology, obstetrics	990 (25.5)	49.5	56.4	57.8	5	5	5			

Supplemental Digital Content B-Table 9: "Staffing" subgroup scores using M1, M2 and M3 with their associated ranks. Ordered according to the AHRQ M1 score, *n*=3888

Variable	*N*	(%)	Scores			Ranks			Sums of absolute rank differences (SARD)		
			M1	M2	M3	M1	M2	M3	$SARD_{1,2}$	$SARD_{1,3}$	$SARD_{2,3}$
Age classes (years)									0	0	0
<35	1406	(37.5)	34.1	48.2	52.9	1	1	1			
≥55	415	(11.1)	33.0	47.4	51.3	2	2	2			
35–44	966	(25.7)	31.9	46.6	51.1	3	3	3			
45–54	967	(25.8)	30.2	45.5	50.0	4	4	4			
Occupation									6	4	2
Nurse	1323	(34.7)	38.6	49.5	53.5	1	3	3			
Other healthcare	179	(4.7)	37.6	51.9	55.6	2	2	2			
Other non-healthcare	26	(0.7)	36.4	53.4	55.8	3	1	1			
Nursing assistant	706	(18.5)	34.2	45.8	50.5	4	5	4			
Technical	775	(20.3)	28.3	46.1	50.4	5	4	5			
Physicist	457	(12.0)	28.2	44.7	50.1	6	6	6			
Administrative	347	(9.1)	17.0	41.8	47.2	7	7	7			
Department									12	16	10
K	32	(0.8)	54.1	60.5	61.4	1	1	1			
N	158	(4.1)	38.7	49.9	53.8	2	4	6			
M	97	(2.5)	37.3	50.7	53.9	3	3	4			
L	150	(3.9)	37.3	52.6	55.9	4	2	2			
C	266	(6.9)	36.4	49.8	53.8	5	7	5			
D	591	(15.2)	35.6	49.9	53.8	6	5	7			
J	337	(8.7)	35.5	49.9	54.1	7	6	3			
H	466	(12.0)	33.9	49.2	53.4	8	8	8			
G	310	(8.0)	33.5	46.6	50.8	9	11	11			
E	300	(7.7)	32.8	47.3	51.3	10	10	10			
I	303	(7.8)	31.6	48.5	52.2	11	9	9			
A	302	(7.8)	25.0	41.1	46.4	12	12	12			
F	153	(3.9)	24.7	39.7	46.0	13	13	13			
B	415	(10.7)	22.6	37.6	44.2	14	14	14			
Gender									0	0	0
Female	3016	(80.6)	33.1	47.2	51.6	1	1	1			
Male	725	(19.4)	30.0	46.5	51.5	2	2	2			
Speciality									2	2	0
Medicine, psychiatry	1163	(29.9)	34.1	47.7	51.9	1	2	2			
Biology, imaging, pharmacy	586	(15.1)	33.3	49.6	53.2	2	1	1			
Anesthesia, reanimation, emergency	771	(19.8)	32.8	47.2	51.8	3	3	3			
Surgery, gynecology, obstetrics	990	(25.5)	32.5	47.2	51.8	4	4	4			
Home care, nursing home, follow-up care	378	(9.7)	25.2	39.8	45.8	5	5	5			

Supplemental Digital Content B-Table 10: "Hospital management support" subgroup scores using M1, M2 and M3 with their associated ranks. Ordered according to the AHRQ M1 score, *n*=3888

Variable	*N* (%)	Scores			Ranks			Sums of absolute rank differences (SARD)		
		M1	M2	M3	M1	M2	M3	$SARD_{1,2}$	$SARD_{1,3}$	$SARD_{2,3}$
Age classes (years)								0	0	0
≥55	415 (11.1)	28.0	46.3	49.6	1	1	1			
45–54	967 (25.8)	24.3	44.3	48.1	2	2	2			
35–44	966 (25.7)	19.6	41.3	45.6	3	3	3			
<35	1406 (37.5)	15.9	39.3	44.5	4	4	4			
Occupation								12	14	4
Technical	775 (20.3)	25.1	46.4	49.3	1	2	3			
Nursing assistant	706 (18.5)	22.9	40.8	45.4	2	5	6			
Administrative	347 (9.1)	22.6	48.2	51.1	3	1	1			
Other non-healthcare	26 (0.7)	21.8	46.2	49.3	4	3	2			
Physician	457 (12.0)	18.4	40.8	45.6	5	6	5			
Nurse	1323 (34.7)	16.3	37.8	43.2	6	7	7			
Other healthcare	179 (4.7)	15.8	43.5	47.7	7	4	4			
Department								48	48	4
M	97 (2.5)	36.4	48.5	51.7	1	1	1			
N	158 (4.1)	26.4	45.1	48.9	2	3	3			
L	150 (3.9)	23.9	45.9	49.9	3	2	2			
F	153 (3.9)	22.1	39.5	44.2	4	13	12			
G	310 (8.0)	21.3	41.0	45.6	5	10	10			
C	266 (6.9)	21.3	43.3	47.0	6	5	6			
B	415 (10.7)	20.0	39.5	44.2	7	12	13			
E	300 (7.7)	19.7	40.1	44.8	8	11	11			
J	337 (8.7)	19.2	41.0	45.8	9	9	9			
H	466 (12.0)	19.1	41.7	46.0	10	8	8			
D	591 (15.2)	19.1	42.4	46.6	11	7	7			
I	303 (7.8)	17.8	44.7	48.0	12	4	4			
A	302 (7.8)	15.5	37.1	42.4	13	14	14			
K	32 (0.8)	14.6	42.7	47.5	14	6	5			
Gender								0	0	0
Male	725 (19.4)	22.9	42.4	46.7	1	1	1			
Female	3016 (80.6)	19.7	41.7	46.1	2	2	2			
Speciality								2	2	0
Home care, nursing home, follow-up care	378 (9.7)	22.9	41.6	46.0	1	2	2			
Biology, imaging, pharmacy	586 (15.1)	22.7	45.5	49.0	2	1	1			
Medicine, psychiatry	1163 (29.9)	20.2	41.5	45.8	3	3	3			
Surgery, gynecology, obstetrics	990 (25.5)	19.6	41.1	45.6	4	4	4			
Anesthesia, reanimation, emergency	771 (19.8)	17.8	39.9	44.9	5	5	5			

Supplemental Digital Content B-Table 11: "Teamwork across hospital units" subgroup scores using M1, M2 and M3 with their associated ranks. Ordered according to the AHRQ M1 score, n=3888

Variable	*N* (%)	Scores			Ranks			Sums of absolute rank differences (SARD)		
		M1	M2	M3	M1	M2	M3	$SARD_{1,2}$	$SARD_{1,3}$	$SARD_{2,3}$
Age classes (years)								2	2	0
≥55	415 (11.1)	37.1	53.0	55.9	1	1	1			
45–54	967 (25.8)	37.1	51.9	55.9	2	2	2			
<35	1406 (37.5)	34.4	50.6	55.1	3	4	4			
35–44	966 (25.7)	34.3	50.8	55.1	4	3	3			
Occupation								10	10	0
Other non-healthcare	26 (0.7)	43.8	55.8	59.0	1	1	1			
Physician	457 (12.0)	42.9	53.5	57.1	2	3	3			
Other healthcare	179 (4.7)	37.1	53.7	57.2	3	2	2			
Nursing assistant	706 (18.5)	34.8	50.2	54.6	4	6	6			
Nurse	1323 (34.7)	34.2	49.9	54.6	5	7	7			
Technical	775 (20.3)	34.1	51.8	55.3	6	5	5			
Administrative	347 (9.1)	33.3	52.9	56.6	7	4	4			
Department								24	18	12
M	97 (2.5)	39.5	54.4	57.8	1	2	2			
N	158 (4.1)	38.1	55.1	58.8	2	1	1			
L	150 (3.9)	38.0	51.9	56.2	3	7	5			
G	310 (8.0)	37.6	52.0	56.3	4	5	4			
F	153 (3.9)	37.6	50.3	54.8	5	10	9			
J	337 (8.7)	36.3	52.4	56.6	6	3	3			
C	266 (6.9)	35.8	51.9	55.5	7	6	7			
D	591 (15.2)	35.7	52.2	55.9	8	4	6			
B	415 (10.7)	35.4	50.8	55.1	9	9	8			
H	466 (12.0)	34.5	50.0	54.3	10	11	11			
I	303 (7.8)	34.0	51.8	54.6	11	8	10			
E	300 (7.7)	32.6	49.2	54.0	12	12	12			
A	302 (7.8)	31.0	48.0	52.8	13	13	14			
K	32 (0.8)	28.1	47.3	52.8	14	14	13			
Gender								0	0	0
Male	725 (19.4)	38.5	52.1	56.1	1	1	1			
Female	3016 (80.6)	34.5	51.0	55.2	2	2	2			
Speciality								2	4	2
Home care, nursing home, follow-up care	378 (9.7)	38.1	53.0	57.1	1	1	1			
Biology, imaging, pharmacy	586 (15.1)	36.0	52.4	55.6	2	2	2			
Surgery, gynecology, obstetrics	990 (25.5)	35.0	50.5	54.8	3	4	5			
Medicine, psychiatry	1163 (29.9)	35.0	51.3	55.4	4	3	3			
Anesthesia, reanimation, emergency	771 (19.8)	34.1	50.3	54.8	5	5	4			

Supplemental Digital Content B-Table 12: "Hospital handoffs and transitions" subgroup scores using M1, M2 and M3 with their associated ranks. Ordered according to the AHRQ M1 score, *n*=3888

Variable	*N* (%)	Scores			Ranks			Sums of absolute rank differences (SARD)		
		M1	M2	M3	M1	M2	M3	$SARD_{1,2}$	$SARD_{1,3}$	$SARD_{2,3}$
Age classes (years)								2	2	4
45–54	967 (25.8)	27.0	47.7	52.1	1	1	2			
<35	1406 (37.5)	26.3	47.1	52.2	2	2	1			
35–44	966 (25.7)	24.0	46.1	51.2	3	4	3			
≥55	415 (11.1)	22.2	46.6	50.8	4	3	4			
Occupation								14	12	4
Nursing assistant	706 (18.5)	34.7	48.9	53.6	1	1	1			
Nurse	1323 (34.7)	29.0	46.6	51.7	2	3	3			
Other non-healthcare	26 (0.7)	24.1	45.1	48.2	3	6	7			
Other healthcare	179 (4.7)	23.7	45.7	51.3	4	5	4			
Physician	457 (12.0)	23.2	45.0	49.9	5	7	6			
Technical	775 (20.3)	19.2	47.6	51.8	6	2	2			
Administrative	347 (9.1)	11.4	46.2	51.2	7	4	5			
Department								36	34	14
G	310 (8.0)	31.4	47.1	52.4	1	6	4			
D	591 (15.2)	29.0	48.5	53.0	2	1	2			
F	153 (3.9)	28.7	47.6	52.3	3	4	6			
B	415 (10.7)	28.6	46.8	51.9	4	8	8			
H	466 (12.0)	28.5	47.1	52.2	5	7	7			
J	337 (8.7)	28.4	48.3	52.8	6	2	3			
E	300 (7.7)	26.7	46.0	51.3	7	9	9			
N	158 (4.1)	25.9	47.9	53.2	8	3	1			
A	302 (7.8)	24.8	47.5	52.3	9	5	5			
C	266 (6.9)	21.9	45.3	50.4	10	11	10			
K	32 (0.8)	21.1	41.2	47.3	11	14	14			
M	97 (2.5)	14.4	44.6	48.7	12	12	12			
L	150 (3.9)	13.6	43.7	48.8	13	13	11			
I	303 (7.8)	9.4	45.5	48.5	14	10	13			
Gender								2	2	0
Female	3016 (80.6)	25.6	46.9	51.7	1	2	2			
Male	725 (19.4)	24.6	47.2	52.1	2	1	1			
Speciality								2	0	2
Home care, nursing home, follow-up care	378 (9.7)	32.4	48.5	53.5	1	2	1			
Anesthesia, reanimation, emergency	771 (19.8)	28.8	48.8	53.3	2	1	2			
Surgery, gynecology, obstetrics	990 (25.5)	28.6	47.1	52.1	3	3	3			
Medicine, psychiatry	1163 (29.9)	24.9	45.8	51.1	4	4	4			
Biology, imaging, pharmacy	586 (15.1)	11.5	45.2	49.0	5	5	5			

Supplemental Digital Content C: Unweighted means of M1*, M2† and M3 dimension scores (top) and summary statistics (bottom) at a continental level for 38 international studies on the Hospital Survey On Patient Safety Culture. # [1-38]

	US (*n*=10)	Europe (*n*=15)		Asia (*n*=6)		Middle East (*n*=7)		All studies (*n*=38)		
Dimension scores, mean ± standard deviation	M1* (*n*=10)	M1* (*n*=10)	M2† (i=8)	M1* (*n*=4)	M3‡ (*n*=2)	M1* (*n*=7)	M2† (*n*=2)	M1* (*n*=31)	M2† (*n*=10)	M3‡ (*n*=2)
Overall perception of safety	56.0 ±12.7	58.4 ±7.0	61.7 ±4.5	57.3 ±6.8	64.0 ±0.0	53.8 ±13.7	65.0 ±6.0	56.4 ±10.5	62.4 ±4.6	64.0 ±0.0
Frequency of event reporting	47.9 ±12.1	42.7 ±21.5	56.5 ±12.3	47.5 ±13.4	74.7 ±2.8	51.3 ±13.4	68.9 ±4.4	46.9 ±16.0	59.0 ±12.2	74.7 ±2.8
Supervisor expectations and actions	69.4 ±9.8	58.3 ±13.1	66.4 ±3.5	72.3 ±9.4	69.0 ±0.7	59.9 ±8.9	64.0 ±3.5	63.7 ±11.8	65.9 ±3.4	69.0 ±0.7
Organizational learning	58.9 ±17.5	55.9 ±9.9	62.2 ±3.1	81.3 ±7.3	63.3 ±0.9	75.0 ±8.3	72.4 ±0.2	64.6 ±15.4	64.4 ±5.2	63.3 ±0.9
Teamwork within hospital units	64.2 ±19.6	71.7 ±8.6	69.8 ±5.0	88.0 ±4.5	93.3 ±0.9	74.0 ±8.8	73.1 ±2.7	71.9 ±14.5	70.5 ±4.7	93.3 ±0.9
Communication openness	48.5 ±16.7	55.9 ±9.5	68.6 ±3.3	48.3 ±15.9	63.0 ±0.5	50.2 ±10.2	60.1 ±5.5	51.3 ±12.9	66.9 ±5.0	63.0 ±0.5
Feedback and communication about errors	49.8 ±15.0	48.4 ±7.3	64.7 ±8.5	54.4 ±6.6	49.5 ±1.4	57.7 ±14.3	69.6 ±1.9	52.0 ±12.1	66.1 ±7.4	49.5 ±1.4
Nonpunitive response to error	37.1 ±12.4	41.9 ±19.0	63.8 ±8.0	42.1 ±15.2	57.3 ±0.0	21.1 ±3.6	40.3 ±2.5	35.7 ±15.9	58.6 ±12.5	57.3 ±0.0
Staffing	48.6 ±15.8	42.1 ±12.2	61.1 ±8.9	37.3 ±6.7	55.0 ±0.7	34.5 ±8.1	46.6 ±0.9	41.6 ±12.8	58.2 ±9.9	55.0 ±0.7
Hospital management support	64.1 ±14.9	38.2 ±12.7	54.8 ±10.1	64.0 ±8.5	62.0 ±0.0	56.0 ±23.4	70.4 ±4.4	53.9 ±19.0	57.9 ±11.2	62.0 ±0.0
Teamwork across hospital units	52.5 ±10.0	40.9 ±11.2	52.8 ±8.2	64.3 ±6.3	60.3 ±0.4	50.8 ±15.9	59.3 ±0.4	50.1 ±13.6	54.2 ±7.6	60.3 ±0.4
Hospital handoffs and transitions	37.8 ±11.6	48.8 ±11.3	60.0 ±3.0	54.2 ±10.7	58.0 ±0.7	46.9 ±12.1	60.3 ±3.9	45.6 ±12.2	60.1 ±3.0	58.0 ±0.7
English language, *n* (%)	10 (100.0)	1 (6.7)		0 (0.0)		1 (14.3)		12 (31.6)		
Data collection, *n* (%)										
2005–2009	8 (80.0)	8 (53.3)		3 (50.0)		0 (0.0)		19 (50.0)		
2010–2014	2 (20.0)	7 (46.7)		3 (50.0)		7 (100.0)		19 (50.0)		
Number of dimensions, *n* (%)										
Less than 12	2 (20.0)	3 (20.0)		1 (16.7)		1 (14.3)		7 (18.4)		
Exactly 12	8 (80.0)	11 (73.3)		5 (83.3)		6 (85.7)		30 (79.0)		
More than 12	0 (0.0)	1 (6.7)		0 (0.0)		0 (0.0)		1 (2.6)		
Healthcare facilities, median [range]	33 [1–2717]	8 [1–100]		13.5 [1–42]		11 [1–68]		10.5 [1–2717]		
Analyzed questionnaires, median [range]	1773 [45–50513]	583 [148–3779]		974 [248–6963]		1316 [99–6807]		788 [45–50513]		
Population selection, *n* (%)										
Everyone	3 (30.0)	5 (33.3)		4 (66.7)		6 (100.0)		18 (48.7)		
Only healthcare occupations	5 (50.0)	8 (53.3)		0 (0.0)		0 (0.0)		13 (35.1)		
Only non-healthcare occupations	1 (10.0)	1 (6.7)		0 (0.0)		0 (0.0)		2 (5.4)		
Mixed	1 (10.0)	1 (6.7)		2 (33.3)		0 (0.0)		4 (10.8)		