Spread of hospital-acquired infections and emerging multidrug resistant enterobacteriaceae in healthcare networks : assessment of the role of interfacility patient transfers on infection risks and control measures

Narimane Nekkab

le cnam

ÉCOLE DOCTORALE SCIENCES DES METIERS DE L'INGENIEUR

LABORATOIRE MODELISATION, EPIDEMIOLOGIE ET SURVEILLANCE DES RISQUES SANITAIRES

THÈSE présentée par :

Narimane NEKKAB

soutenue le : **25 JUIN 2018**

pour obtenir le grade de : **Docteur du Conservatoire National des Arts et Métiers**

Discipline/ Spécialité : Discipline : Biologie des populations et écologie / Spécialité : Sécurité sanitaire

SPREAD OF PATHOGENS IN HEALTHCARE NETWORKS:

ASSESSMENT OF THE ROLE OF INTER-FACILITY PATIENT TRANSFERS ON INFECTION RISKS AND CONTROL MEASURES

THÈSE dirigée par :

Madame TEMIME Laura	Professeur, Cnam	*Directrice*
Monsieur CRÉPEY Pascal	Professeur, EHESP	*Co-encadrent*

RAPPORTEURS :

Monsieur DONKER Tjibbe	Chercheur, Oxford University
Madame MERLE Véronique	Professeur, Université de Rouen

JURY :

Monsieur DESENCLOS Jean-Claude,	Directeur scientifique, InVS/SPF	*Président du jury*
Madame COLIZZA Vittoria	CR1, INSERM, UPMC	

INVITÉ :

Monsieur ASTAGNEAU Pascal	Responsable, CPias

Thèse préparée dans le cadre du Réseau doctoral en santé publique animé par l'EHESP

Thesis prepared in the framework of the public health doctoral network hosted by EHESP

Dedication

I would like to dedicate my thesis to my family whom I lovely dearly.

To Papa, for having pushed me to come to France and to attain a PhD and for all that we have been through as a family. I am who I am because of you. My strength comes from your strength. Thank you for everything.

To Mama, for always having been so patient and for supporting me every step of the way and for all that you have done for us. Words cannot express my gratitude.

To Yasmine, my little artist. I am proud of who you have become. The world is at your feet and I cannot wait to see you pursue your dreams.

To Nesrine, my angel in heaven. You are always in my heart. You are the ray of sunshine on stormy days. You are my everlasting motivation for living compassionately and working compassionately to help others.

Acknowledgements

I would like to express my deep gratitude to my thesis director, Laura Temime, and co-supervisor Pascal Crépey for their guidance these past few years.

I would like to offer my special thanks to Pascal for his encouragement during my masters' internship to pursue a thesis. Thank you for all that you have taught me, your humour, your sarcasm, your hospitality, and for giving me the opportunity to teach by your side.

I would also like to thank Laura for welcoming me into your lab in which we all have become like family, for your kindness, and for your exceptional guidance and advice. This work would not have been possible without the both of you.

I would like to express my gratitude to Pascal Astagneau for also encouraging me to pursue my internship as a thesis, for sharing your expertise, and especially for believing in me and this project. Thanks are also in order to you and other PRINCEPS members for funding the thesis.

I would also like to thank William Dab for welcoming me into your lab. Your encouragement and advice have been greatly appreciated.

I am particularly grateful for the fruitful collaboration given by Lulla Opatowski. Your suggestions regarding my work have been exceptionally helpful and much appreciated.

I wish to thank Rania Assab in particular for her friendship during the thesis. We have shared many special moments together in a lot of places especially that moment at the beach with the birds… Your humour and personality has made this experience very enjoyable and I would not have wanted to share these past three years with anyone else. I will treasure all of the laughs! Thank you!

To my lab mates Audrey, Jonathan, Lionelle, Tom, Karim, Mounia, Kévin, Sonia, Isabelle, Hélène, and Jerôme, I greatly appreciate having worked with you. Thank you for all of the laughs, food conversations, emotional support, advice, and helping me improve my French!

To my fellow Réseau doctoral friends, Annabelle, Solenn, Jonathan, Mélanie, and Gloria, it has been a pleasure organising the seminar and sharing this thesis experience together.

I would like to acknowledge the assistance from members of Santé Publique France, Bruno Coignard and Valérie Pontiès for your expertise and providing us with the data for our analyses.

I would also like to acknowledge those at EHESP: Fei Gao for technical assistance, the Réseau doctoral for enriching my thesis experience, and members of the IDEA course.

Special thanks to Vittoria Colizza and Bruno Coignard for your support as members of my thesis follow-up committee. Your advice, expertise, and encouragement are much appreciated.

I would also like to extend my thanks to the reviewers of the publications produced during the thesis whose comments have been very helpful.

I am particularly grateful to my friends Lindsey, Alexa, Aisha, Courtney, Louisa, and the students in the MPH who have always been very supportive and have provided great company. I have learned a lot from you.

And finally, thank you T. for all of your support in particular.

Summary

The spread of pathogens and multi-drug resistance in healthcare networks is a major public health issue. Evaluating the role of inter-facility patient transfers that form the structure of these networks may provide insights on novel infection control measures. Identifying novel infection control strategies is especially important for multi-drug resistant pathogens such as Carbapenemase-producing *Enterobacteriaceae* (CPE) due to limited treatment options. The increasing use of inter-individual contact and inter-facility transfer network data in mathematical modelling of pathogen spread in healthcare settings has helped these models become more realistic; however, they remain limited to a few settings and pathogens. The main objectives of this thesis were two-fold: 1) to better understand the structure of the healthcare networks of France and their impact on pathogen spread dynamics; and 2) to assess the role of transfers on the spread of CPE in France during the 2012 to 2015 period.

The French healthcare networks are characterized by centralized patient flows towards hubs hospitals and a two-tier community clustering structure. We also found that networks of patients with HAIs form the same underlying structure as that of the general patient population. The number of CPE episodes have increased over time in France and projections estimate that the number of monthly episodes could continue to increase with seasonal peaks in October. The general patient network was used to show that, since 2012, patient transfers have played an increasingly important role over time in the spread of CPE in France. Multiple spreading events of CPE linked to patient transfers were also observed.

Despite subtle differences in the flows of patients with a healthcare-associated infection (HAI) and the general patient population, the general patient network may best inform novel infection control measures for pathogen spread. The structure of healthcare networks may help serve as a basis for novel infection control strategies to tackle HAIs in general but also CPE in particular. Key healthcare hubs in large metropoles and key patient flows connecting hospital communities at the local and regional level should be considered in the development of coordinated regional strategies to control pathogen spread in healthcare systems.

Keywords: hospital-acquired infections; healthcare networks; Enterobacteriaceae; patient transfers; infection risk; infection control

Résumé

La propagation des pathogènes, notamment liées aux bactéries multi-résistantes, au sein du réseau des hôpitaux, est un grand enjeu de santé publique. L'évaluation du rôle joué par les transferts inter-établissements des patients sur cette propagation pourrait permettre l'élaboration de nouvelles mesures de contrôle. L'identification de nouvelles mesures de contrôle est particulièrement importante pour les bactéries résistantes aux antibiotiques comme les entérobactéries productrices de carbapenemase (EPC) pour lesquelles les possibilités de traitement sont très limitées. L'utilisation des données de réseaux de contact inter-individus et de transferts inter-établissement dans la modélisation mathématique ont rendu ces modèles plus proches de la réalité. Toutefois, ces derniers restent limités à quelques milieux hospitaliers et quelques pathogènes. La thèse a eu pour objectifs de 1) mieux comprendre la structure des réseaux hospitaliers français et leur impact sur la propagation des pathogènes ; et 2) évaluer le rôle des transferts sur la propagation des EPC.

Les réseaux hospitaliers français sont caractérisés par des flux de patients vers des hubs et par deux niveaux de communautés des hôpitaux. La structure du réseau de transfert des patients présentant une infection nosocomiale (IN) n'est pas différente de celle du réseau général de transfert des patients. Au cours des dernières années, le nombre d'épisode d'EPC a augmenté en France et les prédictions prévoient une poursuite de cette augmentation, avec des pics de saisonnalité en octobre. Ce travail a également montré que, depuis 2012, les transferts de patients jouent avec les années un rôle de plus en plus important sur la diffusion des EPC en France. Des évènements de propagation multiple liée aux transferts sont également de plus en plus souvent observés.

En conséquence, la structure du réseau des hôpitaux pourrait servir de base pour la proposition des nouvelles stratégies de contrôles des IN en général, et des EPC en particulier. Les hôpitaux très connectés des grandes métropoles et les flux des patients entre les communautés locale et régionale doivent être considérés pour le développement de mesures de contrôle coordonnées entre établissements de santé.

Mots clés : infections nosocomiales ; réseaux hospitaliers ; entérobactéries; transferts des patients ; risques infectieux ; mesures de contrôle

Résumé longue

Introduction

La propagation des infections nosocomiales (IN), notamment liées aux bactéries multi-résistantes (BMR), au sein du réseau des hôpitaux, est un grand enjeu de santé publique. Partout dans le monde, les systèmes de santé font face aux IN qui menacent la sécurité des patients hospitalisés, encombrent les personnels de santé, et augmentent les coûts des soins. Les IN défient les frontières internationales et se dispersent dans les établissements de santé malgré des systèmes de surveillance robustes et des recommandations de mesures de prévention et de contrôle. Pour combattre la propagation des IN et les BMR, il y a donc un besoin de nouvelles stratégies de prévention et de contrôle. Notamment, le transfert de patients d'un pays à un autre et d'un établissement de santé à un autre est associés avec la propagation de bactéries très résistantes aux antibiotiques – notamment pour des entérobactéries productrices de carbapenemase (EPC) : le transferts des patients peut être l'un des facteurs majeurs lié à la propagation spatio-temporelle des pathogènes mais aussi des épidémies. Il est ainsi primordial de mieux comprendre l'impact des transferts inter-établissement sur la propagation des pathogènes.

Cette thèse a eu pour objectifs de 1) mieux comprendre la structure des réseaux hospitaliers français et l'impact des transferts de patients sur la propagation des IN ; et 2) évaluer le rôle des transferts sur la propagation des EPC.

Pour répondre à ces objectifs, la première étape a été de faire une revue de la littérature sur l'utilisation des données de contacts inter-individu et des transferts inter-établissements. Nous avons évalué comment l'intégration des données sur des réseaux réels dans la modélisation mathématique a amélioré la compréhension et la capacité prédictive de la propagation des IN dans des milieux hospitaliers. Nous avons également fait une revue de la littérature concernant l'utilisation de la base de données Programme de Médicalisation des Systèmes d'information (PMSI) que nous avons exploitée pour la création des réseaux hospitaliers français.

Récemment, des publications de modèles mathématiques qui intègrent des réseaux inter-établissements ont approfondi la connaissance sur la manière dont les IN se propagent et sur l'optimisation des stratégies de contrôle. Néanmoins, les différences entre les structures de réseaux construits avec des patients avec un IN et des patients en général n'ont jamais été illustrées. La deuxième étape de cette thèse a donc été d'élaborer une première description des

réseaux hospitaliers français puis de comparer la structure des réseaux de transfert des patients présentant une IN à celle du réseau général de transfert des patients. Cette étape a pour objectif de nous permettre de mieux comprendre les différences entre ces réseaux et aussi l'impact de ces structures sur la propagation des pathogènes dans le système de santé.

Un modèle mathématique a ensuite été développé pour mieux comprendre comment les caractéristiques des hôpitaux dans le réseau hospitalier français peuvent prédire les épidémies des pathogènes. Les résultats de ces simulations seront décrits.

Enfin, concernant l'épidémie EPC en France, aucune étude n'a prédit le nombre d'épisodes attendu dans un futur proche. Pour répondre au deuxième objectif, nous avons décrit et prédit l'épidémie EPC avec des séries temporelles et des modèles SARIMA. Pour mieux comprendre l'impact des transferts sur l'épidémie, nous avons aussi évalué si le réseau hospitalier peut expliquer l'incidence observée et la propagation des épisodes d'EPC en France. Dans une dernière étape, nous avons utilisé des méthodes bayésiennes pour reconstruire des chaines de transmission des épisodes EPC afin de mieux comprendre les dynamiques de l'épidémie en France.

L'épidémiologie des infections nosocomiales

Les IN sont des infections qui apparaissent chez les patients au moins 48 heures après leur admission à l'hôpital. Ces infections sont liées à l'exposition à un environnement contaminé par des pathogènes infectieux, des patients infectés, ou des patients et professionnels de santé colonisés – dits « porteurs » des pathogènes. L'Organisation Mondiale de la Santé (OMS) estime la prévalence mondiale des IN à 7,6% dans les hôpitaux des pays à haut revenu et à 10,1% dans les pays à moyen et bas revenu en 2011.(1) Les *Staphylocoques dorés* et *Escherichia coli* sont les pathogènes les plus souvent identifiés dans les établissements de santé. Les IN se propagent avec une prévalence élevée dans les services de réanimation où les patients le plus vulnérables sont traités : dans les services de réanimation, il est estimé que 30% des patients dans des hôpitaux des pays à haut revenu ont eu au moins une IN. Ce chiffre est deux à trois fois plus dans les pays à moyen et bas revenu.(1) En Europe, 5% des patients sont admis dans les services de réanimation mais ils représentent 16,5% des IN.(2) Le taux de mortalité dans les services de réanimation a été estimé entre 18,5% et 29,3% aux Etats-Unis.(1) En Europe, 37 000 décès par an sont attribués directement aux IN.(1) De plus, chaque année en Europe, les IN sont la cause 16 millions de jours de séjour hospitalier et coûtent 7 milliard d'euros.(1, 3)

L'épidémiologie des infections nosocomiales en France

En France, le Réseau d'Alerte Investigation et de Surveillance des Infections Nosocomiales (RAISIN) géré par l'Institut de Veille Sanitaire (InVS)/Santé Publique France est le système national de surveillance des IN sur le territoire français. Il s'agit de cinq différents réseaux gérés par 17 Centres d'appui pour la Prévention des infections associées aux soins (CPias).

En 2012, la prévalence globale des IN en France a été estimée à 4.9% avec une incidence de 2.7%.(2) Les IN touchent 324 000 patients par an en France et sont la cause d'entre 6% and 15% des décès dans les hôpitaux.(2, 4) Les coûts annuels associés aux IN ont été estimé entre 750 millions d'euro et 1,8 milliard d'euro.(5)

En France, l'épidémiologie concernant les BMR a changé depuis 2002. Le taux d'incidence des *Staphylococcus aureus* résistant à la méticilline (SARM) a diminué de 0.63 en 2002 à 0.26 en 2015.(6) En revanche, l'incidence des bactéries productrices de béta-lactamases à spectre élargi (BLSE) a augmenté de 0.13 en 2002 à 0.67 en 2015.(6) Les quatre BLSE le plus importantes dans les établissement de santé français sont les *E. coli*, *Klebsiella pneumoniae*, *Enterobacter cloacae*, et *Enterobacter aerogenes*.(6) Dans les services d'animation, 9% des cas ont eu un SARM alors que 11% qui ont eu un BLSE en 2015.(6)

Les entérobactéries productrices de carbapenemase

Les entérobactéries, comme *E. coli* et *K. pneumonia*e, sont des pathogènes responsables de plusieurs infections dans la population. Ils sont souvent associés aux IN et sont souvent traités par des antibiotiques. Malgré le succès de ces traitements, il y a eu l'émergence de différents mécanismes de résistances contres ces antibiotiques au fil du temps. Les bêta-lactamases, par exemples, sont des enzymes utilisés par les pathogènes pour empêcher l'action des beta-lactame antibiotiques. Les carbapénèmes sont eux-aussi des antibiotiques pour lesquels ces pathogènes ont développés une résistance. Ces entérobactéries productrices de carbapenemase (EPC) sont souvent résistantes aux autres traitements et sont donc difficiles à soigner.

La consommation des antibiotiques comme les carbapénèmes, céphalosporines de troisième génération, céphalosporines de quatrième génération, et les fluoroquinolones mais aussi le transfert des patients à travers les frontières internationales ont été identifiés comme facteurs de risque de colonisation et infection par le Centre Européen de Prévention et Contrôle des Maladies (ECDC).(7) À cause des transferts internationaux des patients, les EPC se sont

installés dans beaucoup de pays à travers le monde en conservant leurs différents mécanismes de résistances. Notamment, deux épidémies mondiales ont été décrites par Nordmann et al. : Une épidémie d'une souche de *K. pneumoniae* avec différents types de mécanismes de résistances et une épidémie d'*E. coli* principalement avec le mécanisme de résistance OXA-48.(8)

En France, les premiers cas sont apparus suite à une importation internationale en 2004. Depuis, plusieurs épidémies ont été observées et le nombre d'épisodes augmente chaque année. En France, la plupart des EPC sont productrices d'OXA-48, un mécanisme qui est apparu en Turquie en 2001 et qui est devenu très prévalent en Afrique du Nord – une région avec de forts échanges de patients avec la France.(9) La mortalité associée aux EPC en France est estimée entre 30% et 70% chez les patients infectés en 2017.(10)

La propagation des infections nosocomiales

Le poids des IN sur le système de santé est important et il y a un besoin de mieux comprendre la transmission de ces pathogènes pour mieux les contrôler. Les infections sont en général transmises par trois voies : directement, indirectement, et dans l'air. En prenant en compte ces modes de transmission dans un milieu hospitalier, nous pouvons avoir trois échelles de transmission des IN : 1) les contacts inter-individuels (patients et professionnels de santé) et les contacts avec l'environnement ; 2) les contacts et les mouvements des patients et personnels inter- services au sein des établissements ; et 3) les transferts des patients entres des différents établissements de santé. Ces trois échelles créent des réseaux de contact ou de transfert. Ces réseaux peuvent servir à expliquer les dynamiques de transmission des IN. La numérisation a permis aux chercheurs de profiter des différentes sources de données, par exemple les capteurs de mouvement ou les systèmes électroniques de dossiers médicaux, pour recréer ces réseaux avec des données réelles. L'analyse de ces réseaux et leur intégration dans des modèles mathématiques a permis aux chercheurs de mieux illustrer comment les hétérogénéités des contacts peuvent aider à mieux comprendre la dynamique de transmission des IN.

Les réseaux des transferts des patients permettent de mieux comprendre la transmission des IN. En 2001, le lien entre le transferts des patients et la transmission des IN a été démontré.(11) Quelques années plus tard, un modèle mathématique a pris en compte plusieurs hôpitaux pour estimer la probabilité de portage des infection IN parmi les admissions des patients.(12) En 2010, le premier modèle individu-centré basé sur un réseau hospitalier nationale est publiée.(13) Avec d'autres publications (14-16), les réseaux hospitaliers ont apporté de nouvelles

connaissances sur la propagation et le contrôle des IN : les hôpitaux universitaires jouent un rôle très important dans la transmission des IN, la coordination régionale entre les différents établissements est primordiale pour obtenir le meilleur contrôle des pathogènes. Il y a en effet des hétérogénéités des flux de patients et des communautés d'hôpitaux qui partagent des patients. De plus, la modification du nombre de transferts, le changement de la direction d'admission, le choix des hôpitaux sentinelle, et les stratégies de contrôle coordonnée et basé sur la structure des communautés sont toutes des nouvelles propositions de stratégies de contrôle qui sortent de ces études.

L'utilisation des données de réseau dans des modelés de la propagation des IN

Le développement des outils numériques a permis aux épidémiologistes d'utiliser différentes sources de données pour mieux comprendre l'impact des activités humaines sur la santé. Il était alors nécessaire d'évaluer l'utilisation des données réelles des réseaux inter-individuels et inter-établissements dans les modèles mathématiques, et c'est ce qui a justifié notre première publication scientifique.(17) Nous avons évalué les différentes sources de données et les méthodes utilisées dans des modèles ainsi que la manière dont ils ont amélioré les connaissances sur la transmission des IN dans le milieu hospitalier.

Nous avons exploité trois bases de données pour recenser les publications sur des modèles mathématiques ou mécanistiques de la diffusion de pathogènes dans les établissements de soins : MEDLINE, Web of Science Core Collection et Institute of Electrical and Electronic Engineers (IEEE) Xplore Digital Library. Deux cent seize publications ont été identifiées dont 28 intégrant des données de contacts, 26 des données de transferts, et 22 basées sur des données des réseaux théorétiques.

L'utilisation de données de réseaux de contact inter-individuels et de transferts inter-établissements dans la modélisation mathématique a permis à ces modèles d'être plus proches de la réalité. Le nombre de publication sur des modèles dans un milieu hospitalier qui utilisent des données réelles a tendance à augmenter au fil de temps. Ces modèles représentent 27% de toutes les publications identifiées jusqu'au 26 janvier 2017. La diversité en termes de différents types de source de donnes a aussi augmenté. Les données sur les transferts proviennent des dossiers médicaux électroniques des hôpitaux, d'un système de santé national, ou des bases de données d'assurance. Les données de contacts entre individus proviennent d'observations, de questionnaires, des dossiers médicaux ou plus récemment des capteurs.

Cependant, les publications sur l'intégration des données de réseaux dans des modèles en milieu hospitalier sont limitées aux pays à haut revenu, limitées aux unités de soins de courte durée et de soins intensifs comme la réanimation, et limitées aux quelques pathogènes. Par exemple, 48% des modèles ont étudié les SARM. De plus, les méthodes d'estimation des paramètres et la validation des modèles ont été très peu utilisées dans ces articles.

Malgré ces limites, ces modèles ont proposés des stratégies de contrôle des IN plus efficaces et plus précises grâce aux données. Par exemple, les données ont permis aux chercheurs de mieux comprendre comment les variations des contacts inter-individuels peuvent mieux expliquer la propagation des pathogènes. Beaucoup de publications ont étudié l'impact du respect des règles d'hygiène des mains et des différentes stratégies de dépistages des patients en prenant en compte les contacts. Les modèles de transferts des patients affirment l'importance de la coordination régionale des stratégies de contrôle entres les hôpitaux pour mieux gérer les épidémies d'IN. Ces publications ont aussi décrit la structure de ces réseaux en utilisant l'analyse de réseaux sociaux pour identifier des « hubs » – des hôpitaux très connectés qui jouent un rôle très important dans la structure du réseau – et une structure de communauté des hôpitaux qui partagent la même population des patients.

L'utilisation des données de réseau dans la modélisation est devenue plus fréquente. Les modèles en milieux hospitaliers ont apporté beaucoup de connaissance sur la propagation des pathogènes mais aussi leur contrôle. De nouvelles innovations sur le recueil des données et leur utilisation dans la modélisation sont nécessaires pour mieux comprendre des dynamiques des IN dans les établissements de santé.

La base de données PMSI

Pour construire les réseaux hospitaliers français, nous avons exploité la base de données Programme de Médicalisation des Systèmes d'information (PMSI). Cette base de données est exhaustive sur les activités hospitalières et contient, pour chaque séjour d'un patient hospitalisé, un résumé de sortie standardisé (RSS). Chaque établissement de santé est identifié avec un Fichier National des Etablissements Sanitaires et Sociaux et chaque séjour est numéroté, ce qui nous permet de reconstruire le flux des patients entres les établissements.

Nous avons fait une revue de la littérature pour évaluer l'utilisation de la PMSI pour les études épidémiologiques des IN. Sept publications ont été identifiées sur la base MEDLINE. Plusieurs publications ont montré que l'utilisation du code Classification Internationale des Maladies (CIM-10) pour les IN – Y95 – n'a été pas suffisant en termes de sensibilité et de spécificité pour

identifier les patients avec des IN dans le PMSI. D'autres codes diagnostiques recommandés par le système de surveillance ont été évalués et une étude a montré que ces codes ont amélioré la détection des IN en termes de sensibilité et de spécificité.(18) Le PMSI est une base très exhaustive avec beaucoup d'information sur les trajectoires des patients ; cependant, pour identifier les patients avec un IN, l'utilisation des autres codes diagnostiques a été nécessaire.

Les réseaux hospitaliers français

Les études identifiées dans notre revue de la littérature ont montré que la transmission des IN était dépendante de la structure des réseaux hospitaliers et des transferts de patients(15). Cependant, ces réseaux sont basés sur les transferts de tous les patients, sans distinction des pathologies. Notre objectif a été de mettre en comparaison le réseau de transferts de tous les patients et celui des patients présentant une IN. En utilisant les données des transferts de 2,3 millions de patients en 2014 en France, trois réseaux hospitaliers français ont été construits et décrits : un réseau de transferts de tous les patients, un réseau de transferts des patients présentant une IN codée avec le code Y95, et un réseau des transferts des patients suspects d'avoir présenté une IN avec des codes additionnels tels que décrits par Gerbier et al.(16) Les trois réseaux ont été comparés en termes de flux des patients, topologie du réseau, et communauté des hôpitaux.

Le réseau de tous les patients était composé de 2 063 hôpitaux et 50 026 liens. Le réseau des patients suspectés d'avoir présenté une IN était composé de 1 975 hôpitaux et 18 812 connexions de 128 681 transferts. Le réseau des patients présentant une IN avec le code Y95 était composé de 1 266 hôpitaux et 3 722 connexions de 13 627 transferts. Le nombre moyen de connexions des hôpitaux au sein de chaque réseau était de 48, 19 et 5,88. Le moyen nombre de patients transférés dans chaque lien était de 14, 5 et 2,3 patients.

Les réseaux ont été caractérisés comme des réseaux invariants d'échelle (dont les degrés suivent une loi de puissance) avec des flux hétérogènes mais très centralisés. Les degrés représentent le nombre de connexions d'un hôpital à l'autre au sein du réseau. Des réseaux avec des invariants d'échelle ont une distribution des degrés dont la plupart des hôpitaux sont très peu connecté comparer aux autres qui sont très connectés. Ces hôpitaux très connectés sont des « hubs » – des hôpitaux privés, des centres hospitaliers, et des centres hospitaliers universitaires. Les hôpitaux dans les trois réseaux ont été classés par leur degré et ont été comparés avec un test de Wilcoxon. Il n'y avait pas une différence significative entres les « hubs » dans les trois réseaux malgré la différence de taille des réseaux et la population des patients. Les réseaux ont eu aussi

un effet du « petit monde » dont entre trois à cinq transferts ont été suffissent pour transférer des patients à tous les hôpitaux dans le réseau.

Etant donné que les réseaux sont de taille différentes, nous avons développé une analyse pour mieux comprendre les différences entres les structures de ces trois réseaux. Nous avons comparé des caractéristiques du réseau de tous les patients avec les valeurs moyennes de mille réseaux construits du même nombre de patients des réseaux des patients présentant des IN. Ces mille réseaux ont été construits avec le même nombre de patients sélectionnés aléatoirement dans la population globale de tous les patients. Pour le diamètre (la distance entre deux hôpitaux étant définie comme la distance la plus longue des plus courts chemins entre deux hôpitaux), la moyenne des courts chemins entres les hôpitaux, et le coefficient de « clustering » (nombre de triangles dans le réseau), ces caractéristiques ont été plus similaires aux réseaux des patient présentant des IN que le réseau de tous les patients. Ça veut dire que les différences observées entre ces trois réseaux ont été associées à la différence de taille et ne pas aux transferts. Ces analyses ont montré que la structure du réseau de transfert des patients présentant une IN n'est pas différente de celle du réseau général de transfert des patients.

Nous avons évalué la structure des communautés dans les réseaux hospitaliers français. Une « communauté » est définie comme un groupe d'hôpitaux qui partagent la même population des patients. En utilisant l'algorithme Greedy (19), nous avons identifié 18 communautés d'hôpitaux. L'analyse des distances géographiques (grâce à la géolocalisation) entre les hôpitaux de la même communauté a montré que ces réseaux correspondent aux régions administratives de la France. Quatre-vingt-treize pour cent des patients ont été partagés entre des hôpitaux de la même communauté. Nous avons aussi utilisé un autre algorithme pour identifier des communautés prenant en compte la direction des mouvements des patients et aussi le nombre de patients dans chaque lien (20). Nous avons observé 132 communautés au niveau local qui correspondait aux départements de la France. Quatre-vingt pourcent des patients ont été partagées entres des hôpitaux de ces communautés. En France, nous avons montré qu'il y a des communautés à deux niveaux : au niveau local des transferts entre des départements et au niveau régional. Les tranferts suivent un mouvement centralisé en France : d'abord vers un ou deux hôpitaux dans le département, puis vers le centre hospitalier universitaire de la région, puis entre les régions puis inter-régional avec une prédominance centrée sur l'Ile-de-France vers Ile-de-France.

Une étude de simulation des épidémies sur les réseaux hospitaliers

Une étude préliminaire a été menée en début de thèse pour mieux comprendre comment des pathogènes peuvent se propager au sein des réseaux hospitaliers. En utilisant les réseaux hospitaliers de 2012 (similaires aux trois réseaux de 2014), nous avons développé un modèle mathématique SIS (les établissements de santé sont dans un état soit Susceptible soit Infecté selon la probabilité des patients d'être infectés ou d'être porteur d'une infection dans l'établissement). Les simulations ont montré que des établissements très connectés en termes de degrés (nombre de connexions), puissance (nombre de patients transférés), et « betweenness » (rôle intermédiaire) ont prédit de très fortes probabilités d'épidémie. Ces caractéristiques peuvent prédire si un établissement peut entretenir une épidémie dans le réseau. Des établissements très connectés doivent être des cibles pour des mesures de contrôle des IN à une grande échelle. Cependant, une recherche complémentaire est nécessaire pour mieux évaluer l'impact des diffèrent mesures de contrôle sur ces réseaux.

Prédiction de nombres des épisodes d'EPC en France

Les IN couvrent un large spectre de pathogènes. Pour répondre à l'un des problèmes majeurs de santé publique – la résistance aux antibiotiques – la thèse a eu pour objectif de décrire l'épidémie des EPC en France. Le nombre d'épisodes d'EPC n'a eu cesse d'augmenter depuis leur introduction en 2004. Il y a eu 2 346 épisodes entre septembre 2011 et décembre 2015 collectés par le système de surveillance le Réseau d'Alerte, Investigation et de Surveillance des Infections Nosocomiales (RAISIN).

La majorité des épisodes d'EPC en France ont le mécanisme de résistance OXA-48. Un total de 1 110 épisodes a été lié à des cas importés pendant toute la période. L'épisode avec le plus grand nombre de cas (n = 194) est survenu en septembre 2012 d'une souche OXA-48 suivi d'une épidémie de 149 cas en octobre 2012 d'OXA-48. Ces deux épidémies n'ont pas été liées à l'importation internationale de souches OXA-48. La majorité des épisodes (n = 469) a eu lieu à Paris dans la région Ile-de-France avec la majorité (n = 324) ~~étant~~ liée à l'importation internationale. Les deuxième et troisième épisodes les plus fréquents se sont produits dans les départements voisins du Val-de-Marne (n = 83) et des Hauts-de-Seine (n = 159). Ces épisodes ont été dominés par OXA-48 (n = 347 à Paris et n = 208 en Val-de-Marne respectivement) suivi d'un autre mécanisme de résistance – NDM (n = 95 à Paris et n = 39 dans les Hauts-de-Seine respectivement). La plupart des cas de KPC (un mécanisme de résistance) se trouvent dans le Val-de-Marne (n = 23) suivi de Paris (n = 20). Les cas OXA-48 étaient les plus fréquents et les

plus dispersés couvrant 87 départements sur un total de 101 départements (95 départements continentaux dont la Corse, Monaco et 5 départements d'outre-mer).

Nous avons utilisé des modelés SARIMA pour prédire le nombre d'épisodes attendu à la fin de 2019. Le modèle a prédit un nombre croissant d'épisodes au cours du temps : une moyenne estimée de 122 épisodes d'EPC par mois à la fin de 2016, 151 épisodes EPC par mois à la fin de 2017, 177 épisodes EPC par mois à la fin de 2018, et 204 épisodes EPC par mois à la fin de 2019. Le pic de nombre d'épisodes a été attendu en octobre de chaque année. Nous prévoyons une augmentation du nombre d'épisodes d'un seul cas (jusqu'à 200 épisodes par mois d'ici la fin de 2019) et la stabilisation des épisodes avec plus d'un cas (moins de 25 épisodes par mois à la fin de 2019). Le nombre d'épisodes non associés à un cas importé devrait augmenter à un rythme plus élevé que celui des épisodes associés à un cas importé, ce qui suggère que la propagation locale soutiendrait l'épidémie. Les épisodes d'OXA-48 étaient prédits de continuer à dominer à un taux plus élevé que celui des épisodes NDM, KPC et VIM. En 2014, le nombre d'épisodes NDM était plus élevé que celui du KPC et du VIM. À la fin de 2016, 2017, 2018 et 2019, une moyenne de 86, 98, 111, et 124 OXA-48 épisodes par mois a été prédite. Les cas de NDM devraient également augmenter avec une moyenne prévue de 26 épisodes (95% PI [17-34], 80% PI [20-31]) par mois d'ici la fin de 2019.

En conclusion, le nombre d'épisodes d'EPC en France a augmenté au fil des ans. Le nombre d'épisodes devrait augmenter jusqu'à atteindre le double du nombre d'épisodes d'ici la fin de 2019 par rapport à la fin de 2015. Les épisodes non liés à l'importation internationale devraient prédominer et soutenir l'épidémie. Des précautions doivent être prises pour contrôler la propagation locale des EPC dans les années à venir.

Le rôle du réseau des transferts dans la transmission et incidence d'EPC en France

La plupart des épisodes d'EPC en France observés au cours des années ont été associés aux transferts transfrontaliers et locaux entre les établissements de santé (81). Par conséquent, afin de mieux comprendre l'épidémie d'EPC et comment le mieux contrôler, cette thèse a également impliqué l'évaluation de l'impact des modèles de transfert de patients sur la propagation des EPC. En utilisant le réseau de tous les patients décrit précédemment en France en 2014 (188), l'étude reposait sur la méthode statistique publiée précédemment pour tester empiriquement la contribution de ce réseau sur la propagation d'EPC au cours de la période 2012 à 2015 (197).

Sur les 2 273 épisodes d'EPC signalés entre 2012 et 2015, nous avons identifié l'épisode le plus probable d'avoir infecté chaque cas d'incident non-importé (n = 1 251) en sélectionnant les

épisodes candidats ayant la distance au sein du réseau la plus courte de chaque épisode incident. La distribution des distances les plus courtes a été comparée à 500 simulations de permutations des données. La distribution spatiale des contaminateurs potentiels et des épisodes d'incidents ont été également décrits.

Quatre-vingt-dix pourcent des épisodes d'incidents avaient un épisode potentiellement infectieux identifié pour toute la période d'étude ; toutefois, lorsque nous avons stratifié les données par année, seulement des épisodes en 2013, 2014 et 2015 avaient significativement des distances plus courtes que les permutations. Ceci suggère que l'épidémie d'EPC en France est passée d'une épidémie soutenue par l'importation d'épisodes d'autres pays avant 2013 à une épidémie soutenue par des événements de transmission locale soutenus par des transferts de patients. De plus, le nombre d'événements reliant les « infecteurs » potentiels à des épisodes multiples a augmenté au fil des ans, ce qui suggère que des métropoles fortement connectées peuvent avoir provoqué des épidémies par transfert de patients. Des événements de transmission survenant plus localement à des distances géographiques rapprochées ont également été observés.

En conclusion, les travaux réalisés au cours de la thèse décrivent la structure du réseau hospitalier français et nous avons évalué que les transferts des patients peuvent avoir un rôle important dans la transmission des IN. L'étude a révélé qu'une grande majorité des épisodes d'EPC observés pouvaient être liés aux transferts de patients. L'exemple particulier du rôle des transferts de patients dans la transmission d'EPC sert à mettre en évidence l'importance de considérer la structure du réseau dans le développement de nouvelles stratégies de prévention et de contrôle. En conséquence, l'étude suggère que les stratégies de prévention et de contrôle des infections coordonnées devraient maintenant se concentrer sur les transferts de patients à risque d'être porteurs d'EPC pour réduire la transmission régionale et interrégionale.

Les chaines de transmission des EPC

Sur la base de notre analyse de la contribution des transferts de patients à la transmission d'EPC en France, on peut conclure qu'en 2015 au moins, les EPC se sont diffusés sur le réseau hospitalier français. Nous avons donc tenté de reconstruire les chaînes de transmission des épisodes d'EPC entre les hôpitaux français. Cette analyse a été réalisée en adaptant une méthode bayésienne initialement développée pour reconstruire des épidémies basées sur des données épidémiologiques et génomiques (paquet Outbreaker 2 sur le logiciel R, développé par T. Jombart, F. Campbell, R. Fitzjohn) (194, 195), mais sans données génomiques et pour une

grande épidémie à l'échelle nationale. La méthode a été utilisée pour déterminer les chaînes de transmission possibles en fonction de la connaissance des dates des épisodes d'EPC, du poids des liens dans le réseau hospitalier et du statut d'importation des épisodes. L'étude visait à identifier les sources des EPC dans le réseau hospitalier et à déterminer la proportion des épisodes EPC secondaires expliquées à partir des données de transferts.

Quelques observations peuvent être faites sur certaines de ces analyses préliminaires : tous les épisodes importés ne faisaient pas partie d'une chaîne de transmission, beaucoup n'étaient liés à aucun épisode, et nombre de cas secondaires ont eu de multiples sources potentielles, même si l'incertitude de la source était souvent haute.

Étant donné que les épisodes d'EPC peuvent donner lieu à de nombreux autres cas secondaires, la deuxième analyse a évalué l'ensemble de données complètes de chaque cas d'EPC individuel. Puisque plusieurs épisodes partageaient la même date de notification mais n'étaient pas produits à la même date, différentes méthodes ont été développées pour modifier les dates des cas.

Lorsque les dates de cas étaient transformées à l'aide d'une distribution de Poisson, la plus grande composante connexe était obtenue pour un temps de génération de 37 jours. Ce résultat était similaire aux résultats de l'étude précédente évaluant le rôle des transferts de patients sur la propagation des EPC, où l'analyse de sensibilité permettait de prévoir un intervalle de temps entre les infections potentielles et les épisodes d'incidence entre 20 et 30 jours. Cependant, en fonction de la transformation utilisée, d'autres fenêtres temporelles telles que 10 à 20 jours ou 40 à 50 jours auraient aussi pu être pertinentes. La taille des composants connectés ou le nombre de connexions peuvent ne pas être les critères les plus appropriés pour la sélection du temps de génération; par conséquent, d'autres critères épidémiologiques et pertinents doivent être pris en compte.

Cette étude a montré la possibilité de reconstruire de grandes épidémies sans données génomiques. Les réseaux hospitaliers mobilisés pour construire des chaînes de transmission des EPC ont été en mesure d'expliquer le pourcentage d'épisode EPC secondaires maximum possible et d'identifier les « hotspots » hospitaliers de propagation des EPC; cependant, poursuivre ce travail est nécessaire est nécessaire pour affiner la compréhension des liens entre les épisodes.

Discussion

L'amélioration des connaissances actuelles sur la dynamique et les mécanismes qui conduisent à la propagation des IN dans les établissements de santé et la proposition de mesures innovantes de prévention et de contrôle des pathogènes sont d'une importance majeure pour la santé publique. Les IN se produisent dans les systèmes de santé dans le monde entier en raison de plusieurs facteurs : l'augmentation de la charge de morbidité des patients qui risquent de rester plus longtemps à l'hôpital et risquent de décéder, ce qui représente un poids conséquent pour le personnel soignant ainsi qu'une charge financière. Parmi tous les pathogènes présents dans les établissements de santé, les souches très résistantes aux antibiotiques constituent une menace majeure pour la sécurité des patients et dans certains cas, les mesures de prévention peuvent être les seules méthodes pour lutter contre leur dissémination lorsque le traitement antibiotique n'est plus envisageable. Cette thèse cherchait à élucider de nouvelles pistes de recherche sur le contrôle des pathogènes ciblant les entérobactéries multi-résistantes en modélisant leur dynamique de propagation dans les milieux de soin. Pour atteindre ces objectifs, trois articles dans des revues internationales à comité de lecture ont été publiées ou soumises: une revue systématique de la modélisation mathématique de la propagation des pathogènes en milieu de soin, une analyse approfondie des réseaux hospitaliers français, et une évaluation du rôle des transferts de patients dans la transmission des EPC en France. La discussion qui suit vise à éclairer les implications de ces découvertes pour la santé publique et comment elles peuvent fournir un travail de soutien à de nouveaux moyens de prévention et de contrôle des pathogènes dans les établissements de santé à l'échelle nationale, en particulier pour la France, mais ainsi que dans d'autres contextes.

Un domaine de recherche dans laquelle des mesures innovatrices de prévention et de contrôle ont été proposées a été le domaine de la modélisation mathématique. Depuis 15 ans les modèles mathématiques ont fourni un cadre théorique pour comprendre la dynamique complexe de la transmission dans les établissements de soins (102-105). De plus, ils ont fourni une approche quantitative pour estimer l'impact de diverses stratégies pour lutter contre les infections, la colonisation et leurs effets combinés (103-105, 201). Le nombre de publications sur les modèles mathématiques des pathogènes dans les établissements de santé est devenu plus fréquent au fil des ans. Plusieurs facteurs peuvent avoir conduit à cette augmentation observée : y compris l'utilité perçue des modèles comme outils pour analyser l'impact de la prévention et du contrôle des infections dans le domaine de la santé, pour comprendre les causes des récentes épidémies importantes telles que l'épidémie de SARS 2002-2003 (202-205) et l'épidémie d'Ebola 2014-

2015 (206-209) ou la prise de conscience des facteurs contribuant à l'impact global de la résistance aux antibiotiques (210). En raison de l'utilisation des données numériques dans le domaine de l'épidémiologie au fil des ans, comme la disponibilité accrue des dossiers médicaux numérisés et le développement de la technologie des capteurs pour surveiller les contacts interindividuels, les chercheurs ont pu construire des modèles plus réalistes. D'autres innovations dans le recueil des données sur la structure des réseaux et l'incidence des IN, la mise en œuvre des données de modélisation et l'étalonnage et la validation des données dans les modèles sont nécessaires pour renforcer les recommandations existantes et évaluer de nouvelles stratégies de contrôle dans les établissements de santé.

L'analyse des réseaux hospitaliers français a fourni une première description des schémas de transfert des patients au niveau national. L'étude a montré que les profils de transfert des patients qui ont présenté une IN pendant leur séjour et d'autres patients avec différentes maladies et comorbidités étaient soumis à la même dynamique de réseau. Comparés aux aux autres réseaux hospitaliers en Angleterre (162, 170), aux Pays-Bas (147), en Écosse (163) ou aux Etats-Unis (172), les réseaux hospitaliers français sont des systèmes très centralisés. Les centres hospitaliers universitaires et les hôpitaux privés des principales métropoles françaises dominent le flux des patients. Des études menées en France ont montré que les hôpitaux fortement connectés peuvent héberger davantage de cas de bactériémie à SARM (147, 160, 170, 211) et que les IN étaient les plus répandues dans les centres anticancéreux, les hôpitaux universitaires et les hôpitaux militaires (212). Par conséquent, ces établissements de santé peuvent être les plus susceptibles de transmettre des infections nosocomiales dans l'ensemble du réseau par le transfert des patients infectés ou colonisés. Les pathogènes peuvent se propager à un rythme plus élevé que prévu par hasard en raison de la centralisation du mouvement des patients et du faible nombre moyen de transferts requis pour que les patients puissent se déplacer dans le réseau.

Ces résultats concordaient avec les résultats préliminaires du modèle SIS, qui montrent que la probabilité la plus élevée d'une épidémie persistante survient lorsque les hôpitaux « hub » sont infectés au départ. Un modèle de simulation du réseau anglais a également identifié les hôpitaux universitaires comme centres d'incidence du SARM. L'étude a recommandé ces hôpitaux comme cibles idéales pour des mesures d'intervention comme le dépistage des patients sortis de ces hôpitaux comme moyen de contrôle plus efficace (170). Parallèlement, l'étude du réseau français indique également que les centres hospitalier sont des cibles pour la surveillance sentinelle, en plus des cibles prioritaires des stratégies pour réduire au mieux la transmission

des pathogènes dans le pays (154). Dans toutes les situations où les patients sont transférés dans centres de soins concentrés, une attention particulière devrait être accordée à tous les risques potentiels liés au portage des pathogènes potentiellement infectieux pendant l'admission.

Ces dernières années, des études qui ont modélisé la propagation des pathogènes dans les réseaux hospitaliers ont plaidé pour des stratégies régionales coordonnées comme l'un des moyens les plus efficaces pour réduire leur dissémination au niveau régional et national. L'étude du réseau français a été similaire à celle du réseau anglais où la majorité des transferts se faisait par le partage intra-régional de patients et les flux de patients étaient centrés sur le centre hospitalier universitaire ou régional au sein de la communauté (170). Nous avons également démontré qu'il existait une structure de communauté hospitalière à deux niveaux en France. Les communautés hospitalières ont été identifiées à la fois au niveau régional, en cohérence avec les régions administratives françaises, et au niveau subrégional ou départemental. Les différences entre les communautés départementales du réseau des patients suspectent d'avoir présenté des IN et les communautés départementales du réseau général peuvent être importantes pour distinguer les hôpitaux ayant un potentiel plus élevé d'héberger des patients IN, avec des conséquences possibles en termes de prédiction de propagation des pathogènes responsables des IN; cependant, cela nécessite une étude plus approfondie. Des mesures de contrôle coordonnées localement telles que le dépistage des transferts de patients à risque et des précautions de contact basées sur la centralité d'un hôpital de sortie avec des cas connus ou à risque de portage peuvent être la première ligne de défense contre la transmission des pathogènes responsables des IN dans les régions.

Des trajectoires intermédiaires importantes peuvent jouer un rôle clé dans la propagation des pathogènes entre les hôpitaux « hub » et entre les communautés. Une étude a montré que la modification du nombre de patients transférés entre les communautés peut, par exemple, réduire la propagation du SARM (162). L'étude a également montré que même si les connexions directes d'un établissement étaient des facteurs de risque importants pour un voisin, des connexions plus faibles offraient également des voies indirectes idéales pour que les pathogènes circulent plus loin et plus vite dans le réseau. En outre, dans le cas des EPC, « Les chercheurs de cette étude ont constaté qu'en termes de nombre absolu de patients colonisés admis à l'hôpital par des transferts de la même région par rapport à les transferts provenant de l'extérieur de leur région, les transferts survenant dans la même région représentaient plus de menaces (214). Par conséquent, ces études sont parallèles aux observations d'une structure de réseau à deux niveaux dans laquelle les risques d'infection et dépendent des deux niveaux. Des liens

légèrement plus faibles au niveau subrégional peuvent jouer un rôle plus important dans la dynamique de transmission en termes de nombre absolu de transferts et de risque de transmission par rapport aux liens intercommunautaires fortement pondérés et interconnectés entre les hubs. La simulation de la propagation d'un pathogène hypothétique dans les réseaux a montré que même si les hôpitaux introduisaient le pathogène de la grande métropole dans les régions, le nombre le plus élevé d'hôpitaux infectés provenait des hôpitaux universitaires qui diffusaient l'infection aux clusters locaux.

La structure des transferts à deux niveaux peut éclairer des stratégies coordonnées à un niveau plus local où les établissements de santé identifient non seulement les patients à risque transférés des centres hospitaliers universitaires, mais aussi les risques provenant des centres hospitaliers locaux. Les stratégies de contrôle des infections - pour le contrôle à court terme - devraient davantage s'appuyer sur la dynamique au niveau des départements pour minimiser les épidémies au niveau hospitalier et la transmission aux hôpitaux voisins. À long terme, la dynamique des communautés régionales peut donner des indications sur la propagation progressive de souches spécifiques des pathogènes IN. Des études sont nécessaires pour valider ces recommandations et quantifier les mesures de contrôle. En outre, d'autres études sont également nécessaires pour évaluer la dynamique temporelle de la propagation des pathogènes dans les réseaux afin d'identifier les schémas des flux de saisonnalité potentiels et la manière d'empêcher les bactéries multi-résistantes émergentes de devenir endémiques.

Des mesures de prévention à long terme sont également nécessaires pour empêcher que les nouveaux pathogènes ne deviennent endémiques. Réduire la connectivité hospitalière afin de réduire le risque de propagation dans les réseaux est au cœur de nombreuses propositions de contrôle des infections (129, 170). La décentralisation du système de santé et plus spécifiquement celle des ressources humaines et des services de santé spécialisés vers les régions et départements peut aider à réduire la forte connectivité des centres hospitaliers dans les métropoles et à rediriger les transferts de patients. La France s'est orientée vers des stratégies de régionalisation avec la création d'agences régionales de santé, mais peu efficaces (215, 216). En outre, le nombre d'hôpitaux universitaires peut être insuffisant en France, il est d'ailleurs inférieur à celui du Royaume-Uni, un pays d'une population de même taille. Par conséquent, une solution structurelle pour alléger la propagation des pathogènes responsables des IN pourrait être d'augmenter le nombre d'établissements fournissant des services spécialisés et de les distribuer au niveau local pour aider à réorienter le flux des patients et potentiellement éviter la dispersion à grande échelle.

Les épisodes d'EPC en France continuent d'augmenter chaque année. La région Ile-de-France, qui comprend Paris, présente l'incidence d'épisodes d'EPC la plus élevée, y compris le plus grand nombre d'épisodes liés à des cas importés au niveau international. Paris est un centre de santé et attire un grand nombre de patients qui recherchent des soins spécialisés, ce qui peut les exposer à un risque plus élevé d'infection par les EPC. Les risques d'infection ou de colonisation sont donc double : le patient a besoin de soins spécialisés ce qui peut augmenter les facteurs de risque d'infection ou colonisation individuels, car les services spécialisés peuvent nécessiter une intervention chirurgicale et d'autres interventions invasives ; mais aussi, exposer le patient à un plus grand nombre de contacts potentiels avec les porteurs d'EPC en raison de l'incidence élevée. Cependant, la forte concentration des hôpitaux universitaires en Île-de-France peut biaiser la surveillance des EPC parce que ces hôpitaux peuvent avoir plus de ressources et sont davantage à-même de détecter des cas d'EPC.

L'étude a prédit une stabilisation des épidémies EPC en France pour les prochaines années. Par conséquent, on peut supposer que les mesures déjà en place au cours de la période 2010-2015 pour contrôler les épidémies ont été efficaces et devraient se poursuivre afin de contrôler la transmission afin d'éviter les épidémies récurrentes dans les hôpitaux. Il convient de noter, cependant, que plusieurs épisodes représentaient entre deux et 200 cas d'EPC chacun. Par conséquent, un plus grand nombre d'événements de transmission de personne à personne peuvent continuer à se produire; cependant, le nombre d'épidémie devrait se stabilisée. L'étude a également prédit une augmentation du nombre d'épisodes uniques ; mais cela peut être dû au fait que le système de surveillance n'a pas réussi à relier les événements de transmission ou que les données de surveillance sont incomplètes.

Les stratégies de prévention et de contrôle de la propagation des pathogènes sont particulièrement pertinentes dans l'épidémie d'EPC. La dynamique de la transmission des EPC en France a évoluée. Les plus fortes preuves de transmission d'EPC soutenues par le réseau de transferts ont été observées en 2015. Ces résultats suggèrent qu'entre 2013 et 2014, les transferts régionaux et interrégionaux ont de plus en plus contribué à la propagation des EPC. En outre, le délai de notification des épisodes d'EPC entre hôpitaux dû au transfert de patients est estimé entre 20 et 30 jours après la notification initiale dans un hôpital de référence. Par conséquent, cette étude a non seulement lié les événements des EPC aux transferts de patients au cours des dernières années, mais a également estimé le délai de notification de ces événements. Ces deux constatations renforcent les recommandations soulignant l'importance de considérer le transfert des patients comme un facteur de risque critique pour l'introduction des EPC dans les

établissements de santé, ce qui pourrait aider le système de surveillance à estimer les périodes à risque pour les épidémies liées aux transferts de patients hospitalisés avec des cas d'EPC.

La structure du réseau peut également expliquer les observations de ces événements au niveau local. Dans notre étude, la majorité d'épisodes que nous avons identifiés comme reliés s'étaient produits dans le même département ou à une courte distance géographique. Comme nous l'avons déjà mentionné, nous nous attendions à ce que le partage des patients dans les communautés locales au niveau du département joue un rôle important dans la dynamique de la propagation des pathogènes. Le nombre croissant de ces événements de transmission des EPC possibles dans un même département peut s'expliquer par le fait que la plupart des transferts de patients dans le réseau hospitalier ont eu lieu au niveau local. Par conséquent, la plus grande proportion de transmissions d'EPC peut avoir eu lieu entre des hôpitaux voisins et les autorités de santé publique l'importance de surveiller les transferts locaux de patients pour les porteurs potentiels d'EPC à risque devrait être prise en considération afin d'avoir l'impact le plus efficace sur l'EPC.

Bien que les transferts de patients ne soient certainement pas la seule explication de l'augmentation des épisodes d'EPC observés entre 2013 et 2015, ces études suggèrent qu'ils ont joué un rôle de plus en plus important au fil du temps. Les épisodes d'importation internationale pourraient également avoir contribué à près de la moitié de la propagation des EPC en France par exemple. Ces résultats sont en accord avec les descriptions des épidémies observées dans la littérature scientifique dans lesquelles les cas importés et non importés ont conduit à des cas secondaires d'EPC dans différents hôpitaux. En outre, l'hétérogénéité des politiques de contrôle des infections entre les différents types d'établissements de santé en France, et la mise en œuvre limitée de stratégies spécifiques de contrôle des EPC ont pu conduire à un mauvais contrôle des EPC et par conséquent à une dissémination dans le temps (217).

Aucune association n'a été observée entre le nombre de cas par épisode infectieux potentiel et le nombre d'épisodes secondaires. D'une part, cela pourrait signifier que les mesures de contrôle ont empêché les grandes épidémies hospitalières au cours de la période 2014-2015 ; d'autre part, comme mentionné précédemment dans l'étude de prédiction, la plupart des rapports étaient des épisodes d'un seul cas suggérant un échec potentiel des autorités de surveillance à identifier les cas uniques dans la même chaîne de transmission d'autres épisodes signalés.

La dynamique du réseau peut également aider à expliquer les épidémies d'EPC. Le nombre d'épisodes d'EPC impliquant plusieurs épisodes a augmenté avec le temps. Par exemple, un épisode OXA-48 importé à Paris a été lié à neuf autres épisodes dans neuf départements différents en France en 2015. Paris a été identifié comme le plus grand centre, non seulement pour les épisodes EPC liés à l'importation, mais aussi pour le transfert des patients. D'autres exemples mettent davantage en évidence le rôle important des liens avec un grand nombre de transferts de patients dans la connexion des hôpitaux géographiquement éloignés en termes de transmission des pathogènes. Ces observations soulignent l'importance pour les autorités sanitaires d'améliorer les efforts de contrôle dans les métropoles et les établissements de soins fortement connectés. Ces efforts vont également de pair avec la coordination entre les systèmes de surveillance régionaux, les laboratoires d'experts locaux et les autorités sanitaires régionales afin de détecter rapidement les cas d'EPC. Les établissements de santé devraient également être invités à notifier rapidement tous les cas ou contacts potentiels. En outre, des mesures de dépistage, des précautions de contact et une cohorte stricte de patients (qui se sont révélés particulièrement efficaces dans une éclosion d'EPC particulière (81, 83)) devraient être mises en œuvre une fois les cas identifiés.

Conclusion

Les réseaux hospitaliers ont joué un rôle important dans l'élucidation du rôle des transferts des patients dans la propagation des infections nosocomiales. Le principal facteur limitant de ce travail a été l'absence d'une étude de modélisation détaillée quantifiant l'impact de la propagation des pathogènes dans les réseaux de soins de santé ; Cependant, cet ensemble de travaux fournit une base pour les travaux futurs. En outre, les diverses conclusions des études menées au cours de la thèse peuvent aider à éclairer les implications des transferts de patients sur la santé publique en matière de propagation, afin de fournir un travail de soutien aux nouvelles méthodes de prévention et de contrôle des pathogènes à l'échelle nationale. Ces études soutiennent les efforts coordonnés au niveau local, régional et national entre les établissements de santé qui nécessitent l'aide du système de surveillance afin de coordonner ces efforts et de mobiliser les hôpitaux pour mettre en œuvre de nouvelles mesures en conséquence. De plus, ces efforts nécessitent la collaboration des hôpitaux universitaires qui jouent un rôle important dans la structure du réseau de santé. L'identification de transfert de patient à risque en utilisant des mesures de topologie de réseau peut s'avérer utile dans le cas de pathogènes spécifiques telles que des bactéries multi-résistantes. Les infections à EPC ont été liées à l'importation internationale et à la propagation locale ; par conséquent, il est primordial que la

dynamique des réseaux de santé soit prise en compte dans le processus de prévention et de contrôle des cas. Cette thèse sert à souligner l'importance de la structure du réseau hospitalier dans le développement de mesures efficaces de prévention et de contrôle des infections.

Table of Contents

List of tables

List of figures

Published and submitted papers

Article 1

Assab* R, Nekkab* N, Crépey P, Astagneau P, Guillemot D, Opatowski L, Temime L. Using data on network structures to inform transmission dynamics of infections in healthcare settings: a review of mathematical models. Published August 2017 (*Current Opinion in Infectious Diseases*)

* *the first authors contributed equally to the paper*

Article 2

Nekkab N, Astagneau P, Temime L, Crépey P. Spread of hospital-acquired infections: a comparison of healthcare networks. Published August 2017 (*PLoS Computational Biology*)

Article 3

Nekkab N, Crépey P, Astagneau P, Opatowski L, Temime L. Assessing the role of inter-hospital patient transfer in the spread of carbapenemase-producing *Enterobacteriaceae*: the case of France between 2012 and 2015. (*submitted Eurosurveillance*)

Article 4

Crépey P, Nekkab N, Temime L, et al. Modeling the evolution in time of CPE episodes in France. *(in preparation) (author list not yet definitive/liste des auteurs non définitive)*

Article 5

Nekkab N, Crépey P, Temime L, et al. Transmission tree construction of CPE episodes in France, 2010-2015. *(in preparation) (author list not yet definitive/liste des auteurs non définitive)*

Introduction

Worldwide, healthcare systems are confronted with the spread of pathogens that can lead to healthcare-associated infections (HAIs) threatening the safety of patients, burdening healthcare staff, and increasing healthcare costs. Efforts have been made to quantify the international burden of HAIs in particular with the best estimate of 7.6% and 10.1% pooled prevalence reported by the World Health Organisation (WHO) in 2011 in high-income and low- and middle-income countries respectively.(1) Various pathogens including those responsible for HAIs have also been shown to defy international borders. Even with robust surveillance systems and up-to-date prevention and infection control recommendations like that of France, spread within countries continues to occur. For example, in intensive-care units (ICUs) where the most vulnerable patients are being treated and the highest precautions are taken when it comes to preventing and controlling infection and colonisation, these pathogens can be highly prevalent. Novel infection control strategies are needed to combat the ongoing spread of pathogens in healthcare settings and multi-drug resistant organisms (MDROs) in particular. Patient transfers both across international borders and within countries across different healthcare facilities have been associated with the highly resistant Carbapenemase-producing *Enterobacteriaceae* (CPE). Patient transfers may also be the main driver of both the spatial dissemination of pathogens and may sustain these endemics. Therefore, it is essential to better understand the impact of intra- and inter-facility transfer of patients on pathogen spread.

The main objectives of this thesis are two-fold: 1) to better understand the structure of patient transfer networks, henceforth referred to as healthcare networks; in particular, the healthcare networks of France, for which we aimed at giving a first description in order to study the potential impact of the network structure on pathogen spread in terms of network topology; and 2) to examine the role of the French healthcare network on the dissemination of CPE over time, an HAI which remains untreatable with most currently available antibiotics.

In order to address these two main objectives, the mechanisms in which pathogens spread in the healthcare setting were explored. The first step of the thesis was to describe in general terms how mathematical and mechanistic models have helped explain the propagation of pathogens in different healthcare settings by conducting a systematic review. In addition, the review aimed to explore how the integration of network data in the different modelling studies may improve the understanding and predictive capacity of pathogen spread in healthcare settings.

In recent years, healthcare networks have provided insightful ways on improving the understanding of how common HAIs spread and how control efforts can be optimised. However, the differences between the structures of the healthcare networks based on different patient populations had not been previously explored. The second step of the thesis was to give a first description of the French healthcare networks and to compare the healthcare networks of the general patient population and HAI-diagnosed patient population in order to better understand the differences. This step also entailed understanding the heterogeneities present in the healthcare network structure and the potential impact they could have on the risk of common HAI pathogen spread.

Finally, we focused on CPE, which has become an important public health issue in France. The epidemic has been described in detail; however, forecasting trends have not yet been explored. In order to address the second aim of the thesis, we first described and attempted to predict the temporal trends of CPE in France. In order to better understand the impact of patient transfers on CPE spread, we assessed if the healthcare network could explain the observed incidence and spread of CPE in France during the 2012 to 2015 period. Lastly, we also used Bayesian methods to reconstruct CPE chains of transmission to further understand CPE spread dynamics.

This manuscript is organized into five parts:

- Part One covers the epidemiological context of common HAI pathogens and CPE in the international and French context (Chapters 1 and 2);
- Part Two (Chapters 3, 4, 5, and 6) introduces the major themes covered by the thesis (pathogen spread dynamics, mathematical modelling, and healthcare networks), and reviews work previously done in this field of research (first article);
- Part Three (Chapters 7 and 8) addresses the first thesis objective, by first describing the database used to construct the French healthcare networks and then analysing these networks (second article);
- Part Four (Chapters 9 and 10), addresses the second thesis objective, with the description and prediction of CPE spreading trends, the assessment of the contribution of healthcare networks to these trends (third article), and description of CPE chains of transmission; and finally,
- Part Five (Chapters 11, 12, and 13) covers the discussion of the main results and perspectives.

Part One: The epidemiology of hospital-acquired infections (HAIs)

Chapter 1. Epidemiological context of HAIs

Pathogens in the healthcare setting can often lead to hospital-acquired infections (HAI), also referred to as healthcare-associated infections (HCAI) or nosocomial infections (NI), which are defined as an infection acquired (usually by a patient, though HAIs are not exclusive to patients and can be acquired by healthcare workers (HCW)) in a healthcare facility in which the infection was not present before the time of admission and was developed after admission (usually at least 48 hours after admission) due to exposure to the healthcare facility environment.(1, 21) HAIs can be further classified into specific types of infections, for example urinary tract infections (UTI), surgical site infections (SSI), lower respiratory tract infections (LRI), and bloodstream infections (BSI).(22) Many types of microorganisms can be agents of infection though bacteria from the *Staphylococcus* genus, *Enterococcus* genus, and *Enterobacteriaceae* family are the most prevalent in the healthcare environment.(21)

Since the discovery of penicillin in 1928, antimicrobial agents or antibiotics have been the most common way to treat many bacterial infections in the healthcare setting. The mid-century development of several antimicrobials led to the cure of many once-serious infections and saved countless lives. However, the misuse and overuse of these antimicrobials over time has put a selective pressure on bacteria to develop resistance mechanisms to prevent antimicrobial action in order to survive. The selection of these mechanisms is encoded in genes that can be passed on vertically (through proliferation) or horizontally (across bacterial species usually through plasmids transferring genetic material between them).(23) In consequence, antimicrobial agents eventually led to a selection of bacteria that were resistant to one or more antimicrobials more rapidly than the rate at which they were being developed. Today, MDROs such as carbapenemase-producing *Enterobacteriaceae* (CPE), *Pseudomonas aeruginosa*, *Acinetobacter baumannii*, vancomycin-resistant *Enterococci* (VRE), and methicillin-resistant and vancomycin-resistant *Staphylococcus aureus* (MRSA, VRSA) are a few of a number of HAIs that were declared critical and high priority for research and development for new antimicrobials by the World Health Organisation (WHO).(24)

1.1 Public health importance

Despite significant progress in infection control, the last decades have seen a worldwide (re)emergence and spread of virulent infectious agents within healthcare settings, including viruses such as SARS, MERS-CoV and Ebola, as well as multi-drug resistant bacteria such as

MRSA or extended-spectrum β-lactamase producing *Enterobacteriaceae* (ESBL-E). Indeed, the conditions of the healthcare environment makes healthcare facilities breeding grounds for infection spread. Patients are susceptible to infections due to a number of risk factors. Being immunocompromised thus more vulnerable to infection, having had an invasive procedure resulting in an open wound, sharing a room with other potentially vulnerable or infectious patients, or being in frequent contact with HCWs with potential hand carriage of infectious agents, are examples of how the shared hospital environment can expose patients to pathogens. In addition, certain pathogens can remain viable in the environment for long periods of time and can be carried asymptomatically (without any clinical symptoms or immune response to an infection). Therefore, many factors can make microorganisms difficult to control in the healthcare setting. In consequence, it is essential to both improve the understanding of how pathogen spread in these settings and how to best prevent and control them in order to improve patient safety in the healthcare environment.

1.2 International context

1.2.1 Epidemiological burden

Pathogens leading to HAIs put patients at risk for complications which can lead to increased morbidity and mortality, as well as longer hospital stays; thus, posing both an epidemiological and financial burden on the healthcare system. Due to frequent asymptomatic carriage, inaccessibility to laboratories, disorganisation and lack of detail in medical records, lack of expertise, understaffing, and overcrowding, the prevalence of HAIs can be difficult to estimate – one of the challenges of preventing and controlling HAIs and MDROs pathogens.

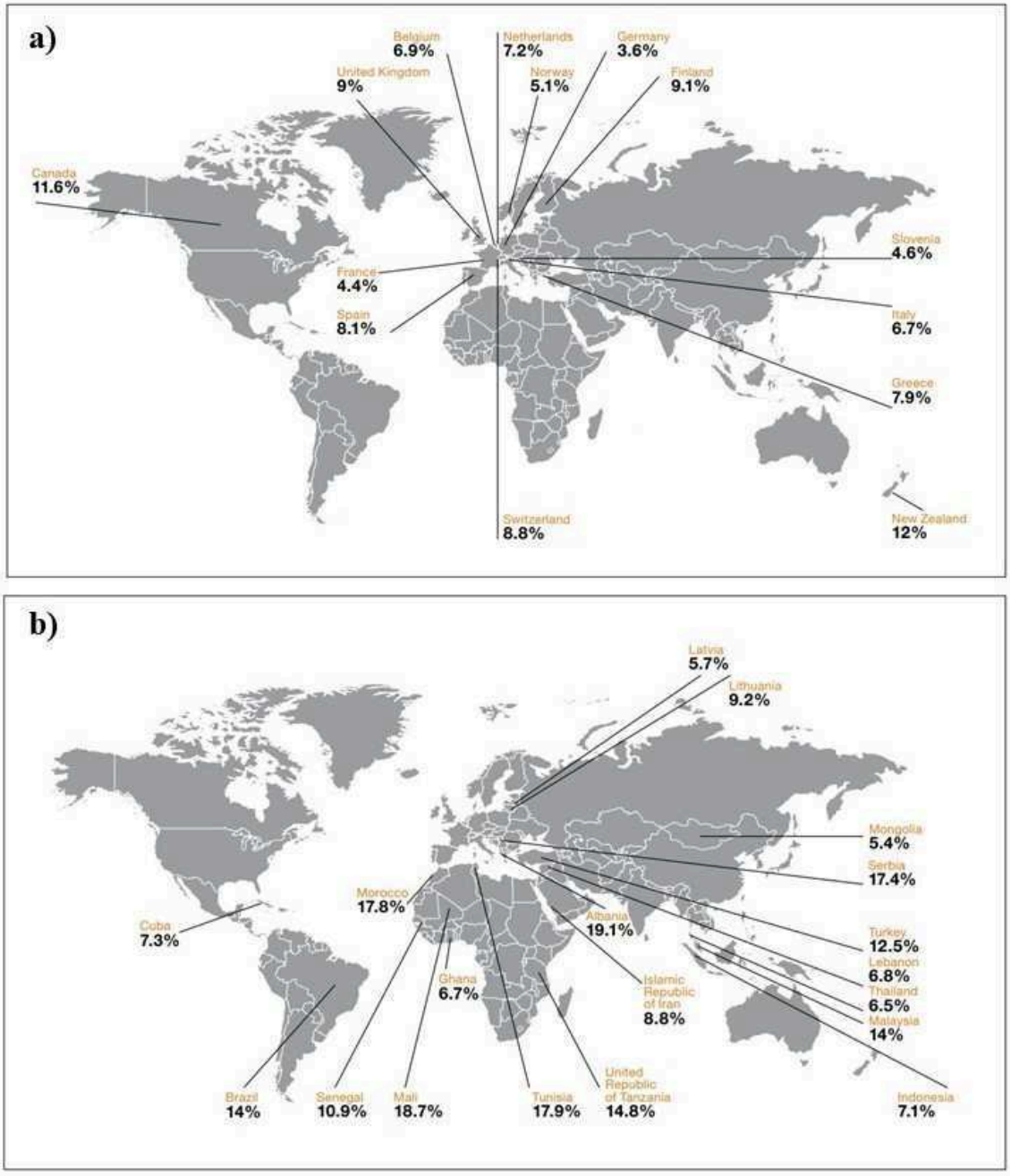
1.2.1.1 Incidence and prevalence

The epidemiological burden of HAIs in terms of pooled hospital-wide prevalence in mixed patient populations was estimated at 7.6% in high-income countries and at 10.1% in low- and middle-income countries by the WHO in 2011.(1, 2) In high-income countries, studies estimated between 3.5% to 12% prevalence of hospitalised patient who acquired at least one HAI (Figure 1a). In low- and middle-income countries, studies estimated between 5.4% to 19.1% prevalence of hospitalised patient who acquired at least one HAI (Figure 1b). The HAI incidence density ranged from 13.0 to 20.3 episodes per 1000 patient-days in high-income countries with a cumulative incidence of 17.0 episodes per 1000 patient-days among adult high-risk patients.(1) The incidence density in low- and middle-income settings was estimated to be

higher, ranging from 4.1 to 91.7 episodes per 1000 patient-days with a pooled cumulative incidence of 42.7 episodes per 1000 patient-days.(1)

The most recent national HAI estimates for the United States (US) (which was not included in the WHO report) estimated the incidence rate of HAIs at 4.5% in 2002 affecting 1.7 million patients and leading to 9.3 infections per 1000 patient-days.(1)

Figure 1. Prevalence of HAIs in high and low-and middle-income countries, 1995-2010



High-income country-level hospital-level prevalence estimates (a) and low- and middle-level country hospital-level prevalence estimates (b) from the most recent studies identified between 1995 and 2010 by the WHO.(1)

Acute-care settings respond to immediate and life-threatening health conditions and play an important role in prevention of death and disability.(25) These settings also have an important HAI prevalence and incidence.

In the US, an estimated 4.0% of inpatients in acute-care hospitals had at least one HAI in 2011 which corresponded to 648 000 patients with 721 800 HAIs.(26) A median of 6 days was reported between hospital admission and onset of symptoms in the same study.(26) A total of 25.6% of infections were associated with a medical device.(26)

At the European level, surveys conducted by the European Centre for Disease Control (ECDC) gave a more detailed description of HAIs in acute care settings. In a 2011-2012 point prevalence study conducted by the ECDC (2), the prevalence of patients with at least one HAI was estimated at 5.7% in the acute-care setting in Europe. Acute-care HAIs were most common in tertiary hospitals (7.4%), followed by specialised hospitals (6%), primary hospitals (5%), and secondary hospitals (5%). The incidence rate of HAIs in terms of patients acquiring at least one HAI per year was estimated at 3.5% in 2012 which represented an estimated 3.2 million patients per year in Europe. The point prevalence estimate for the same time period and population was estimated at 3.5 million HAIs per year. On any given day, the study estimated that over 87 000 patients had an HAI.

The burden of HAIs affects especially high-risk patients such as those admitted to intensive care units (ICU) in the acute-care setting, and new-borns.(1, 27) In high-income countries, an estimated 30% of patients in ICUs had at least one HAI episode and in low- and middle-income countries it was estimated to be 2 to 3 times higher.(1)

In Europe, ICUs account for 5% of the patient population but have 16.5% of the HAIs.(2) Almost 20% of the ICU patient population had at least one HAI, almost a four-fold increase compared to the average of other specialities or wards.(2) In a more recent survey by the ECDC based on 2015 data (28), it was estimated that after at least two days in the ICU 8.3% of patients contracted an ICU-HAI.

1.2.1.2 HAI pathogens

S. aureus and *Escherichia coli* species are the most commonly isolated HAI pathogens.(1) In Europe, *E. coli* was the most commonly isolated HAI pathogen (accounting for 15.9% of HAIs) followed by *S. aureus* species (accounting for 12.3% of HAIs).(2) In the US, *Clostridium difficile* was the most commonly reported pathogen in the acute-care setting accounting for 12.1% of HAIs followed by *S. aureus*, *Klebsiella pneumoniae*, and *E. coli*.(26) *C. difficile*

infections are also important HAI pathogens in Europe, accounting for 48% of the five most common HAIs: hospital-acquired (HA) pneumonia, SSI, UTI, BSI, and gastro-intestinal infections (GI).

Antimicrobial resistance is common among HAI pathogens. The best estimates in developing countries showed methicillin resistance in up to 54.5% of *S. aureus* isolated strains.(1) Among ICU-HAI isolates assessed by the ECDC, 23.1% were resistant to oxacillin among *S. aureus* species, 3.4% were resistant to vancomycin among *Enterococcus* species, 23.7% were resistant to ceftazidime among *P. aeruginosa*, and 20.2% were resistant to third-generation cephalosporins among three species (*E. coli*, *Klebsiella*, and *Enterobacter*). Resistance to carbapenems among five species ranged from 0.5% to 42.9%.(28)

1.2.1.3 Mortality

HAI can cause considerable morbidity but can also lead to death in the most vulnerable patients. The best world-wide estimates (excluding North America) have found a crude excess mortality in adult ICU patients at 18.5% for catheter-related (CR)-UTIs, 23.6% for CR-BSI, and 29.3% for ventilator-associated pneumonia (VAP).(1) In the United States, 99 000 deaths were attributed to HAIs in 2002.(1) More than 2.6 million new cases, 2.5 million disability-adjusted life years (DALYs), and 91 000 deaths were attributed to six types of HAIs (HA pneumonia, BSI, UTI, SSI, HA *C. difficile*, HA neonatal sepsis) each year in Europe.(29) Another Europe-level estimate found 37 000 annual deaths due directly to HAIs that led to an addition 110 000 deaths due to the infections.(1)

In developing countries, HAIs have been estimated to be responsible for 4% to 56% of all causes of deaths of hospital-born babies in the neonatal period.(1) The majority of these deaths occurred in South-East Asia and sub-Saharan Africa.

1.2.2 Impact of HAIs: length of stay and associated-costs

HAIs can lead to increased length of stay (LOS) for patients which could have consequences in terms of increased costs for hospitals. The estimate by the WHO for the increase in length of stay due to HAI in developing countries was between 5 to 29.5 days.(1) The impact of HAIs on patient LOS can vary depending the type of HAI infection. In Europe, the LOS was highest in HA-pneumonia and lowest for BSI.(3) CR-BSI were estimated to add an additional 4 to 14 days in LOS.(1) In 2008, HAIs were estimated to add 16 million extra days of hospital stay in Europe.(1)

An important impact of HAIs are the significant associated costs to the healthcare system. The financial costs associated with HAIs estimated by the ECDC in 2008 (3) found that HAIs were estimated to have a direct cost of 7 billion euros (based on the assumption that the daily costs to a hospital was 435 euros). Based on five major sites of HAIs, a study in 2009 estimated that the direct medical costs to US hospitals ranged from 28.4 to 45 billion US dollars depending on the adjustment to account for inflation.(30)

Costs have also been estimated for specific HAIs. VAP attributable costs have been estimated at 10 000 to 25 000 US dollars per case. Of approximately 250 000 CR-BSI occurring every year hospital-wide in the US, an annual cost was estimated up to 2.3 billion US dollars. A study on the cost associated with HAIs published in 2009 for Europe (31) estimated the total costs related to CR-BSIs for four European countries between 100 and 130 million euros in France, between 60 and 78 million euros in Germany, around 82 million euros in Italy, and between 29 and 54 million euros in the United Kingdom.

1.3 French context

1.3.1 The French HAI surveillance system

In 1992, the national technical committee of nosocomial infections (Comité Technique National des Infections Nosocomiales) pushed for the establishment of a national surveillance of HAIs in French healthcare facilities. As a result, in 1993 France developed a national alert system (Healthcare-Associated Infections Early Warning and Response System (HAI-EWRS) or in French, the Réseau d'Alerte, Investigation et de Surveillance des Infections Nosocomiales (RAISIN)) managed by the health surveillance institute (Institut de Veille Sanitaire (InVS)), now a part of Public Health France (Santé Publique France). The system was composed of five networks based on the following: surgical site infections (ISO-Raisin), MDROs (BMR-Raisin), adult ICU infections (REA-Raisin), antibiotic consumption (ATB-Raisin), HCW blood exposure accidents (AES-Raisin), and nosocomial bacteraemia (BN-Raisin). These networks have been coordinated by 17 regional centres for prevention of HAIs (Centre d'appui pour la Prévention des infections associées aux soins (CPias) previously known as Centre de Coordination de la Lutte contre les infections nosocomiales et associées aux soins (CClin) and Antennes régionales de lutte contre les infections nosocomiales et associées aux soins (Arlin)).

1.3.2 HAI burden in France

1.3.2.1 Incidence and prevalence

The ECDC point prevalence study of HAIs in acute-care settings estimated that in France 4.9% [95% CI 4.3%-5.6%] of patients had an HAI on any given day, a figure consistent with an earlier estimate of 4.4% made by WHO during the 1995-2010 period (1, 2). Out of 12 million discharges in France, the total number of patients to get an HAI was estimated at over 324 000 with an estimated 2.7% HAI incidence.(2)

France, in comparison to other European countries, had the highest number of single rooms (50% of beds are single room beds).(2) Single beds can be useful in isolating patients infected with the most dangerous bacterial strains. On the other hand, an estimated 9.3% of French patients had HAI diagnoses that were classified as rapidly fatal which was the highest estimate among participating European member states.(2) France also had the highest rate of UTIs (30.7% of infections of which 94.1% were microbiologically confirmed) and the highest percentage of *E. coli* HAIs reported (26.6%).(2)

1.3.2.2 Excess mortality, length of stay, and associated costs

In a 2002 prospective study conducted by CPias (CClin at the time), researchers found that between 26% and 30% of patients who died in the study had an HAI and, of these deaths, between 6% to 15% were attributable to the infection.(4) When extrapolating to all healthcare facilities in France, they estimated that between 7 000 to 20 0000 HAI attributable deaths occur in France per year.(4) Other studies estimated a higher HAI attributable death fraction at 21.3% overall, between 30% and 44% for ICU patients, a 6% death rate for CR-HAIs, and between 2.5% and 6% death rate among SSI.(4)

Overall, the total cost of avoidable HAIs in France has been estimated between 730 million and 1.8 billion euros per year.(5) Avoidable costs range between 500 euros for UTIs to 40 000 euros for the most severe ICU-BSI which were found essentially to be due to prolonged LOS.(5) Other estimates in the ICU setting found that HAIs cost over 36 000 euros and added four days to the patients' LOS on average.(5) In addition, antimicrobial resistant strains were found to cost more than sensitive strains (i.e. MRSA costing 30 000 euros versus methicillin-sensitive *S. aureus* costing 19 000 euros) and had an estimated 71% increase in LOS.(4)

1.3.2.3 Epidemiology of SSI

The SSI surveillance network, ISO-Raisin, established in 2001, has been surveying 11 types of surgical units (listed in Table 1). In 2015, the surveillance system surveyed a total of 106 737 interventions and estimated the incidence rates for each of these units (Table 1).(32) Orthopaedic surgery, the most surveyed intervention, had an estimated a 1.15% incidence rate among the surveyed units in which half of these SSIs were deep incisional SSIs.(32) For general surgery, the global SSI incidence rate was estimated at 1.74%. The highest rate of SSI in general surgery was for colorectal surgery (incidence rate of 6.82%) that included 45% of the interventions for colorectal cancer. SSIs for colorectal surgery were highest for general surgery during the entire 2011 to 2015 period. Obesity and diabetes were identified as significant risk factors for an SSI in general surgery. Among all surgical unit types, the highest SSI incidence rate was estimated for coronary surgeries at 4.18%. The SSI incidence density per 1000 follow-up days was estimated at 2.05 for this surgery type. The second highest SSI incidence rate was estimated for reconstructive surgeries (Table 1). Death rates of SSI patients were under 1% in all units in 2015 which was lower compared to a 2002 study that found the SSI associated mortality between 2.5% and 6%.(4, 32)

Table 1. Surgical site infection incidence rates in France by surgical unit speciality

Surgical unit type	**SSI incidence rate [95% CI]**	**Mortality**	**Units surveyed**	**Interventions**
Orthopaedic	1.15 % [1.02-1.27]	0.4%	276	29 293
General	1.74 % [1.59-1.90]	0.4%	254	29 178
Gynaecology-obstetrics	1.63 % [1.46 – 1.79]	<0.1%	229	23 102
Trauma	0.73 % [0.52-0.93]	<1%	77	6 607
Vascular	0.38 % [0.22-0.54]	<1%	79	5 558
Urology	2.76 % [2.32 – 3.19]	<1%	89	5 548
Neurosurgery	1.07 % [0.73-1.42]	<0.1%	43	3 454
Bariatric	1.47 % [0.95-1.99]	<0.1%	42	2 106
Coronary	4.18 % [3.00-5.36]	<1%	10	1 149
Thoracic	1.30 % [0.03-2.57]	<1%	10	540
Reconstructive	3.47 % [0.90-6.03]	0	5	202

1.3.2.4 Epidemiology of ICU-HAI

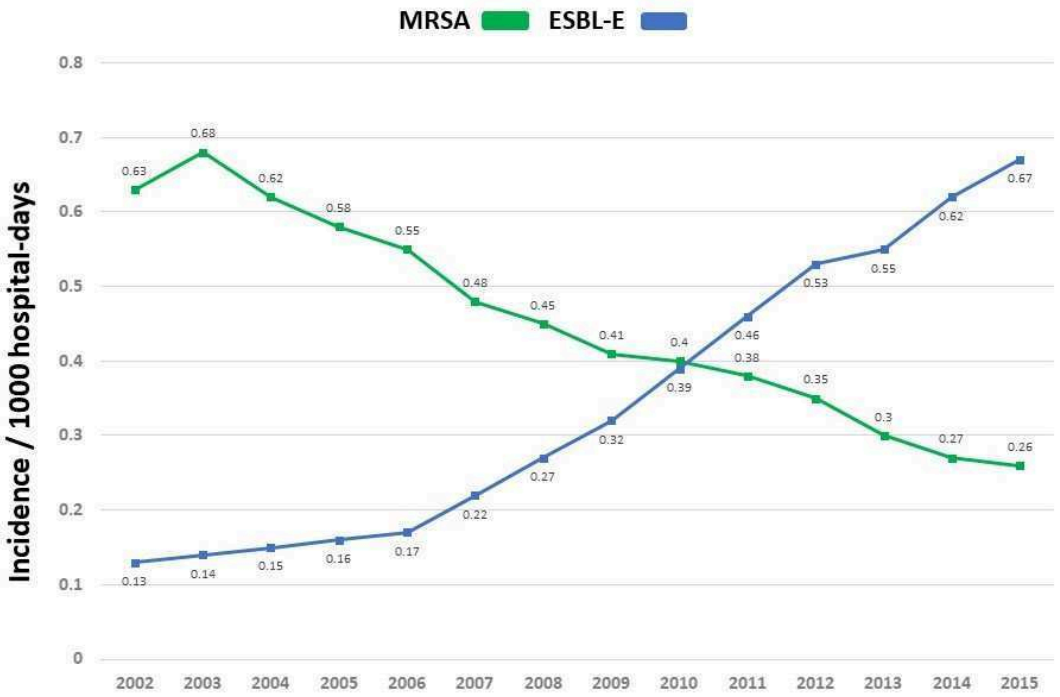
For adult ICU infections, the REA-Raisin also conducted a survey in 2015 of 188 ICUs of 63 240 patients.(33) In these ICUs, 15.8% of patients were immunocompromised, 55.9% of patients received an antibiotic at admission, and the ICU-mortality was 17.8%. The average LOS of patients in the ICU was 11 days. Invasive medical devices were very common among patients: 86.6% had a urinary catheter, 65.1% had a central venous catheter, and 63% had intubation. An infection was present in 10.62% of patients. The REA-Raisin report estimated that out of the isolated strains, 16% of the strains were MRSA and 17.8% were ESBL-E. *Pseudomonas aeruginosa* (15.1%), *S. aureus* (10.6%), *Staphylococcus epidermidis* (9.6%) were also commonly isolated strains. An estimated 15.05 VAP infections occurred per 1 000 intubation days, 3.52 BSI per 1 000 ICU-days, 0.66 CRSI per 1 000 catheter-days, and 0.55 CRBSI per 1 000 catheter-days.

1.3.2.5 Epidemiology of MDROs

The French national surveillance network of MDROs in healthcare facilities (BMR-Raisin) was established in 2001 and conducted annual surveys of MRSA and ESBL-E and the effectiveness of prevention measures on the spread of pathogens.(6) The latest report in 2015 surveyed over 1 400 facilities that covered 77% of the complete hospital stays for that year.(6) The majority of MRSA and ESBL-E were identified in the acute medical care setting, followed by surgical units, and long-term care settings (accounting for 39%, 23%, and 18% of MRSA isolates and 36%, 23% and 19% of ESBLE-E isolates respectively).

Since the first annual survey in 2002, the number of participating healthcare facilities increased significantly and has allowed for comparison of HAI rates over time in France.(6) The incidence rate of MRSA has decreased more than two-fold since 2002 (from 0.63 in 2002 to 0.26 in 2015) and especially in acute-care, long-term care and ICU settings (reduced by -60%, -63%, and -63% respectively) (Figure 2). On the other hand, the incidence rate for ESBL-E has increased over time from 0.13 in 2002 to 0.67 in 2015 (Figure 2). Four main species of ESBL-E (*E. coli*, *K. pneumoniae*, *Enterobacter cloacae*, and *Enterobacter aerogenes*), became more present over time. *E. coli* represented 56.5% of isolates and *K. pneumoniae* 26% of isolates in 2015. Among the same healthcare facilities surveyed regularly between 2011 and 2015, there was significant increase in the number of ESBL-E isolates over time; most notably, ESBL-producing *K. pneumoniae* increased by 76%.

Figure 2. MRSA and ESBL-E incidence rate per 1 000 hospital-days, France 2002-2015



Taking into account the incidence rates, number of hospital stays, and hospital beds, the report estimated that between 23 000 to 39 000 cases of MRSA occurred annually in which 3 800 to 5 700 were of bacteraemia cases.(6) Concerning ESBL-E, between 49 000 to 103 000 cases were estimated to occur annually. The overall trends and rates were consistent with other reports.(34)

In the ICU, 9% of incident cases had MRSA and 11% had ESBL-E. The average percentage of methicillin resistance among isolated *S. aureus* was 16.4%. There were far more ESBL-E UTI isolates than MRSA UTIs (77% versus 21%). The incidence rate of ESBL-E was higher in the ICU than MRSA (0.93 incidence of MRSA per 1000 hospital-days versus 2.72 incidence of ESBL-E per 1 000 hospital days) (Figure 3).

Figure 3. MRSA and ESBL-E incidence rate per 1 000 hospital-days by setting, France 2015

Chapter 2. Carbapenemase-producing *Enterobacteriaceae* (CPE)

2.1 *Enterobacteriaceae*

Enterobacteriaceae are a family of Gram-negative bacteria that are among the most important human pathogens and are responsible for a broad range of infections in humans. Although they are frequent asymptomatic colonizers of the gastrointestinal and oropharyngeal tract of humans and animals, they also are one of the leading causes of HAIs. *E. coli*, *K. pneumonia*e, *Enterobacter* species, and *K. oxytoca* are some of the most common HAIs and accounted for 20% of HAI isolated in 2006 to 2007 in the US.(35) They have been implicated in HAI-UTIs, SSIs, BSIs, and VAP infections. Although there has been a report that the overall proportion of HAI due to *Enterobacteriaceae* declining over time in the US (35, 36); the prevalence of antimicrobial resistance has increased among bacteria of this family, causing major concern in the healthcare settings all over the world.

2.2 Beta-lactam antimicrobials and resistance

Beta-lactam antimicrobials are chemicals characterized by a beta-lactam ring within their structure. These antimicrobials are considered to be broad-spectrum antimicrobials, meaning that they can be used against a large variety of bacterial species. Beta-lactams, used to treat patients with various infections (i.e. skin infections, dental infections, lower respiratory tract infections, and urinary infections), can be classified into five main antimicrobial groups: penicillins, cephalosporins, monobactams, carbapenems, and beta-lactamase inhibitors.(37, 38) The first use of penicillin to treat patients was in 1941, 13 years after its initial discovery.

Beta-lactamases, which are enzymes produced by resistant bacteria that inhibit beta-lactam activity, are the most important mechanism of resistance to beta-lactam antimicrobials. Beta-lactamases can be grouped into four Ambler classes which share structural similarities: class A, B, C, and D.(39) Their main mechanism of action is to inhibit the peptidase domain of penicillin-binding proteins (PBPs) required for cell wall synthesis which leads to autolysis and cell bursting.(40)

2.2.1 Carbapenem antimicrobials and the emergence of CPE

Carbapenems are a class of beta-lactam antimicrobials that are effective against both gram positive and gram negative bacteria.(41) Thienamycin, or thienpenem, was the first naturally occurring carbapenem antimicrobial to be discovered in 1976 and to be described in detail in 1979.(40, 42, 43) Thienpenem is both potent and has broad-spectrum activity; however, it is

chemically unstable.(40) In addition to other carbapenems, imipenem, a derivative of thienpenem, was found to be both more stable and effective against gram positive and gram negative bacteria such as the *Enterobacteriaceae*.(40, 41) Imipenem was released in 1985 as the first carbapenem to be used to treat infections.(40) Over 80 compounds of carbapenems have been developed.(40) Combining carbapenems with other antimicrobials has become a common method to treat multi-drug resistant bacteria.(40)

Resistance to carbapenems is an important public health issue because carbapenems were often considered to be the treatment of "last-resort" for infections when other antimicrobials were not successful.(3, 44) Carbapenemases have the ability to hydrolyse carbapenem antimicrobials and some are referred to as a specific type of beta-lactamase – the metallo-beta-lactamases.(40) In order for carbapenem development to be successful against these emerging resistant strains, new antimicrobials must be able to overcome both beta-lactamases and metallo-beta-lactamases.(40) Combination therapy, colistin, and tigecycline are some of the few remaining options for treating patients with CPE. Unfortunately, the development of antimicrobials active against CPE has not yet been achieved.(45-47)

With the exception of one enzyme, all carbapenemases are of class A, B, and D beta-lactamases.(8) The mechanism of resistance to carbapenems (carbapenemase enzymes), the year of their identification, encoding gene location, country of first isolation, and class are shown in Table 2.(48, 49)

Table 2. Isolation of some of the first described carbapenemase enzymes

Enzyme	Class	Encoding	Country	Year	Ref
OXA-1	D	Plasmid	Japan	1967/1981*	(50)
SME-1	A	Chromosome	UK	1982	(51, 52)
IMP-1	B	Plasmid	Japan	1983	(53, 54)
IMI-1	A	Chromosome	USA	1984	(55)
ARI-1/OXA-23	D	Plasmid	Scotland	1985	(56, 57)
NMC-A	A	Chromosome	France	1990	(58)
SME-2	A	Chromosome	USA	1992	(59)
VIM-2	B	Plasmid	France	1996	(60)
VIM-1	B	Plasmid	Italy	1997	(61)
KPC-1	A	Plasmid	USA	1997	(62)
IMP-2	B	Plasmid	Italy	1997	(63)
GES-1	A	Plasmid	French Guiana	1998	(64)
KPC-2	A	Plasmid	USA	1998	(65)
SPM-1	B	Plasmid	Brazil	1999	(66)

GES/IBC-1	A	Plasmid	Greece	1999	(67)
VIM-3	B	Plasmid	Taiwan	1999	(68)
GES-2	A	Plasmid	South Africa	2000	(69)
KPC-3	A	Plasmid	USA	2000	(70)
OXA-48	D	Plasmid	Turkey	2001	(71)
GIM-1	B	Plasmid	Germany	2002	(72)
SIM-1	B	Plasmid	South Korea	2003	(73)
KPC-4	A	Plasmid	Scotland	2003	(74)
SME-3	A	Chromosome	USA	2004	(75)
NDM-1	B	Plasmid	India/Sweden**	2008	(76)
OXA-181	D	Plasmid	India	2010	(77)

* Yamamoto et al. described the *oxa* gene in a paper published in 1981 using bacteria strains originally isolated from Egowa et al. published in 1967 (78)
** bacteria strain isolated and examined in Sweden, patient most likely acquired infection in India
GES: Guiana extended spectrum
GIM: German imipenemase
IBC: integron-borne cephalosporinase; now revised to GES type (49, 79)
IMI: imipenem-hydrolyzing beta-lactamase
IMP: active on imipenem
KPC: *Klebsiella pneumoniae* carbapenemase
NDM: New Delhi metallo-beta-lactamase
NMC-A: not metallo-enzyme carbapenemase
OXA: oxacillin-hydrolyzing
SIM: Seoul imipenemase
SME: *Serratia marcescens* enzyme
SPM: Sao Paulo metallo-beta-lactamase
VIM: Verona integron-encoded metallo-beta-lactamase

Carbapenemase enzymes were first detected in gram positive bacilli and then gram negative bacilli such as *P. aeruginosa* and *A. baumannii.*(40, 49) Carbapenemase activity in *Enterobacteriaceae* was reported in the 1980s and became a concern in the 1990s when it was identified in *K. pneumoniae*.(80) Initially carbapenemases were considered an issue of clonal spread until transmission of carbapenemase genes were identified between different species.(49, 81) A combination of the production of beta-lactamases, efflux pumps, and mutations that alter porins and PBP have led to high levels of resistance in species such as *K. pneumoniae* and *A. baumannii.*(40)

2.3 Risk factors for spread

In 2010, the ECDC conducted a risk assessment on the spread of CPE through patient transfer in healthcare facilities. There was strong evidence to support that transferred patients colonised or infected with CPE across international borders increased the risk of introduction of CPE in the healthcare facilities in another country.(82) Consumption of carbapenems, third and fourth generation cephalosporins, and fluoroquinolones were identified as risk factors for CPE

colonisation or infection. The ECDC urged for more prudent use of antimicrobials. Although the assessment found that there was little evidence on the effectiveness of control measures, they argued that measures targeting CPE, if implemented, should have similar effectiveness as other measures for MDROs. European-level data on CPE in long-term care settings have been limited compared to the many studies conducted in acute-care settings.(82)

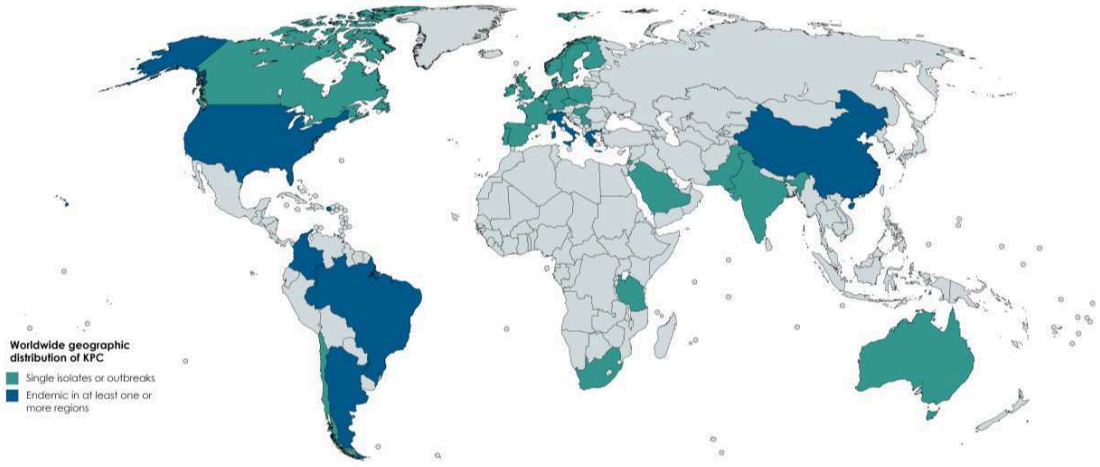
2.4 International context of the CPE epidemic

CPE prevalence has increased rapidly in the last two decades worldwide.(82) Horizontal gene transfer through mobile genetic elements such as plasmids and transposons, moving populations and patient transfers within and across borders have contributed to the spread of many classes of carbapenemase-producing genes, if not all classes, in the United States, Greece, the Mediterranean and European regions, and India.(82)

2.4.1 Worldwide reports of CPE

The first description of a carbapenemase-producing bacteria isolated from the *Enterobacteriaceae* family was reported in 1993 in France.(8, 58) Isolation of the first carbapenemase-producing *K. pneumoniae* (KPC) was reported in the United States soon after in 1997. KPC was reported across the United States, in South America, in Europe, and several provinces in China among other countries (Figure 4).(8) These KPC-producers have been isolated from patients in hospital settings in most cases; though rare, they have also been identified in the community. One clone of *K. pneumoniae*, sequence type-258, was identified worldwide.(8)

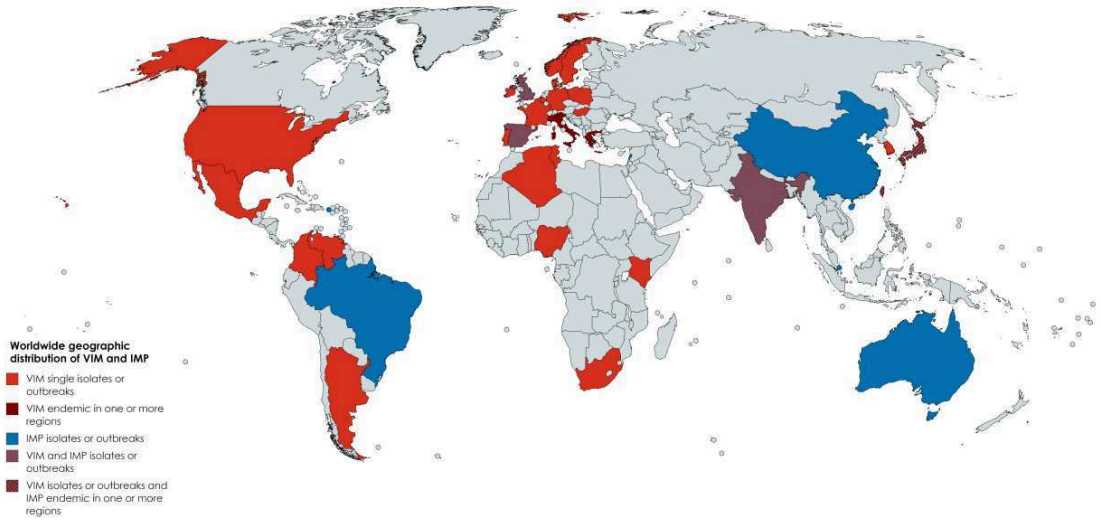
Figure 4. Worldwide geographic distribution of KPC



Retrieved from studies published using data up to 2016 (8, 83, 84)

For Ambler class A producing enzymes, VIM and IMP producers have also been described worldwide (Figure 5). There has been high prevalence of IMP producers in Japan and Taiwan while Greece had one of the highest prevalence of VIM in 2011. Outbreaks have been reported for both CPE types in many countries.

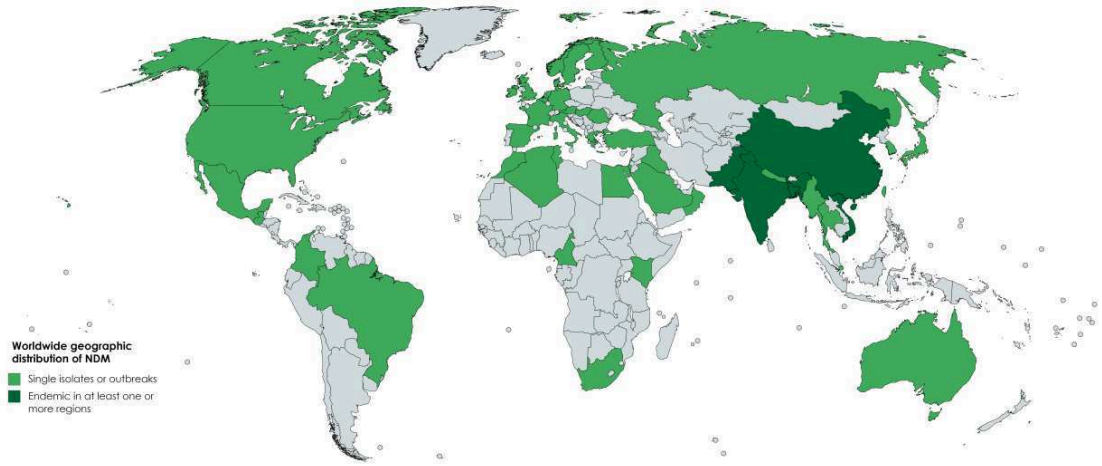
Figure 5. Worldwide geographic distribution of VIM and IMP



Retrieved from studies published using data up to 2011 (8)

Within four years of its' first identification, NDM-1 producing *Enterobacteriaceae* have been observed in high numbers in India, Bangladesh and Pakistan where it originated and also in the Middle East and Europe (Figure 6). The main challenges that are faced with controlling NDM-1 pertain to the beta-lactamase gene $bla_{\text{NDM-1}}$ being expressed in many unrelated species, its spread in the environment, its acquisition in two commonly transmitted pathogens *K. pneumoniae* and *E.coli*, and the presence of a potentially large human population reservoir in India.(8)

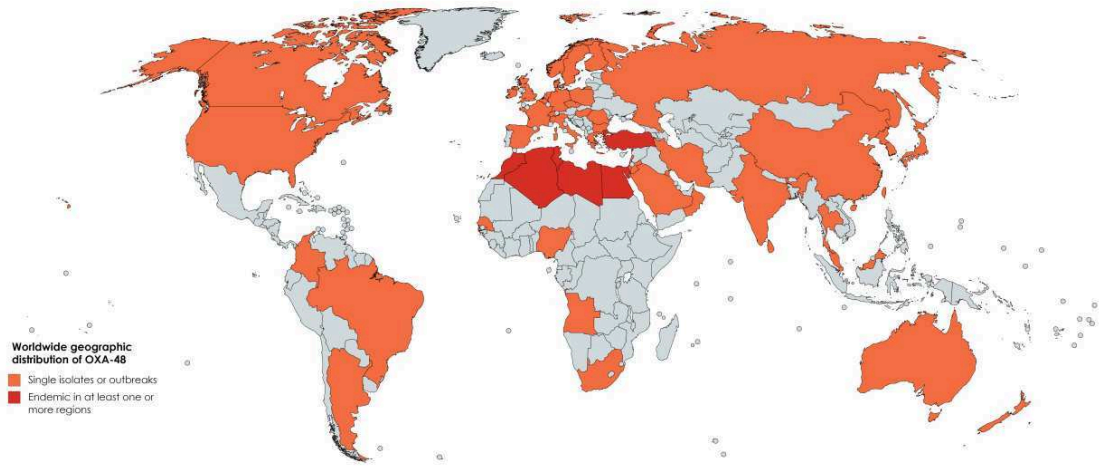
Figure 6. Worldwide geographic distribution of NDM



Retrieved from studies published using data up to 2014 (8, 85)

The true prevalence of OXA-48, one of the most commonly identified OXA producers, may actually be underestimated due to the difficulty of identifying this type of carbapenemase producer.(8) This Ambler class D enzyme was first identified in Turkey in 2001 and spread throughout the Middle East, North Africa, in particular Morocco, and Europe (Figure 7).(71, 86, 87) A study has shown that spread may not have been due to one single *K. pneumoniae* clone but many.(71) Nordmann et al. argue that two types of CPE epidemics may be occurring worldwide: a nosocomial CPE epidemic mainly through *K. pneumoniae* bacteria for all mechanism types and one of community spread via *E. coli* bacteria encoding OXA-48 and NDM genes.(8)

Figure 7. Worldwide geographic distribution of OXA-48



Retrieved from studies published using data up to 2017 (8, 88)

In 2016 the ECDC published a rapid risk assessment based on CPE 2014 data from the European Antimicrobials Surveillance Network (EARS-Net) to describe collected isolates in Europe.(89) Less than one percent *of E. coli* isolates were CPE in all participating European countries while in Greece more than half of *K. pneumoniae* isolates were KPC. For the 2014-2015 period, four countries were identified as having an endemic CPE situation (stage 5): Greece (since 2010), Italy (since 2013), Malta (since 2013), and Turkey. Nine countries were classified at stage 4 with inter-regional spread of CPE.

2.4.2 Incidence and prevalence

The prevalence of CPE varies depending on the region and country. In New York where KPC emerged and became highly prevalent, carbapenemase-resistance among *K. pneumonia* was reported as high as 36% in 2006 but then decreased to 13% in 2013-2014. Although data in the US regarding KPC is rare, there have been estimates of 2.94 annual incident KPC-producing CRE cases per 100 000 population.(83) In a study conducted in seven Central and South American countries, out of 21% of patients with CPE, 38% were identified with a KPC-producing gene. The attributable death rate of a KPC infection have been reported as high as 50%.(8)

In 2016, EARS-Net reported on antimicrobial resistance in Europe.(90) All countries had less than 0.1% resistance to carbapenems among *E. coli* isolates with the exception of Romania with 1% resistance. The prevalence of CPE resistance among *K. pneumoniae* isolates was estimated at 6.1% with the highest prevalence in Greece at 67% followed by Italy at 34% and Romania at 31%. Among *P. aeruginosa* isolates, 15% were carbapenem resistant and among *Acinetobacter* species, 35% of isolates were carbapenem resistant. In European ICUs, carbapenem resistance was reported in 11% of *Klebsiella* species isolates, 24% of *P. aeruginosa* isolates and 69% of *A. baumannii* isolates. Another European-level study found that based on population-weighted averages, 1.3 patients per 10 000 hospital admissions and 2.5 patients per 100 000 hospital patient-days had a carbapenemase-producing *K. pneumoniae* or *E. coli*. The highest incidence was found in Greece, Italy, Montenegro, Spain, and Serbia.(91) In a French study, the incidence density rate of CPE was estimated at 0.0041 per 1000 hospital-days and 0.0027 per 100 admissions.(92)

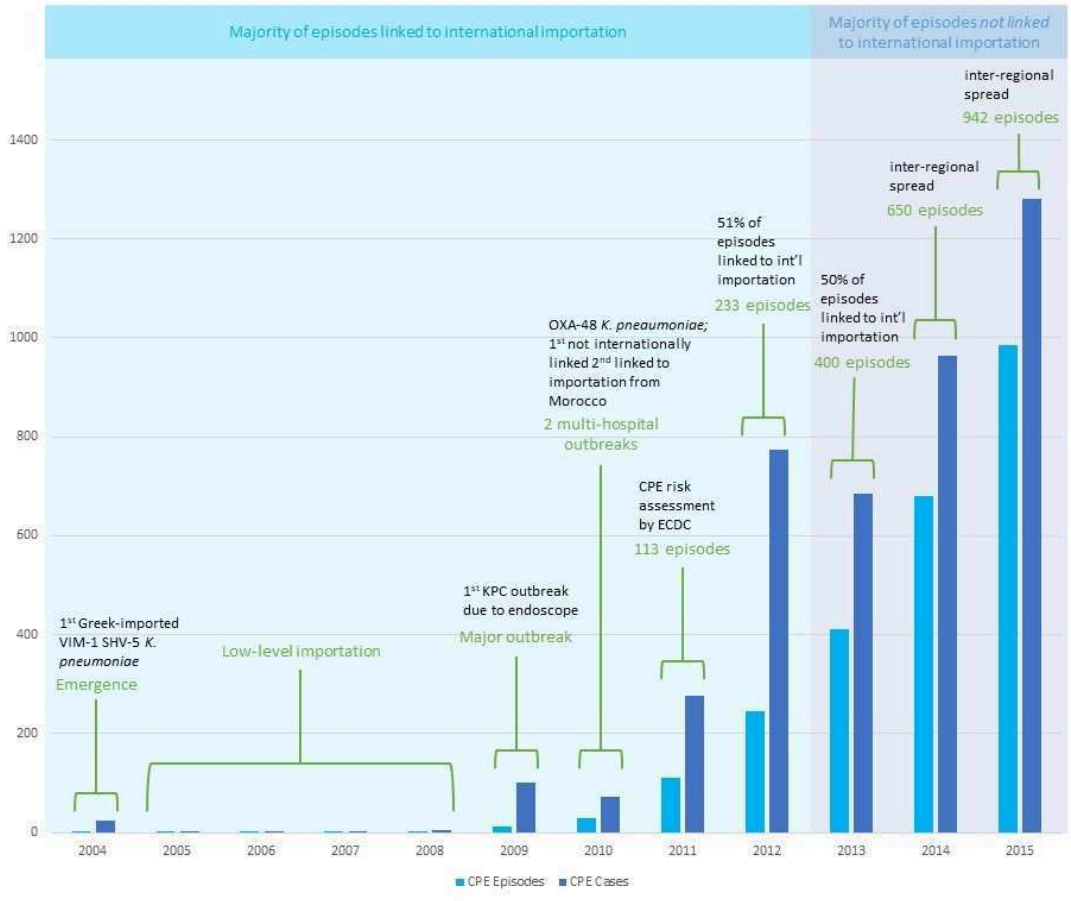
2.5 French context of the CPE epidemic

Mandatory notifications of all suspected isolates of CPE, both infections and colonisations, are reported to regional health agencies (Agence National de Santé (ARS) and CPias) in France.(93)

The mechanisms of resistance are verified in local or expert laboratories or transferred to the national reference centres (Centre National de Référence (CNR)) for confirmation.(9) Public Health France publishes reports on the epidemiological situation of CPE every semester.(93)

The first identified case of CPE in France was in 2004 concerning a patient admitted in December 2003 to a French hospital ICU who had been transferred from Greece (Figure 8).(94) The patient was colonized with *K. pneumoniae* producing VIM-1 and SHV-5 and transmitted the strains to seven other patients in the surgical ward and ICU from February to August of 2004.(9, 94) Despite transmission to other patients, the outbreak was controlled quickly according to the authors who first described the study.(95) Another seven episodes (defined as grouped signals of one or more cases within the same chain of transmission determined by an epidemiological investigation by surveillance authorities) were identified between 2004 and 2008.(96, 97)

Figure 8. CPE epidemic in France, 2004-2015



An important outbreak occurred in 2009 which was the first reported KPC-producing *K. pneumoniae* outbreak due to a contaminated endoscope (Figure 8).(9, 98) Over 340 contacts

were screened in five different healthcare facilities.(9) Two large OXA-48-producing *K. pneumoniae* outbreaks were reported in 2010 in which one did not have a link to any cross-border transfer and the other was linked to importation from Morocco.(99, 100) The first outbreak led to the screening of contacts in 15 healthcare facilities and the second outbreak implicated 32 cases in which contacts were screened in 10 facilities.(9)

Among the first 29 episodes reported between 2004 and 2010 in France, 18 of them implicated either colonisation or infection of *K. pneumoniae*, followed by four *E. cloacae* strains, three *E. coli* strains, two *E. aerogenes* strain, one *Citrobacter freundii* strain, and one *Proteus mirabilis* strain.(97) The KPC mechanism of resistance was most common (n=12), followed by OXA-48 (n=7), VIM (n=5, one of two mechanisms identified in one strain), NDM-1 (n=4), IMP (n=1, one of two mechanisms identified in one strain), and one was not identified.(97) Out of the 29 episodes, 24 were linked to a transferred patient who had been previously discharged from a foreign country. Eleven of the episodes had an index case who was transferred from Greece followed by India (n=4) and Morocco (n=3).(97) Twenty episodes occurred in the Ile-de-France and the northern regions and six occurred in the south eastern regions of France.

Due to the rising occurrence of cross-border transfer of various carbapenemase-producing strains in France, including the first report of intercontinental spread of KPC from the United States to France in 2005 (101), a KPC-2 and SHV-12-producing *K. pneumoniae* strain from Greece (102), OXA-48 transferred from Turkey (87), and the consequent growing number of secondary cases, France voiced its concerns of CPE spread in 2010 at the European Centre for Disease Prevention and Control (ECDC) Advisory Forum.(82) The growing concern lead to a European-wide risk assessment of CPE by the ECDC in 2011 and again more recently in 2016.(89)

In 2011, France observed another sharp rise in the number of CPE episodes totalling to 113 episodes.(96) In the following years, the number of episodes continued to rise: 233 in 2012, 400 in 2013, 650 in 2014, and 942 in 2015 (Figure 8). Seasonality trends were observed starting in 2013 where the number of episodes peaked during August, September, and October.

The CPE epidemic from 2004 to 2014 affected almost 4 000 patients during which 48% of episodes occurred in the Ile-de-France region, 12% in the Provence-Alpes-Cote d'Azur region and less than 10% in each remaining region.(96) Eighty percent of patients were colonised and 20% were infected with a CPE strain. The number of cases per reported episode reduced over time; however, four episodes affecting over 100 cases were followed over the years and were

still active in 2015. Since 2013, most infection sites of CPE have been UTI (46%), BSI (24%), or pneumonia (14%). Most colonisations were identified in the gastrointestinal tract (77%) or urinary tract (26%). From the 2004-2010 to 2011-2015 period, *K. pneumoniae* were the most commonly isolated strain of CPE representing 58% of all isolated strains in both periods. *E. coli* were the second most commonly isolated bacteria representing 36% of isolated strains. In addition, OXA-48 and OXA-48-like CPE represented 71% of the identified mechanisms of carbapenem resistance. Concerning episodes not linked to international importation, 82% have been reported with OXA-48-like mechanism of resistance.(96) There was also an important diminution of the number of episodes linked to foreign importation over time: from 80% in 2009, 71% in 2010, 51% in 2012, 50% in 2013, 47% in 2014, to 42% in 2015.(96) Morocco, Algeria, Tunisia, India and Egypt have been the top five countries were cross-border importation of CPE has been linked in France.

The European-wide risk assessment of CPE by the ECDC in 2016 showed that like other European countries France has updated their national guidelines (103) and strategies concerning CPE; however from 2013 to the 2014-2015 period France advanced from an epidemiological stage 3 of regional spread to a stage 4 inter-regional spread of CPE and may risk reaching a CPE endemic situation (stage 5).(89)

2.5.1 Prevalence and mortality

In the 2016 EARS-Net French data, less than 0.1% of *E. coli* isolates were resistant to carbapenemases and less than 0.4% of *K. pneumoniae* were resistant to carbapenems. Among 1 968 isolates of *P. aeruginosa* 15.6% were carbapenem resistant and among 450 isolates of *Acinetobacter* species 7.1% were resistant.(90) The percentage of carbapenemase resistance among *Enterobacteriaceae* with decreased susceptibility to carbapenems increased from 23.1% in 2012 to 28.6% in 2013 and 36.2% in 2014.(104)

In the 2004 to 2010 period, the 29 CPE episodes in France resulted in a 30% mortality rate.(97) Reports of rates of mortality between 30% to 70% were found for CPE bacteraemia.(10) Out of the CPE episodes occurring in La Reunion (a French overseas island) between 2010 and 2015, a study found an 89% mortality rate among those colonised and an 11% mortality rate among those infected.(105)

Part Two: Spread of pathogens in healthcare settings

Chapter 3. Understanding pathogen spread dynamics

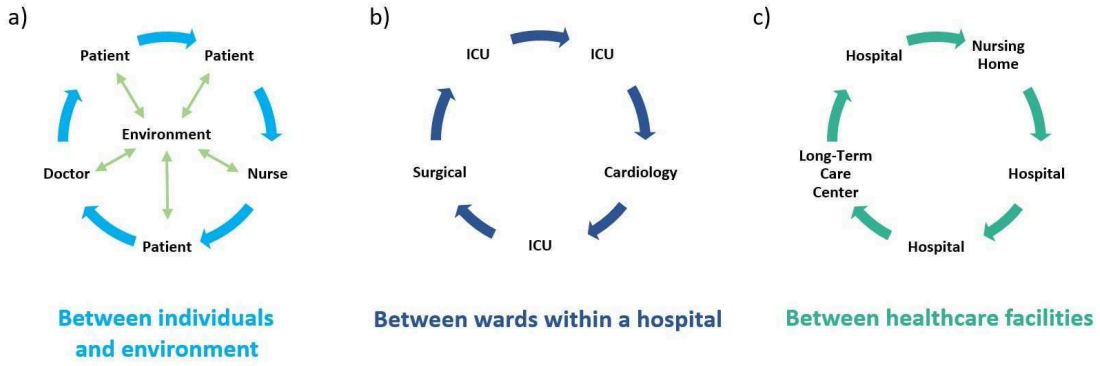
3.1 Mechanisms of spread in healthcare settings

There are several ways in which pathogens can spread within healthcare settings, through various routes of contact with colonised or infected individuals or with the hospital environment.(106) Reservoirs of potential HAI pathogens in the healthcare setting include the patients and HCWs (usually in the nasal cavity, skin, or gastrointestinal tract) and contaminated objects (fomites such as a sink, bed, ventilator machines, intravenous devices, etc.).(35) In particular, bacteria (which are responsible for a large part of HAIs) may colonize individuals in their gut or skin without any clinical symptoms, and survive on the surfaces of objects, such as hospital beds or surgical equipment, going undetected for long periods of time. Humans are natural hosts to many bacteria that may lead to infection; however, this depends on various host factors that can lead to the development and the severity of the infection.(35)

The three main modes of pathogen transmission in the healthcare environment are direct, indirect, and airborne transmission. Mayhall and colleagues (35) described them as follows: direct transmission follows direct skin contact that can quickly become infectious if a wound becomes contaminated; most often, in acute care hospitals, direct transmission occurs during patient contacts with the contaminated hands and clothing of HCWs. Airborne transmission is a type of indirect transmission route when infected droplet nuclei are released into the air and breathed in by other individuals. Finally, indirect transmission occurs when a patient's body is in contact with contaminated objects such as contaminated surfaces, food, biological fluids, or medical devices.

Considering the healthcare system as a whole, pathogen spread may occur at several levels: within hospital wards, at the scale of an entire hospital, or even between different healthcare institutions. Figure 9 illustrates three possible pathways that can lead to pathogen spread:

Figure 9. Examples of how pathogens may spread in the healthcare system



In Figure 9, a) is an example of individual-to-individual contact structure with interactions with the environment, b) represents an individual-to-individual contact structure where the contact depend on the ward (i.e. contact rates can be homogenised within each ward among individuals in models), and c) shows a healthcare facility-level structure where we are no longer concerned with individual or ward-level contacts but rather patient transfers that can link the various connected facilities. All three structures, however, form networks in which the components represent individuals, wards, or healthcare facilities, and their links are defined by contact, admission/discharge, or transfer rates.

For intra-hospital movement, the term "contact" can have many different definitions and some of the specificities can include direct physical touch or indirect touch, sharing of a hospital room or nurse, the physical distance between individuals, and duration of said contact. Some studies have gone beyond estimating an average contact rate among individuals to the use of wearable devices to measure interactions between individuals that can allow for identification of "super-spreaders" which are individuals who have a higher number of contacts than average and connect many individuals.(107) Another example is a large-scale study conducted to better understand the social contacts and mixing patterns of individuals on the potential spread of respiratory infections within different groups and social settings using diaries to assess the direct contact rates.(108) A recent study described in detail the contact patterns of patients and HCWs in terms of frequency and duration and their potential implications on infection risk in a long-term care setting.(109) Generally, the definition of a contact varies due to the complexity encompassing the variety of ways bacteria can spread within the hospital environment; however, experts in the field have attempted to collect rich data to better understand the heterogeneities.

Finally, the mobility of patients and HCWs plays an important role in the dynamics of spread of pathogens. These movements can be divided into two main channels: intra- and inter-hospital movement. HCWs move around treating different patients, in different rooms, and also in different hospital wards. In addition, patients can move between rooms, hospitals, and between the healthcare facility and the community. Despite geographic distance, hospitals with shared patients moving between them can play an important role in the sharing of bacterial strains and thus could lead to HAI pathogen spread.(110) For example, a study showed that patient sharing was significantly associated with the genetic similarity between MRSA isolates.(111) Therefore, both contact and transfer network structure can help better understand pathogen spread dynamics in healthcare settings.

3.2 Role of patient transfers in pathogen spread

Understanding pathogen spread at the health facility-level based on the patient transfer network structure – hereinafter referred to as a "healthcare" network – relies on transmission dynamics occurring in the other two network structures (Figure 9a and Figure 9b) and vice-versa. As a result, the healthcare network structure can connect geographically distant healthcare, patient populations, and can potentially lead to the sharing of common HAI pathogens. Authors Eveillard and colleagues published one of the first papers in 2001 (11) showing evidence for an association between transferred patients from hospitals and HAI prevalence. The authors found that transferred patients were five times more likely to be infected compared to those not transferred. This is due to the fact that patients can have a certain probability of acquiring a bacterial strain in one hospital facility that they can potentially carry (commonly referred to as "carriage") to another facility, leading to contamination of the environment of other individuals at a given probability.

As medical records became more commonly stored in electronic form, patient transfers became easier to document and follow. In consequence, researchers have been able to exploit these data sources by using algorithms to more easily and quickly retrace patient transfers and construct transfer networks. How patient transfer patterns and healthcare network structure have been used to explain the transmission of pathogens in healthcare systems is discussed in more detail in the next chapters (Chapter 4 and 5).

3.3 The role of the surveillance system and infection control

Once the impact of the healthcare network structure on infection spread is better understood, novel recommendations concerning infection prevention and control practices at all three levels

(Figure 9) can be developed. In order for these levels to work harmoniously together, cooperation and communication are vital since the levels are both inter and intra-connected. In this respect, managers of the surveillance system play a critical role in monitoring and implementing infection prevention and control measures at the various levels. In particular, communication among different healthcare facilities is required, though not the most feasible of tasks, in order to tackle any pathogen transmission risk concerning patient transfers. For instance, a working group recently updated the French guidelines on infection control regarding extensively-drug resistant bacteria (eXDR), including CPE.(103) They proposed to target for screening patients who were direct hospital transfers from a foreign hospital, patients with a history of hospitalisation in a foreign country, patients with a history of eXDR carriage and patients who had contact with eXDR patients. The definition of contacts, the difficulty of screening and microbiological diagnostics to identify all CPE types, prevention measures, and detection of risk of eXDR transmission were also elaborated by the working group. Although the French recommendations are comprehensive, there is little cooperation between countries to manage eXDR.(103) One major challenge faced by France and other nations concerns the feasibility of implementing prevention and infection control recommendations. Isolation measures may vary in practice between different healthcare facilities. Unavailability of single rooms, insufficient staff-to-patient ratios, and financial and managerial constraints may pose a challenge for many facilities leading to poor infection control.(103)

In conjunction with understanding the impact of the healthcare network on pathogen spread dynamics, the thesis aims to propose changes to how the surveillance system functions when responding to infection risk in order to take into account novel information that may not have been previously considered in the system's protocol.

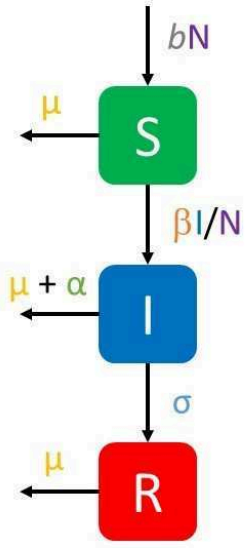
Chapter 4. Mathematical modelling of pathogen spread

4.1 An emerging discipline

In recent years, mathematical modelling has become an important tool to analyse the structure of intra- and inter-healthcare facility movement of patients and HCWs leading to new ideas on how this structure can impact the transmission of infections. Mathematical modelling has played an important role in improving the understanding of the propagation of pathogens in healthcare settings over time and how control efforts could be optimized.(112-115)

Modelling the complexity of interactions between an infectious agent and its' host, the interactions between patients and HCWs, and the dynamic structure of the healthcare system has been essential in better understanding the epidemiology of common HAI pathogens. A compartmental model of an SIR (Susceptible-Infected-Recovered) framework is shown in Figure 10 to explain the structure of a basic model. The total population, *N*, is composed of *S* number of susceptibles, *I* number of infected, and *R* number of recovered individuals. The parameters that determine the flows between the compartments include the transmission rate β, the birth rate *b*, the death rate μ, the disease-induced death rate α, and the recovery rate σ. The differential equations describing these transitions can also be described (Figure 10).

Figure 10. SIR model framework



Variables	Equations
N : total population S : susceptible state I : infected state R : recovered state	$dS/dt = bN - \frac{\beta SI}{N} - \mu S$
Parameters	$dI/dt = \frac{\beta SI}{N} - \sigma I - (\mu + \alpha)I$
β : transmission rate = p*c p : probability of transmission c : contact rate *b* : birth rate μ : death rate α : disease-induced death rate σ : recovery rate	$dR/dt = \sigma I - \mu R$ $N = S + I + R$

In 1999, Austin and Anderson summarized some of the earliest work of mathematical modelling of the spread of pathogens in healthcare settings.(116) The risks and costs of antibiotic resistance spread motivated researchers to develop simple models to better understand the

mechanisms driving resistance (i.e. antimicrobial consumption) and dissemination. Austin and Anderson detailed some of the first descriptions of models addressing the evolution of resistance within the host and antimicrobial therapy, the intra-hospital indirect transmission and colonisation, inter-hospital outbreaks, and the evolution of resistance in the community.(116) Soon after, other experts proposed that mathematical modelling may be useful in both the quantification of the transmission process and the impact of infection control practices in the hospital.(113) In order to address the potential reluctance of healthcare professionals in applying the recommendations resulting from these models, modellers aimed to elucidate theoretical framework to health practitioners in order to defend the usefulness of modelling in helping improve control interventions. (113, 114)

Mathematical modelling of pathogens in healthcare settings has made headway over the years. By 2013, close to 100 papers were published on dynamic mathematical models in the healthcare setting.(115) The majority of models of common HAI pathogens were set in high-income countries, with the majority evaluating infection control effectiveness, and usually of antimicrobial resistance organisms such as MRSA and VRE. The methods also evolved over time: from compartmental models (where the model contains compartments that separate the population in to different groups, groups that individuals can transition from) to agent-based models (in which each individuals is characterized independently); and from deterministic models (where the output is determined by the initial conditions and parameter values) to stochastic models (in which randomness in infection processes is taken into account, allowing to provide prediction intervals). Applying more advanced methods such as fitting models to data and applying sensitivity analyses also gained ground over time.(115) Although limited to a few settings and pathogens, mathematical modelling of common HAIs helped increase the understanding of the complexity of the healthcare system and transmission dynamics.

4.2 A digital revolution in epidemiology

The use of digital data in the field of epidemiology has grown over the years as both the public health system has digitalized many of the ways it operates and as other non-public health sources of digital data have become available.(117, 118) Early models were parameterized using literature reviews (119) or observed prevalence of carriage.(120) Later, model parameters were estimated using more detailed data such as daily or weekly carriage prevalence.(121, 122) More recently, data on the connections that exist between individuals and institutions have been used for the reconstruction of the underlying existing social network structures.(123) Data on inter-individual connections and social network analysis have been more commonly used to

better understand infectious disease transmission, notably for sexually transmitted infections (124, 125), *Mycoplasma pneumoniae* (126), and SARS.(127) Regarding pathogen spread in healthcare settings, two types of network structure data are becoming more available: contact network data between patients and HCWs and patient referral or transfer data to construct healthcare networks. Some argue that the growing body of work concerning healthcare networks has demonstrated a "next-generation" network-approach to hospital infection prevention and control.(110) Concerning contact networks, models can integrate digital trace data that can measure face-to-face proximity (107, 128) or movement of individuals in a hospital system (129) as these data sources become more widely available.

Chapter 5. Healthcare networks

5.1 A brief introduction

In 2013, Ciccolini and colleagues coined the term "healthcare networks" in response to a growing number of theoretical works linking HAIs spread and patient movement in the healthcare system.(110) They defined healthcare networks as:

> *"Regional and national inter-institutional patient referral[s] connect[ing] hospitals, LTCFs [long-term care centres], and general practitioners. This cooperative system […] can be described using different mathematical approaches. The simplest approach involves counting. For all pairs of hospitals in a cooperative system, the number of patients who, after being discharged from one hospital, were subsequently admitted to the other is counted over a defined time interval."*

In other words, healthcare networks are cooperative healthcare systems where hospitals and other healthcare centres are linked by shared patients through secondary (inter-facility) transfers or referral. Therefore, the structure of the network is composed of healthcare facilities ("vertices") connected by the flow of shared patient between them ("edges"). Patient movement patterns are the basis of the structure of the healthcare network. It should be noted that these networks do not include patient referrals from general practitioners offices.

Patient movement relies on the definition of patient transfer. The term "transfer" like that of "contact" can vary among researchers and should be clearly defined. Patient transfers can be direct or indirect transfers. Direct transfers can be either patients discharged from a sending healthcare facility and admitted to a receiving facility in another jurisdiction moved directly through the facility transportation service (ambulance or helicopter for example) or patients discharged from one medical unit and moved to another in the same healthcare facility jurisdiction. Indirect transfers are cases in which patients are discharged from a healthcare facility, may return home or take a private transportation service, and are admitted to another facility later on. Indirect transfers rely heavily on the definition of the time interval between discharge and admission. Depending on the duration of the carriage of particular pathogens, the inclusion of indirect patient transfers may be useful in improving the sensitivity of the network in including potential carriers; however, the specificity of selecting carriers of potential HAI pathogens may decrease due to the inclusion of potential community-acquired (CA) infections.

Patients are transferred in the healthcare system so that they can receive ongoing medical care. Patients may need specialised care that is unavailable in the first healthcare facility and can

include either specialised healthcare professionals, services, or equipment.(130-132) Transfers may reflect organisational conventions and may not necessarily be motivated by optimum patient outcomes.(133) A study in France evaluated the trajectories of lung cancer patients throughout the healthcare system and identified five pathways: patients who utilise local care in private clinics, patients who are treated in their local region or travelled beyond their region (with a preference for surgeries at teaching hospitals), and patients treated outside of their region accounting for 44% of regional movement.(134)

The number of patients shared between two healthcare facilities depends on the defined time interval. Since transfer counts are cumulative, they can be counted daily, weekly, monthly, annually or include multiple years, depending on objectives of the reconstruction of the network and the refinement of the data available. Although it is not in the scope of the thesis, temporal networks can be formed depending on the time intervals used to cut or construct the networks. They result in a dynamic network in which links may turn "on" and "off" depending on the time interval. Temporal networks may be advantageous in term of network controllability and can provide insights on infection spread dynamics as well.(135)

5.2 Graph theory

In order to better understand healthcare networks, it is essential to cover the basic topology (the layout of the connections) to better understand their structure. In mathematics, networks are commonly referred to as graph, represented as an adjacency matrix – a square symmetric matrix.(136, 137), A network is a graph of V vertices or nodes and E edges that can be described by its' $V \times V$ adjacency matrix A defined as:

$$A_{ij} = \begin{cases} = 1 \text{ if } i \text{ and } j \text{ are connected} \\ = 0 \text{ otherwise} \end{cases}$$

In a healthcare network, vertices (V) are defined as healthcare facilities and edges (E) as the patient trajectories that connect them. For each patient trajectory there is an origin *i* and target *j* facility. The number of patients moving between facility *i* and *j*, is defined as w_{ij}. Also referred to as edge weights, in other words this measure represents the number of patients transferred within the trajectories between two healthcare facilities over a defined period. The sum of the edge weights of the adjacent edges, also called the strength, is given by:

$$s_i^w = \sum_{j \in \Gamma(i)} w_{ij}$$

in which $\Gamma(i)$ is the set of neighbour facilities of i.(136)

It is often important to identify the shortest path among all that are possible between two vertices in a network. The shortest path length ℓ is given by the following distance:

$$\ell(i,j) = \min_{paths(i \to j)} |path|$$

The diameter of a network corresponds to the maximum value of $\ell(i,j)$ and the average shortest path length's behaviour depends on the type of network structure.(136)

In unweighted directed networks, the network distance between any two vertices i and j is considered as the sum of the number of jumps or edges between the vertices where the unit weights of each edge are equal. The shortest path between any two vertices in a network is the minimum distance or sum of edges between them. In a weighted network, the edges can have an unequal weight, where the weight is considered as a "length" of the distance between vertices i and j. As a result, dubbed as "the shortest path problem" in a weighted network, the shortest path between any two vertices is the path with the lowest sum of the edge weights of all edges between them. One way to find the shortest path between any two vertices is to use a classic algorithm proposed by Dijkstra.(138)

To identify the most important vertices or facilities of a network, a series of centrality measures can be calculated. The degree of a facility, k, is the number of facilities one facility is connected to through its patient trajectories (136) defined as:

$$k_i = \sum_j A_{ij}$$

The average degree of a network (136) is given by:

$$\langle k \rangle = \frac{1}{V} \sum_i k_i = \frac{2E}{V}$$

In addition, A_{ij} is often a directed graph in which the directionality of patient transfers from one facility to another is taken into account. Consequently, the indegree (deg^-) and outdegree (deg^+) of any given vertices can be measured in which the degree sum formula is given by:

$$\sum_{v \in V} deg^+(v) = \sum_{v \in V} deg^-(v) \; = \; |E|$$

The network clustering coefficient measures the number of formed triangles in the network. It serves to give an idea of the density of links between vertices in the network and the probability of neighbours of nodes to be connected.

Another measure, betweenness centrality, measures the importance of hospital acting as an intermediary between other facilities defined as:

$$g(i) = \sum_{s \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}$$

where betweenness centrality g(*i*) is equal to the sum of the σ_{st} the number of shortest paths going from *s* to *t* through facility *i* measuring the importance of facility *i* to the organization of flows in the network.(136) The same measure can be calculated for edges:

$$g(e) = \sum_{e \in E} \frac{\sigma_{st}(e)}{\sigma_{st}}$$

where the edge betweenness centrality g(e) is equal to the sum of the σ_{st} the number of shortest paths going from *s* to *t* through edge *e* measuring the importance of edge *e* to the organization of flow in the network.(136)

These basic network topology measures are a few of many measures that can help describe the structure of healthcare networks and the patient transfer patterns forming them and their potential implications on pathogen spread.

5.3 A theoretical basis of pathogen spread in networks

Network topology has been shown to have an effect on the rate and pattern of disease spread.(139-141) Many models predicting spread of infections assume that infected individuals have the same probability to infect all susceptible individuals of a population.(142) Although this assumption allowed early modellers to produce differential equations to facilitate the description of the infection, this did not reflect the true contact patterns.(139, 142) Individuals have heterogeneous contacts in populations and interact with only a small subset of the population. Since contact is required for transmission of pathogens, contact network structures are essential in the understanding of pathogen spread dynamics.(141) Many of the first models were developed in the 1980s and 1990s and assessed the spread of infections in networks.(143-145) Mark Newman was one of the first researchers to describe and solve epidemiological models of infection spread on networks.(139) His work described the variation of disease-causing contacts, the disproportionate effect of highly connected individuals on dissemination,

epidemic and non-epidemic outbreak size, and the effect of structured populations on spread in networks.(139)

One of the first publications applying network theory to spread of pathogens in the healthcare settings was published by Meyers et al. in 2003.(126) Basing their model on a real epidemic of *Mycoplasma pneumoniae* in a physiatrist healthcare facility, the authors recreated a theoretical representation of the HCW and patient contact network to estimate the epidemic threshold. They determined that the extent and duration of the outbreak relied on the degree (number of connections) of the HCWs. Soon after the SARS epidemic in 2002-2003, hospital transmission and super-spreading events were identified to have substantially contributed to the epidemic.(146) Meyers et al. explored the potential epidemiological outcomes of the SARS epidemic and public health interventions using another theoretical contact network framework to inform SARS spread dynamics.(127)

Before the SARS epidemic, at the end of the 20th century, regional and inter-facility HAI outbreaks were being reported: *Staph. aureus* in New York in 1984 (147), MRSA in Canada in 1999 (148), VRE in the USA in 1999 (149, 150), and CPE in New York in 2000 (151). In 2001, Eveillard et al. found an association between HAIs in a 400-bed Parisian hospital and patient transfers.(11) A few years later in 2004, Smith et al. published one of the first theoretical models of HAI pathogen spread and infection control that took into account multiple healthcare institutions linked by patient movement.(152) Along with another work describing the impact of infection control strategies on neighbouring healthcare facilities, Smith et al.'s theoretical work support regional coordination efforts to reduce HAIs, efforts that had been implemented and evaluated in some previously aforementioned outbreak situations.(150, 153, 154)

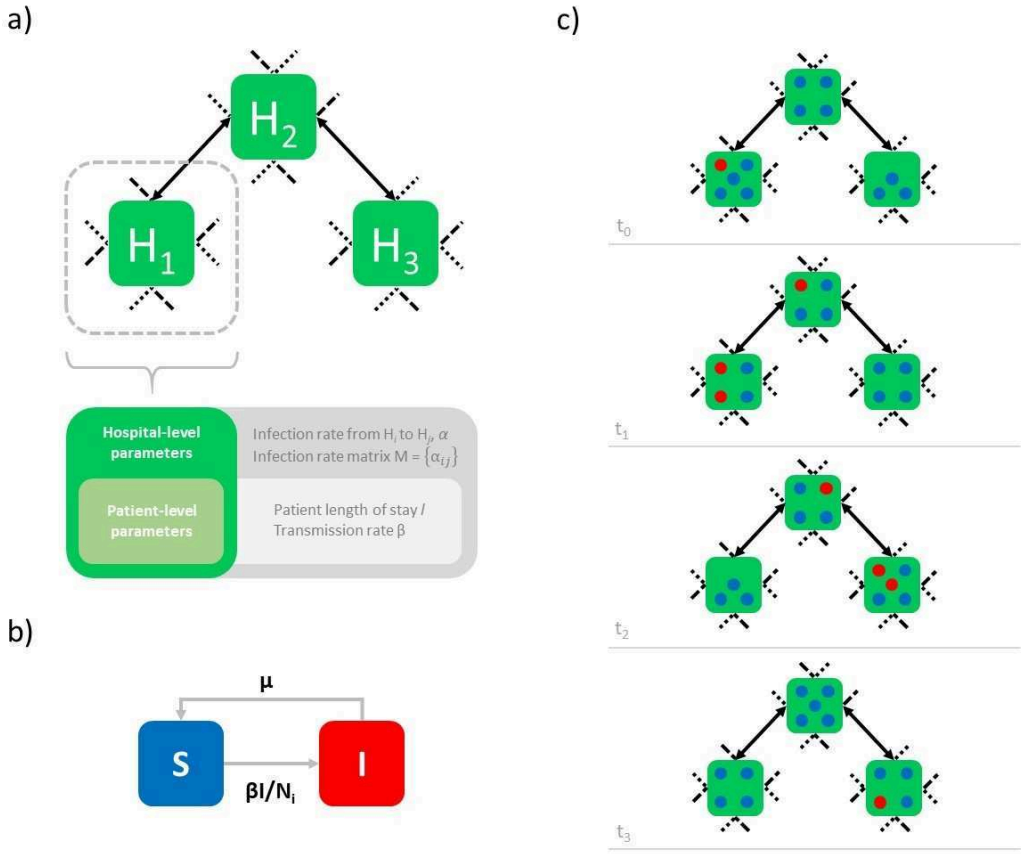
Both theoretical network types (HCW and patient contact networks and healthcare networks based on patient transfers) pointed to the importance of incorporating network structure to better understand transmission dynamics, the predicted size of epidemics, and to support regional coordination among healthcare facilities. Of note, heterogeneities of acquisition and dissemination among health care facilities in the network were still yet to be explored. However, the digital evolution in epidemiology opened new avenues for these types of models. Incorporating real network data in their predictions had the potential to describe the heterogeneities in networks such as the variations in connectivity of individual facilities and the number of patients being exchanged.

5.4 Models of pathogen spread on real healthcare networks

As support for the importance of considering healthcare networks in the dynamics of pathogen spread gained ground, electronic medical records provided an opportunity for researchers to use real data on admissions, discharges, and transfers to create real healthcare networks that better reflected the heterogeneities of transfers.(17) Observed movement in healthcare networks and the catchment area were first integrated into a model in 2007 by Robotham et al. using 1.4 million admissions and discharges from three hospitals in the UK over a seven-year period.(12) The authors found that there was a 44.2% chance of MRSA-positive patients to be readmitted to these hospitals while still being colonised.

The first individual-based model on a nation-wide healthcare network of general, university, and clinical hospitals was reconstructed in 2010 for the Netherlands.(13) This marked also the first time network topology measures such as degree, in-degree, and out-degree were used to explain the prevalence of MRSA in different types of facilities. The individual-based healthcare network model can be generalised in Figure 11. The model considers both the patient-level parameters and transition states but also the network structure of healthcare facilities to inform the infection rate between them through their patient transfers (Figure 11). In order to simulate spread, one can choose a random proportion of patients to infect in an index hospital and the spread of the infection can be followed over time. The authors of the first individual-based healthcare network found that infection spreads quickly from the index hospitals and that due to the high degree or high connectedness and centralisation of patient movement towards teaching hospitals, these hospitals were likely to have a higher HAI prevalence.(13)

Figure 11. Individual-based healthcare network model



An individual based healthcare network model considers the transition states of individual or patients (i.e. SI model given here in b) where a patient transitions from susceptible to infected and back to susceptible when the infection or colonisation is lost) and we also consider the hospital-level dynamics that rely on the healthcare network structure that determine the infection rates between healthcare facilities shown in a). A simulation could give rise to c) where at the initial starting time t_0 only one random patient is infected in H_1; at t_1 this patient infects two contacts (one who stays in H_1 and one who is transferred to H_2); at t_2 the H_2 contact infects an H_2 contact at and then gets transferred to H_3 where they infect another contact; and at t_3 this H_3 contact remains infected while the other patients lose their infection. It should be considered that the rate of patient transfers leading to infection in another hospital relies on a probability that takes into account the number of patient transfers occurring between them. Therefore, a higher number of patient transfers leads to a higher rate of infection between hospitals.

In 2011 Lee and colleagues published the first paper focused on the social network characteristics of healthcare networks.(155) The authors found heterogeneous patient sharing that led to hospitals being either highly connected or poorly connected (in terms of degree, betweenness, density, and network ego measures).

The number of mathematical models integrating electronic discharge summaries of observed patient transfers increased significantly during the 2007 to 2016 period.(17) In particular, a computational software tool called the Regional Healthcare Ecosystem Analyst (RHEA) was developed in 2013 to facilitate individual-based modelling of pathogen spread in healthcare networks.(156) The tool considered admissions to general wards or the ICU both from the

community and from hospital transfers. RHEA was used in particular to model healthcare facilities in Orange County, California, that included acute-care and long-term care settings such as nursing homes and modelled various pathogens such as MRSA, norovirus, VRE and CPE.(157-160)

Overall, a series of publications assessed the role of hospital system-wide spread and control through patient movement.(12, 156-167) In addition, data on transfers within a healthcare network gave insights into sentinel selection for development of more effective sentinel surveillance systems (14, 15, 161, 168) and supported improved coordinated regional control.(158, 162, 166, 169) Finally, some models explicitly assessed the underlying network structure of interactions through social network analyses of patient flows within hospitals (169-174) and between hospitals.(13-15, 156, 160, 168, 175, 176) Finally, many studies show that healthcare networks display a community structure in which communities reflected a shared patient transfer population among healthcare facilities.(13, 16, 110, 175)

In summary, mathematical models of healthcare networks showed their value to inform decision-makers on enhanced coordinated regional and national approaches to infection control strategies, in a context where increasingly centralized healthcare systems favour the spread of HAIs.(110, 112, 161)

Chapter 6. First article - Use of network data in HAI spread models

6.1 Summary

As new and real data (meaning data produced by human activity) on networks have become more widely available for research purposes (117, 177), an evaluation of their use in mathematical models in the healthcare setting was deemed necessary to assess their contribution to the understanding of infection risks. We conducted a systematic review of mathematical models in the healthcare setting using real data on network structure in order to give a review of the sources of network data and the methods (i.e. parameter estimation) used to integrate the data in modelling studies in the healthcare setting over time.(17) The review also evaluated how they may have improved the understanding of HAI transmission dynamics. The review covered the temporal evolution of the scientific publication of these models, the country and type of healthcare facilities modelled, the pathogens studied, the data sources, the objectives of the research (in terms of infection control assessment or understanding HAI spread dynamics), parameter estimation, and model cross-validation.

Network data has allowed mathematical models to have more realistic predictions of infection dynamics in healthcare settings and systems. The review found that the number of publications of models in healthcare settings has increased over the years; however, they remain limited to high-income settings, the ICU and hospital settings, and to a limited number of pathogens. The use of network data has also grown over the years with a wider diversity of data sources being implemented into models such as shadowing or sensor data of HCW and patient contacts and electronic medical records of patient transfers. These models have given rise to new insights into more effective HAI prevention and infection control strategies such as hand-hygiene practices and cohorting. However, settings such as long-term care facilities, data on infection, model parameter estimation, and model cross-validation were found lacking and future research should consider further expanding work in these fields

Mathematical models of infection transmission in healthcare settings: recent advances from the use of network structured data

Rania Assab[a,b,*], *Narimane Nekkab*[a,b,c,*], *Pascal Crépey*[c,d,e], *Pascal Astagneau*[f], *Didier Guillemot*[g,h,i,j], *Lulla Opatowski*[g,h,i], *and Laura Temime*[a,b]

Purpose of review
Mathematical modeling approaches have brought important contributions to the study of pathogen spread in healthcare settings over the last 20 years. Here, we conduct a comprehensive systematic review of mathematical models of disease transmission in healthcare settings and assess the application of contact and patient transfer network data over time and their impact on our understanding of transmission dynamics of infections.

Recent findings
Recently, with the increasing availability of data on the structure of interindividual and interinstitution networks, models incorporating this type of information have been proposed, with the aim of providing more realistic predictions of disease transmission in healthcare settings. Models incorporating realistic data on individual or facility networks often remain limited to a few settings and a few pathogens (mostly methicillin-resistant *Staphylococcus aureus*).

Summary
To respond to the objectives of creating improved infection prevention and control measures and better understanding of healthcare-associated infections transmission dynamics, further innovations in data collection and parameter estimation in modeling is required.

Keywords
hospital-acquired infections, mathematical modeling, networks, systematic review, transmission

INTRODUCTION

Despite advances in biology and medicine, the burden of healthcare-associated infections (HAIs) has increased over the last decades [1]. Indeed, HAIs are the most frequent adverse event in health-delivery settings affecting up to one in three patients in ICU in developed countries [1]. The associated costs are estimated to be seven billion euros in Europe, and approximately six and a half billion dollars in the US [2–4], where 722 000 HAIs occur yearly in acute-care hospitals, resulting in 75 000 deaths [5].

The HAI burden stems notably from the emergence and spread of virulent infectious agents. Multidrug-resistant bacteria such as methicillinresistant Staphylococcus aureus (MRSA) and carbapenemase-producing Enterobacteriaceae (CPE), and viruses such as influenza, severe acute respiratory syndrome (SARS), Middle East respiratory syndrome coronavirus (MERS-CoV) and Ebola have become of concern for public health authorities in most countries [1]. Prevention measures such as hand hygiene, isolation, antibiotic restrictions, staff cohorting, and surveillance may significantly impact HAI rates, decreasing in particular MRSA and *Clostridium difficile* incidence by more than 70% [5].

[a]Laboratoire MESuRS, Conservatoire National des Arts et Métiers, Paris, [b]Unité PACRI, Institut Pasteur, Paris, [c]Département d'Epidémiologie et de Biostatistiques, Ecole des Hautes Etudes en Santé Publique, Rennes, [d]UMR IRD 190, INSERM, EHESP, Aix-Marseille University, Marseille, [e]UPRES-EA 7449 REPERES, Recherche en Pharmaco-Epidémiologie et Recours aux Soins, EHESP, Université de Rennes 1, Rennes, [f]Centre de prévention des infections associées aux soins (C-CLIN), APHP, Paris, [g]Université Versailles St Quentin, Versailles, [h]Institut Pasteur, Paris, [i]INSERM UMR1181, B2PHI and [j]Unité Fonctionnelle de Santé Publique, APHP, Hôpital Raymond-Poincare, Garches, Paris France

Correspondence to Rania Assab, 292 Rue Saint-Martin, F-75003 Paris, France. Tel: +33153018069;
e-mail: rania.assab@cnam.fr, narimane.nekkab@cnam.fr

[*]Rania Assab and Narimane Nekkab contributed equally to this article.

Curr Opin Infect Dis 2017, 30:410–418
DOI:10.1097/QCO.0000000000000390

KEY POINTS

- Mathematical models of infections in healthcare settings have become more frequent over the years.
- Increasing trends of models based on real data on networks of individuals or facilities are due to perceived usefulness as tools for infection prevention and control, increased availability of digitalized medical records or surveys, and development of sensor technology
- The range of pathogens, settings, and situations explored by these models remains to this day highly restrictive which may reflect limited data availability, historical importance of certain infections (i.e., MRSA), and high-risk HAI settings that require more intensive HCW training and precautions (i.e., ICUs).
- The main contributions of models in terms of using real data on networks are to develop more innovative and realistic HAI control strategies and to better understanding the impact of social networks on HAI spread.

DEFINITIONS

- Compartmental model: a model where a population is subdivided into groups corresponding to a status. For example, the susceptible-infected-recovered model is a basic compartmental model composed of three groups of people with the following status: susceptible, infected, and recovered. Each compartment contains a certain number of people from the population presenting the status.
- Agent-based model: rather than grouping people in a compartment in terms of their status, the agent-based model studies each individual separately. These models commonly study the connections between individuals (patients and/or HCWs) with each other in terms of a shared environment (ward, room) or through their contacts (direct, indirect).
- Deterministic model: a model in which the output is fully determined by the initial conditions and parameter values (usually a compartmental model formulated using differential equations).
- Stochastic model: a model including inherent randomness, in which, for a given set of initial conditions and parameter values, an output distribution is provided to account for uncertainty in predictions (often used for small populations in which random fluctuations are important).
- Social network: a network with components and links, and within the scope of our review, they are either contact networks of healthcare workers (nurses, physicians and so on) and patients or of hospitals that are linked by their patient transfers.
- Social network analysis: in the case of our review, it is the assessment of the contacts or healthcare system structures which can help identify 'super-spreaders' that are highly linked and have the most potential to spread disease in the network.

Mathematical models have provided a theoretical framework for understanding complex transmission dynamics within healthcare settings for over 15 years [6–9]. Furthermore, they provide a quantitative approach to estimating the impact of various infection control strategies and their combined effects [6,7,9,10].

Over recent years, detailed data informing on the interactions between patients and healthcare workers (HCWs) or patient transfers within and between healthcare settings have been integrated in such models. Patients transfers between hospitals have been increasingly studied [11], as well as data on contacts between patients and HCWs, in particular, digital trace measuring face-to-face proximity [12,13] or individual movements [14].

Here, we conduct a systematic review of mathematical models in healthcare settings using such real data on networks within institutions and between institutions. We present an overview of the methodological specificities related to the integration of network data in the different modeling studies and we study how they may improve our understanding and predictive capacity of HAI spread in healthcare settings.

METHODS

We conducted a systematic search in three different databases: MEDLINE (1946 to present), Web of Science Core Collection (1956 to present), and Institute of Electrical and Electronic Engineers (IEEE) Xplore Digital Library (1893 to present). Results included all articles published until 26 January 2017, the final day of the search. All results from the search query were independently screened by two reviewers for inclusion criteria eligibility and selection after review of titles, abstracts, and then full texts. Query structure, inclusion and exclusion criteria can be found in Appendix 1 and 2, http://links.lww.com/COID/A20.

We defined four lists to classify our selection results:

(1) L: all studies meeting our first two inclusion criteria comprising of all mechanistic models of pathogen transmission within healthcare settings. We use the term 'HAI' in a generic and inclusive way to encompass multidrugresistant organisms such as MRSA, ESBL (extended-spectrum beta-lactamases) producers, influenza, and VRE (vacomycin-resistant enterococci) among other pathogens.
(2) L_1: all studies from list L incorporating real contact data (within institutions).
(3) L_2: all studies from list L incorporating real transfer data (between institutions and/or wards).

(4) L_3: all studies from list L that incorporate explicit contact or transfer network structure in healthcare settings without real data.

All studies using real data (L_1 and L_2) were analyzed regarding various characteristics such as pathogen studied, data sources, and model parameters. We also compared L_1 and L_2 models characteristics with L models characteristics using Fisher exact and χ^2 statistical tests.

RESULTS

Our search retrieved a total of 5653 distinct records from the three databases (Fig. 1). After screening titles and abstracts, a total of 216 studies were selected for list L, including eight papers added through reference searching. From this list, we identified a total of 28 models using intra-hospital contact data (L_1) [15–40,41▪,42▪], 26 models using inter or intrahospital transfer data (L_2) [43,44▪▪, 45–58,59▪,60–66,67▪▪,68], and 22 contact or transfer network healthcare models without real data (L_3) [69–90].

Publication trends

Publication of mathematical models of pathogen spread in healthcare settings has greatly increased in recent years (Fig. 2; $P < 10^{-11}$, Spearman's rank correlation). The first models including real network data were published in 2002 and used directly observed within-hospital data on inter-individuals contacts [31,34]; the first model including data on interfacility transfers was published in 2007 [47].

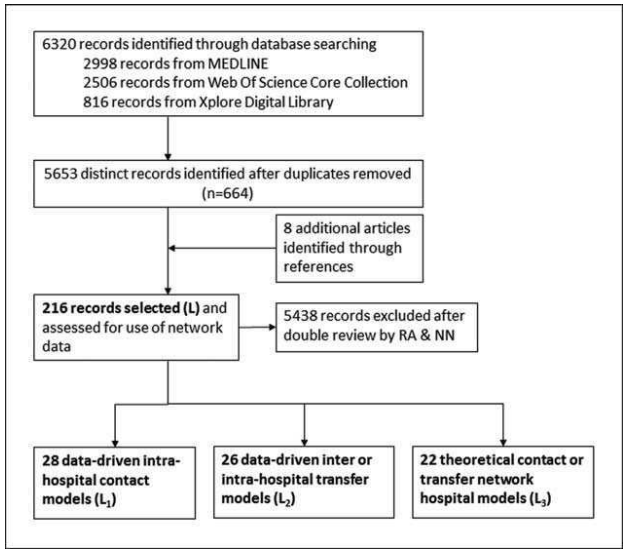


FIGURE 1. PRISMA diagram reviewing literature sources for mathematical models that examined the transmission dynamics in healthcare settings.

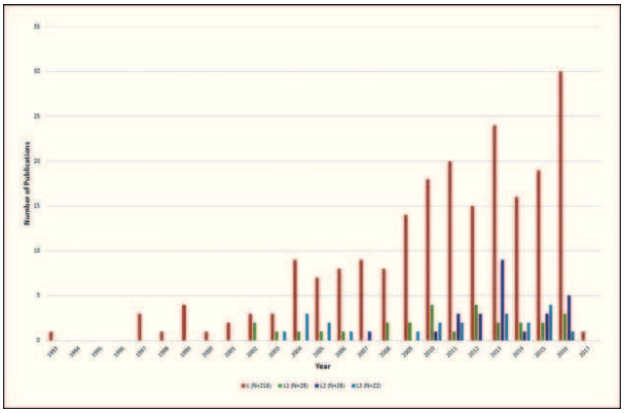


FIGURE 2. Number of mathematical models of healthcare-associated infections spread in healthcare settings published over time. The total of all models published (L, in red), those using real contact data (L_1, in green), those using real transfer data (L_2, in purple), and those focusing on the impact of social networks without real data (L_3, in blue) are depicted.

From these first publications on, the number of yearly published L_1 ($P = 0.03$, Spearman's rank correlation) and L_2 ($P = 0.02$, Spearman's rank correlation) models have been increasing. Overall, since 2002, L_1 and L_2 models represent 27% of L models, with an increasing portion of L_2 models (Suppl Figure 1, http://links.lww.com/COID/A20).

Pathogens studied and epidemic situations

MRSA was the most studied pathogen in L_1 and L_2 models (44.1%) followed by: influenza (13.6%), vancomycin-resistant *Enterococci* spp. (8.5%), HAIs in general (8.5%), *C. difficile* (5.1%) and carbapenemase-producing *Enterobacteriaceae* (5.1%; Fig. 3).

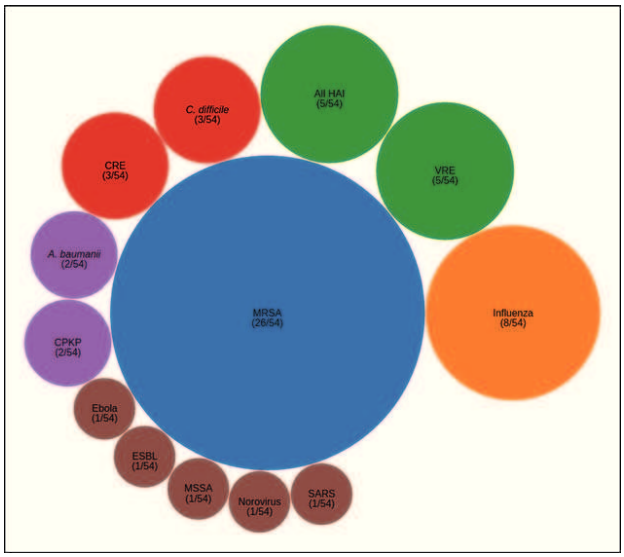


FIGURE 3. Pathogens modeled in models using real contact (L_1) or transfer data (L_2).

The distribution of pathogens studied in L_1 and L_2 models did not differ significantly from that observed in all models ($P = 0.09$, Fisher exact test).

In general, L_1 and L_2 models either simulated outbreaks of these pathogens in a susceptible population or assessed the impact of long-term infection prevention and control on the ongoing epidemic of prevalent HAIs.

Healthcare settings

Out of 54 L_1 and L_2 models, 49 (91%) took place in acute-care settings (Table 1a). The mean number of healthcare settings in general included in the models was 122 [range: 1–3306], with a median at one of L_1 and at 98 of L_2 models. Ward-level description of HAI spread was present in 35% of publications (19/54), of which most modeled one ICU (Table 1b). Only Karkada *et al.* [58] study was an outlier, analyzing a total of 3306 ICUs in the US. L_1 models had a median study size of 100 patients [range: 2–3329] and 34 HCWs [range: 1–19508]. L_2 models incorporated a median of seven million transfers [range: 130 000–13 million].

Data sources

All transfer data were collected using electronic patient records such as national medical and surveillance registers [43,53,54], hospital discharge summaries [45,47–49,51,60,62,65], or insurance databases [58]. Data used to collect the contact patterns between patients and HCWs came from four main sources: shadowing – direct observation of interactions between patients and HCWs, surveys, medical records, and individual wireless proximity sensors recording the identity of other sensors located in a close area. Historically, between 2002 and 2006, shadowing was the first source of data on contact networks in healthcare settings (Fig. 4) [20,31,32,34,37]. During the period 2007–2011, new methods of contact data collection appeared such as medical records [19,25,28] and surveys [17,18,38,40,41▪]. Finally, following technology innovations, proximity sensors were introduced in four studies published over the period 2012–2016 [24,35,36,42▪].

Types of models

L_1 and L_2 models were mostly agent based, rather than compartmental (43 vs. 12 models), and stochastic, rather than deterministic (53 vs. 4 models) [39]. These were significant differences with L models ($P < 10^{-5}$, χ^2 test).

Model objectives

L_1 and L_2 models all aimed at either assessing control interventions or better understanding HAI spread and the impact of social networks. Inclusion of data on social networks allowed simulating more innovative and realistic infection prevention and control strategies, including heterogeneous

Table 1. Type of (a) healthcare settings and (b) wards modelled in L_1/L_2 models: total number of published models and corresponding references. Five articles model studied both nursing homes and hospital settings (a)

(a)

Healthcare setting		Number	References
Hospitals		49	[15–23,26–37,39,41▪,42▪,43,44▪▪,45–50,52–58,59▪,60–66,67▪▪,68]
Long-term care	Nursing homes	9	[25,38,40,44▪▪,45,60,63,66,68]
	Tertiary care	4	[24,51,67▪▪,68]
Total		62 (54 distinct articles)	

(b)

Ward	Number	References
Emergency	1	[39]
General ward	4	[17,22,42▪,52]
Geriatric	1	[18]
Hemodialysis	1	[31]
ICU	10	[16,20,26,34,35,39,41▪,52,58,67▪▪]
Pediatric	1	[36]
Surgical	1	[21]
Total	19	

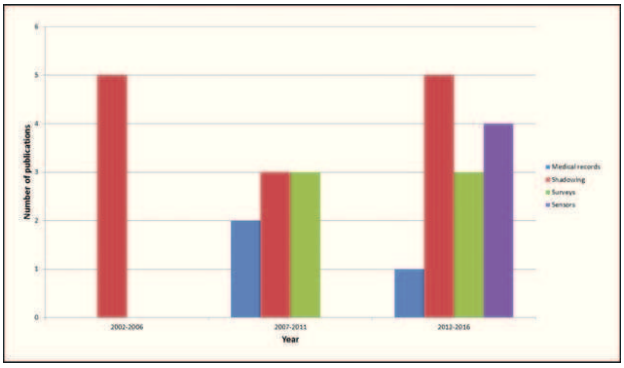


FIGURE 4. Contact data collection sources in models using real contact data (L_1): changes over time.

hand-hygiene compliance or cohorting levels [16,19,22,27,31,32,34,35]. Hand-hygiene compliance was the most common intervention studied [20,22,27,29,31,34,35,37,39,40,41▪,46,56], along with antibiotic exposure [32,33], targeted or screening of patient groups or universal screening of all patients at admission [48,51,57], isolation [26,46,57], and HCW vaccination [17–19]. Other models explored the role of patient-HCW interactions through variations in cohorting by modifying patient: HCW ratios [16,31,34], social interactions in hospitals [24,28,36], and hospital system-wide spread and control through patient movement [45,47–50,59▪,60–66]. In addition, data on transfers within a healthcare network gave insights into sentinel selection for development of more effective sentinel surveillance systems [43,50,55,58] and supported improved coordinated regional control [23,45,59▪, 61,68]. Finally, some models explicitly assessed the underlying network structure of interactions through social network analyses of patient flows in hospitals [19,23,30,35,87,89] and between hospitals [50,53–55,58,64,65].

Parameter estimations and model cross-validation

Around 17% of L_1 or L_2 studies included model parameter estimation using observed infection or colonization data, rather than simple calibration or using values from the literature. Model predictions were rarely cross validated with independent datasets (eight publications overall). In these aspects, L_1 and L_2 models did not differ from L models in general.

DISCUSSION

Mathematical models of infections in healthcare settings have become more frequent over the years. This increase may be because of multiple factors including perceived usefulness of models as tools for understanding the impact of infection prevention and control in the health field, for understanding drivers of recent major epidemics such as the 2002–2003 SARS outbreak [15,78,91,92] and the 2014–2015 Ebola epidemic [42▪,93–95], or growing awareness of factors contributing to the global impact of antibiotic resistance [96]. In parallel, increased availability of digitalized medical records or surveys, and development of sensor technology to monitor interindividual contacts provide researchers with the means to build more realistic models.

Review scope and limitations

In this systematic review, we conducted an exhaustive search of articles studying pathogen spread in healthcare settings through mathematical modeling. Using complementary databases (PubMed, Web of Science, and IEEE Xplore Digital Library) was important and necessary to find the articles analyzed in this review.

However, this review was subject to some limitations. Given that the scope of this review involves both health sciences and computational biology, we could have included more databases in the computer science field. In addition, we only considered publications in English and French, which may have limited the variety of country settings. Statistical models were excluded because they did not meet the objective of the review; however, these models may also improve the understanding of transmission dynamics of pathogen spread in healthcare settings.

Main results of the review and implications for future work

Several points which have been raised by our review may lead to recommendations for future modeling work. The range of pathogens, settings, and situations explored by models based on real data on networks of individuals or facilities remains to this day highly restrictive. Hence, the increased realism in the description of social networks is counterbalanced by the current limitations in the range of investigated questions.

First, 80% of L_1/L_2 models were set in a four developed countries (the US, UK, the Netherlands, and France), while L models considered a wider variety of countries (Suppl Figure 2, http://links.lww.com/COID/A20). This can be explained by their use of more advanced data management technologies, resulting in a better availability of relevant data, as well as by the presence of a very active community of modelers. However, HAIs also represent a major issue in developing countries,

mainly due to high antimicrobial resistance levels and difficulties to afford second-line treatments [97]. Future work should take these settings into account.

Second, the most studied pathogen was MRSA, followed distantly by influenza, HAIs in general and vancomycin-resistant enterococci. Although this was true of all models of HAI spread, the domination of MRSA was even stronger in models incorporating data on observed networks. This may be explained by the large amount of available epidemiological data on MRSA in healthcare settings, reflecting the historical importance of MRSA in HAIs. In addition, data on MRSA carriage are easily collected from nasal or other surveillance swabs, while other pathogens such as *Enterobacteriaceae* require rectal swabs, which can be more difficult to obtain. Although MRSA has indeed represented a major threat over the last decades, the incidence of MRSA infections currently seems to be declining in most developed countries, [97] yet other multiresistant bacteria such as ESBL-producing *Enterobacteriaceae* become more prevalent [98,99]. Future models should definitely consider a wider range of pathogens.

Third, the vast majority of L_1 and L_2 models were set in acute-care settings, with most ward-level descriptions taking place in ICUs. ICUs are frequently modeled because of their high risk of HAIs, raised by a high number of invasive procedures in critical-state patients, and require well informed recommendations regarding control interventions. Consequently, research funding in ICUs is more prevalent and both data collection and implementation of control interventions are facilitated by better informed ICU HCWs compared with other wards. However, HAIs are also an issue in other types of hospital wards, in which lower HCW-to-patient ratios and decreased risk awareness may lead to HAI outbreaks. On the health systems-level, the majority of L_2 models described networks of hospitals linked by their shared patients; only a few recreated transfer networks between the wards of a given hospital to study how the impact of infection prevention and control interventions may vary depending on hospital ward specialties [46,51]. Intrahospital spread has been shown to be one of the major reasons for transmission of SARS in Toronto, Canada, and Taiwan and MERS-CoV in Alhasa, Saudi Arabia, and Korea [100–102]. Future research should attempt to include ward-level modeling as it provides more specific and realistic patients and HCW interactions that are overlooked when modeling at the hospital level, and take into account wards other than ICUs.

Additionally, models of HAI spread in settings outside acute-care should be developed. For instance, the importance of nursing homes in the overall spread of HAIs has been underlined. Factors such as long length-of-stay of nursing home residents have been shown to play an important role in both driving sustained endemics of infections and increasing the risk of epidemics in entire healthcare networks [44▪▪,60,63]. Similarly, the impact of transmission in [28] or readmission from [47] community settings on HAI transmission in healthcare settings is rarely assessed among models using real data. Research should focus on modeling nursing home and community settings with collected data to better understand the complexity of interactions within healthcare networks and their impact on transmission dynamics in healthcare settings.

Another important issue is the inclusion of observed colonization or infection data in modeling works to calibrate or validate model predictions. Although models incorporating data on interindividual contacts or patient transfers are more likely to have access to patient medical records or disease status from HCWs, parameter estimation and model validation using colonization or infection data remains rare overall. A major objective of future research should be to include observed infection or carriage data collected simultaneously with the network data, among the same individuals. Another benefit of simultaneously collecting contact or transfer data and infection data would be the possibility of assessing the pertinence of network data to help predict HAI spread. Indeed, although most published models using network data implicitly assume that interindividual contact and/or interfacility patient transfer networks drive HAI spread, other factors may impact pathogen diffusion in healthcare settings. Depending on the involved pathogen, environmental contamination for instance may play a major part. It is therefore of the utmost importance to further investigate what portion of the pathogen-specific diffusion risk may be explained by network data [24].

CONCLUSION

Our review assessed the use of contact and transfer network data in models over time and its impact on understanding infection transmission dynamics in healthcare settings. Models incorporating such data were limited to a small number of countries, settings, and pathogens, while there is a steady emergence of network graphs to study the contact and structure of patient movement and interactions with HCWs. These models give new insights into more effective HAI prevention and control strategies in both endemic and epidemic situations. Further innovations in data collection and use in modeling

are required to improve understanding of transmission dynamics to reinforce existing recommendations and evaluate new control strategies.

Acknowledgements

None.

Financial support and sponsorship

The work was supported in part by the PRINCEPS program from the Sorbonne-Paris Cité University.

Conflicts of interest

There are no conflicts of interest.

REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the annual period of review, have been highlighted as:

- ■ of special interest
- ■■ of outstanding interest

1. Report on the burden of endemic healthcare-associated infection worldwide. Geneva: WHO; 2011.
2. Allegranzi B, Nejad SB, Combescure C, *et al.* Burden of endemic healthcare-associated infection in developing countries: systematic review and meta-analysis. Lancet 2011; 377:228–241.
3. ECDC. Annual epidemiological report on communicable diseases in Europe 2008: report on the state of communicable diseases in the EU and EEA/EFTA countries. 2008
4. Scott RD. The direct medical costs of healthcare-associated infections in U.S. hospitals and the benefits of prevention. Atlanta, GA; 2009
5. HAI data and statistics. http://www.cdc.gov/HAI/surveillance/?con=&dom=prime&src=syndication. (Accessed 28 April 2017)
6. Bonten M, Austin D, Lipsitch M. Understanding the spread of antibiotic resistant pathogens in hospitals: mathematical models as tools for control. Clin Infect Dis 2001; 33:1739–1746.
7. Grundmann H, Hellriegel B. Mathematical modelling: a tool for hospital infection control. Lancet Infect Dis 2006; 6:39–45.
8. Opatowski L, Guillemot D, Boelle P, Temime L. Contribution of mathematical modeling to the fight against bacterial antibiotic resistance. Curr Opin Infect Dis 2011; 24:279–287.
9. van Kleef E, Robotham J, Jit M, *et al.* Modelling the transmission of healthcare associated infections: a systematic review. BMC Infect Dis 2013; 13:13.
10. Gingras G, Guertin MH, Laprise JF, *et al.* Mathematical modeling of the transmission dynamics of *Clostridium difficile* infection and colonization in healthcare settings: a systematic review. PloS One 2016; 11.
11. Ciccolini M, Donker T, Kock R, *et al.* Infection prevention in a connected world: the case for a regional approach. Int J Med Microbiol 2013; 303:380–387.
12. Cattuto C, Van den Broeck W, Barrat A, *et al.* Dynamics of person-to-person interactions from distributed RFID sensor networks. PloS One 2010; 5:5.
13. Isella L, Romano M, Barrat A, *et al.* Close encounters in a pediatric ward: measuring face-to-face proximity and mixing patterns with wearable sensors. PloS One 2011; 6:6.
14. Huang S, Avery T, Song Y, *et al.* Quantifying interhospital patient sharing as a mechanism for infectious disease spread. Infect Control Hosp Epidemiol 2010; 31:1160–1169.
15. Nishiura H, Kuratsuji T, Quy T, *et al.* Rapid awareness and transmission of severe acute respiratory syndrome in Hanoi French Hospital, Vietnam. Am J Trop Med Hyg 2005; 73:17–25.
16. Wang X, Chen Y, Zhao W, *et al.* A data-driven mathematical model of multidrug resistant Acinetobacter baumannii transmission in an intensive care unit. Sci Rep 2015; 5:8.
17. van den Dool C, Bonten MJ, Hak E, Wallinga J. Modeling the effects of influenza vaccination of healthcare workers in hospital departments. Vaccine 2009; 27:6261–6267.
18. van den Dool C, Bonten MJ, Hak E, *et al.* The effects of influenza vaccination of healthcare workers in nursing homes: insights from a mathematical model. PloS Med 2008; 5:1453–1460.
19. Ueno T, Masuda N. Controlling nosocomial infection based on structure of hospital social networks. J Theor Biol 2008; 254:655–666.
20. Triola MM, Holzman RS. Agent-based simulation of nosocomial transmission in the medical intensive care unit. In 16th IEEE Symposium on Computer-Based Medical Systems: IEEE: 2003:284–288
21. Sypsa V, Psichogiou M, Bouzala GA, *et al.* Transmission dynamics of carbapenemase-producing Klebsiella pneumoniae and anticipated impact of infection control strategies in a surgical unit. PloS One 2012; 7:7.
22. Raboud J, Saskin R, Simor A, *et al.* Modeling transmission of methicillin-resistant Staphylococcus aureus among patients admitted to a hospital. Infect Control Hosp Epidemiol 2005; 26:607–615.
23. Polgreen PM, Tassier TL, Pemmaraju SV, Segre AM. Prioritizing healthcare worker vaccinations on the basis of social network analysis. Infect Control Hosp Epidemiol 2010; 31:893–900.
24. Obadia T, Silhol R, Opatowski L; I-Bird Study Group, *et al.* Detailed contact data and the dissemination of Staphylococcus aureus in hospitals. PloS Comput Biol 2015; 11:16.
25. Lapidus N, Carrat F. WTW: an algorithm for identifying 'who transmits to whom' in outbreaks of interhuman transmitted infectious agents. J Am Med Inform Assoc 2010; 17:348–353.
26. Hall IM, Barrass I, Leach S, *et al.* Transmission dynamics of methicillin-resistant Staphylococcus aureus in a medical intensive care unit. J R Soc Interface 2012; 9:2639–2652.
27. Barnes SL, Morgan DJ, Harris AD, *et al.* Preventing the transmission of multidrug-resistant organisms: modeling the relative importance of hand hygiene and environmental cleaning interventions. Infect Control Hosp Epidemiol 2014; 35:1156–1162.
28. Cooley P, Lee BY, Brown S, *et al.* Protecting healthcare workers: a pandemic simulation based on Allegheny County. Influenza Other Respir Viruses 2010; 4:61–72.
29. Cummings KL, Anderson DJ, Kaye KS. Hand hygiene noncompliance and the cost of hospital-acquired methicillin-resistant Staphylococcus aureus infection. Infect Control Hosp Epidemiol 2010; 31:357–364.
30. Cusumano-Towner M, Li DY, Tuo SS, *et al.* A social network of hospital acquired infection built from electronic medical record data. J Am Med Inform Assoc 2013; 20:427–434.
31. D'Agata EM, Horn MA, Webb GF. The impact of persistent gastrointestinal colonization on the transmission dynamics of vancomycin-resistant enterococci. J Infect Dis 2002; 185:766–773.
32. D'Agata EM, Webb G, Horn M. A mathematical model quantifying the impact of antibiotic exposure and other interventions on the endemic prevalence of vancomycin-resistant enterococci. J Infect Dis 2005; 192:2004–2011.
33. Grima DT, Webb GF, D'Agata EMC. Mathematical model of the impact of a nonantibiotic treatment for Clostridium difficile on the endemic prevalence of vancomycin-resistant Enterococci in a hospital setting. Comput Math Methods Med 2012; 2012:8.
34. Grundmann H, Hori S, Winter B, *et al.* Risk factors for the transmission of methicillin-resistant Staphylococcus aureus in an adult intensive care unit: fitting a model to the data. J Infect Dis 2002; 185:481–488.
35. Hornbeck T, Naylor D, Segre AM, *et al.* Using sensor networks to study the effect of peripatetic healthcare workers on the spread of hospital-associated infections. J Infect Dis 2012; 206:1549–1557.
36. Machens A, Gesualdo F, Rizzo C, *et al.* An infectious disease model on empirical networks of human contact: bridging the gap between dynamic network data and contact matrices. BMC Infect Dis 2013; 13:185.
37. McBryde ES, Pettitt AN, McElwain DLS. A stochastic mathematical model of methicillin resistant Staphylococcus aureus transmission in an intensive care unit: predicting the impact of interventions. J Theor Biol 2007; 245:470–481.
38. van den Dool C, Hak E, Bonten MJ, Wallinga J. A model-based assessment of oseltamivir prophylaxis strategies to prevent influenza in nursing homes. Emerg Infect Dis 2009; 15:1547–1555.
39. Wang J, Wang L, Magal P, *et al.* Modelling the transmission dynamics of meticillin-resistant *Staphylococcus aureus* in Beijing Tongren hospital. J Hosp Infect 2011; 79:302–308.
40. Assab R, Temime L. The role of hand hygiene in controlling norovirus spread in nursing homes. BMC Infect Dis 2016; 16:395.
41. ■ DalBen MF, Teixeira Mendes E, Moura ML, *et al.* A model-based strategy to control the spread of carbapenem-resistant *Enterobacteriaceae*: simulate and implement. Infect ControlHosp Epidemiol 2016; 37:1315–1322.

In this study, hand hygiene comliance and contact precautions are evaluated using a Ross-Macdonald model to describe the transmission and control of CRE in an ICU. The authors estimated the parameters including the per capita contact rate, probability of HCW colonization (with and without contact precuations), and HCW compliance to control measures from the ICU's observed collected data.

42. ■ Vanhems P, Von Raesfeldt R, Ecochard R, Voirin N. Emergence of Ebola virus disease in a french acute care setting: a simulation study based on documented inter-individual contacts. Sci Rep 2016; 6:36301.

The authors model the potential Ebola virus emergence in a hospital ward in a nonoutbreak context. In addition, the authors collected contact sensor data to paramerize the transmission between patients and HCWs and showed that nurses are at highest risk for nosocomial Ebola virus disease.

43. Ciccolini M, Donker T, Grundmann H, *et al.* Efficient surveillance for healthcare-associated infections spreading between hospitals. Proc Natl Acad Sci U S A 2014; 111:2271–2276.

44. van den Dool C, Haenen A, Leenstra T, Wallinga J. The role of nursing homes ■■ in the spread of antimicrobial resistance over the healthcare network. Infect Control Hosp Epidemiol 2016; 37:761–767.

This study is one of the first to reconstruct a national healthcare network with both hospital and nursing homes to demonstrate the importance of infection control after the emergence of a new pathogen before endemicity is reached. Coupling the network with an agent-based model, the authors showed that nursing homes are sufficiently connected to the hospital network to drive national epidemics.

45. Slayton RB, Toth D, Lee BY, *et al.* Vital Signs: estimated effects of a coordinated approach for action to reduce antibiotic-resistant infections in healthcare facilities: United States. MMWR Morb Mortal Wkly Rep 2015; 64:826–831.

46. Sadsad R, Sintchenko V, McDonnell GD, Gilbert GL. Effectiveness of hospital-wide methicillin-resistant Staphylococcus aureus (MRSA) infection control policies differs by ward specialty. PloS One 2013; 8:e83099.

47. Robotham JV, Scarff CA, Jenkins DR, Medley GF. Meticillin-resistant Staphylococcus aureus (MRSA) in hospitals and the community: model predictions based on the UK situation. J Hosp Infect 2007; 65:93–99.

48. Robotham JV, Deeny SR, Fuller C, *et al.* Cost-effectiveness of national mandatory screening of all admissions to English National Health Service hospitals for methicillin-resistant Staphylococcus aureus: a mathematical modelling study. Lancet Infect Dis 2016; 16:348–356.

49. Bartsch SM, Huang SS, Wong KF, *et al.* The spread and control of norovirus outbreaks among hospitals in a region: a simulation model. Open Forum Infect Dis 2014; 1:ofu030.

50. van Bunnik BAD, Ciccolini M, Gibbons CL, *et al.* Efficient national surveillance for health-care-associated infections. BMC Public Health 2015; 15:832.

51. Deeny SR, Cooper BS, Cookson B, *et al.* Targeted versus universal screening and decolonization to reduce healthcare-associated meticillin-resistant Staphylococcus aureus infection. J Hosp Infect 2013; 85:33–44.

52. Deeny SR, Worby CJ, Auguet OT, *et al.* Impact of mupirocin resistance on the transmission and control of healthcare-associated MRSA. J Antimicrob Chemother 2015; 70:3366–3378.

53. Donker T, Wallinga J, Grundmann H. Patient referral patterns and the spread of hospital-acquired infections through national healthcare networks. PloS Comput Biol 2010; 6:e1000715.

54. Donker T, Wallinga J, Slack R, Grundmann H. Hospital networks and the dispersal of hospital-acquired pathogens by patient transfer. PloS One 2012; 7:e35002.

55. Donker T, Wallinga J, Grundmann H. Dispersal of antibiotic-resistant high-risk clones by hospital networks: changing the patient direction can make all the difference. J Hosp Infect 2014; 86:34–41.

56. Gurieva TV, Bootsma MCJ, Bonten MJM. Decolonization of patients and healthcare workers to control nosocomial spread of methicillin-resistant Staphylococcus aureus: a simulation study. BMC Infect Dis 2012; 12:302.

57. Gurieva T, Bootsma MCJ, Bonten MJM. Cost and effects of different admission screening strategies to control the spread of methicillin-resistant Staphylococcus aureus. PloS Comput Biol 2013; 9:e1002874.

58. Karkada UH, Adamic LA, Kahn JM, Iwashyna TJ. Limiting the spread of highly resistant hospital-acquired microorganisms via critical care transfers: a simulation study. Intensive Care Med 2011; 37:1633–1640.

59. Lee BY, Bartsch SM, Wong KF, *et al.* The potential trajectory of carbapenem- ■ resistant Enterobacteriaceae, an emerging threat to health-care facilities, and the impact of the centers for disease control and prevention toolkit. Am J Epidemiol 2016; 183:471–479.

Using a simulation tool that combines patient transfer data and a detailed agent-based model, the authors assess the potential spread of CRE between hospitals and nursing homes. They highlight the importance of regional coordinated surveillance infection control approaches to avoid a CRE-endemic situation.

60. Lee BY, Bartsch SM, Wong KF, *et al.* The importance of nursing homes in the spread of methicillin-resistant Staphylococcus aureus (MRSA) among hospitals. Med Care 2013; 51:205–215.

61. Lee BY, Bartsch SM, Wong KF, *et al.* Simulation shows hospitals that cooperate on infection control obtain better results than hospitals acting alone. Health Affairs (Millwood) 2012; 31:2295–2303.

62. Lee BY, McGlone SM, Wong KF, *et al.* Modeling the spread of methicillin-resistant Staphylococcus aureus (MRSA) outbreaks throughout the hospitals in Orange County, California. Infect Control Hosp Epidemiol 2011; 32:562–572.

63. Lee BY, Singh A, Bartsch SM, *et al.* The potential regional impact of contact precaution use in nursing homes to control methicillin-resistant Staphylococcus aureus. Infect Control Hosp Epidemiol 2013; 34:151–160.

64. Lee BY, Wong KF, Bartsch SM, *et al.* The Regional Healthcare Ecosystem Analyst (RHEA): a simulation modeling tool to assist infectious disease control in a health system. J Am Med Inform Assoc 2013; 20:e139–e146.

65. Lee BY, Yilmaz SL, Wong KF, *et al.* Modeling the regional spread and control of vancomycin-resistant enterococci. Am J Infect Control 2013; 41:668–673.

66. Lesosky M, McGeer A, Simor A, *et al.* Effect of patterns of transferring patients among healthcare institutions on rates of nosocomial methicillin-resistant Staphylococcus aureus transmission: a Monte Carlo simulation. Infect Control Hosp Epidemiol 2011; 32:136–147.

67. van Kleef E, Deeny SR, Jit M, *et al.* The projected effectiveness of Clostridium ■■ difficile vaccination as part of an integrated infection control strategy. Vaccine 2016; 34:5562–5570.

The authors developed an individual-based model of *C. difficile* vaccination incorporating patient transfers between the hospital, community and long-term care facilities. The model parameters come from English national ICU audit data and the authors are one of the few to validate their model. They conclude that vaccinating different target groups might have a better impact on preventing *C. difficile* cases in ICUs.

68. Bartsch SM, Huang SS, Wong KF, *et al.* Impact of delays between Clinical and Laboratory Standards Institute and Food and Drug Administration revisions of interpretive criteria for Carbapenem-resistant Enterobacteriaceae. J Clin Microbiol 2016; 54:2757–2762.

69. Meyers LA, Newman MEJ, Martin M, Schrag S. Applying network theory to epidemics: control measures for *Mycoplasma pneumoniae* outbreaks. Emerg Infect Dis 2003; 9:204–210.

70. Barnes SL, Harris AD, Golden BL, *et al.* Contribution of interfacility patient movement to overall methicillin-resistant Staphylococcus aureus prevalence levels. Infect Control Hosp Epidemiol 2011; 32:1073–1078.

71. Codella J, Safdar N, Heffernan R, Alagoz O. An agent-based simulation model for Clostridium difficile infection control. Med Decis Making 2015; 35:211–229.

72. Economoua A, Gomez-Corral A, Lopez-Garcia M. A stochastic SIS epidemic model with heterogeneous contacts. Phys A Stat Mech Appl 2015; 421:78–97.

73. Ferrer J, Boelle PY, Salomon J, *et al.* Management of nurse shortage and its impact on pathogen dissemination in the intensive care unit. Epidemics 2014; 9:62–69.

74. Hartley DM, Furuno JP, Wright MO, *et al.* The role of institutional epidemiologic weight in guiding infection surveillance and control in community and hospital populations. Infect Control Hosp Epidemiol 2006; 27:170–174.

75. Hotchkiss JR, Strike DG, Simonson DA, *et al.* An agent-based and spatially explicit model of pathogen dissemination in the intensive care unit. Crit Care Med 2005; 33:168–176.

76. Kwok KO, Davoudi B, Riley S, Pourbohloul B. Early real-time estimation of the basic reproduction number of emerging or reemerging infectious diseases in a community with heterogeneous contact pattern: using data from Hong Kong 2009 H1N1 pandemic influenza as an illustrative example. PloS One 2015; 10:12.

77. Lopez-Garcia M. Stochastic descriptors in an SIR epidemic model for heterogeneous individuals in small networks. Math Biosci 2016; 271:42–61.

78. Masuda N, Konno N, Aihara K. Transmission of severe acute respiratory syndrome in dynamical small-world networks. Phys Rev E Stat Nonlin Soft Matter Phys 2004; 69:031917.

79. Milazzo L, Bown JL, Eberst A, *et al.* Modelling of healthcare associated infections: a study on the dynamics of pathogen transmission by using an individual-based approach. Comput Methods Programs Biomed 2011; 104:260–265.

80. Rubin MA, Jones M, Leecaster M, *et al.* A simulation-based assessment of strategies to control Clostridium difficile transmission and infection. PloS One 2013; 8:11.

81. Smith DL, Dushoff J, Perencevich EN, *et al.* Persistent colonization and the spread of antibiotic resistance in nosocomial pathogens: resistance is a regional problem. Proc Natl Acad Sci U S A 2004; 101:3709–3714.

82. Temime L, Opatowski L, Pannet Y, *et al.* Peripatetic health-care workers as potential superspreaders. Proc Natl Acad Sci U S A 2009; 106:18420–18425.

83. Temime L, Kardas-Sloma L, Opatowski L, *et al.* NosoSim: an agent-based model of nosocomial pathogens circulation in hospitals. ICCS 2010 - International Conference on Computational Science 2010; 1:2239–2246.

84. Zhang XT, Ge BF, Wang Q, *et al.* Epidemic spreading characteristics and immunity measures based on complex network with contact strength and community structure. Math Probl Eng 2015; 12.

85. Smith DL, Levin SA, Laxminarayan R. Strategic interactions in multiinstitutional epidemics of antibiotic resistance. Proc Natl Acad Sci U S A 2005; 102:3153–3158.

86. Zowall H, Brewer C, Deutsch A. A model of Clostridium difficile infection: dynamic transmission between hospitals, long-term care facilities and communities. Value Health 2014; 17:A280–A281.

87. Barnes S, Golden B, Wasil E. A dynamic patient network model of hospital-acquired infections. In 2010 Winter Simulation Conference December 05–08; Baltimore, MD: IEEE: 2010:2249–2260.

88. Caudill L, Lawson B. A hybrid agent-based and differential equations model for simulating antibiotic resistance in a hospital ward. In Winter Simulation Conference on Simulation - Making Decisions in a Complex World December 08–11; Washington, DC: IEEE: 2013:1419–1430.

89. Ferrer J, Salmon M, Temime L. Nosolink: an agent-based approach to link patient flows and staff organization with the circulation of nosocomial pathogens in an intensive care unit. 2013 International Conference on Computational Science 2013; 18:1485–1494.

90. Triola MM, Holzman RS. Computer simulation of pathogen transmission in the medical intensive care unit: a comparison of two probabilistic methods. Stud Health Technol Inform 2004; 107:1277–1281.

91. Webb GF, Blaser MJ, Zhu HP, *et al.* Critical role of nosocomial transmission in the Toronto SARS outbreak. Math Biosci Eng 2004; 1:1–13.

92. Lloyd-Smith JO, Galvani AP, Getz WM. Curtailing transmission of severe acute respiratory syndrome within a community and its hospital. Proc Biol Sci 2003; 270:1979–1989.

93. Camacho A, Kucharski AJ, Funk S, *et al.* Potential for large outbreaks of Ebola virus disease. Epidemics 2014; 9:70–78.

94. Barbarossa MV, Denes A, Kiss G, *et al.* Transmission dynamics and final epidemic size of Ebola virus disease outbreaks with varying interventions. PloS One 2015; 10:21.

95. Dong FL, Xu DL, Wang Z, Dong MW. Evaluation of Ebola spreading in West Africa and decision of optimal medicine delivery strategies based on mathematical models. Infect Genet Evol 2015; 36:35–40.

96. O'Neill J. Antimicrobial resistance: tackling a crisis for the health and wealth of nations: review on antimicrobial resistance. London: 2014.

97. Jarlier V, Trystram D, Brun-Buisson C, *et al.* Curbing methicillin-resistant Staphylococcus aureus in 38 French hospitals through a 15-year institutional control program. Arch Intern Med 2010; 170:552–559.

98. Carbonne A, Arnaud I, Maugat S, *et al.,* MDRB Surveillance National Steering Group (BMR-Raisin). National multidrug-resistant bacteria (MDRB) surveillance in France through the RAISIN network: a 9 year experience. J Antimicrob Chemother 2013; 68:954–959.

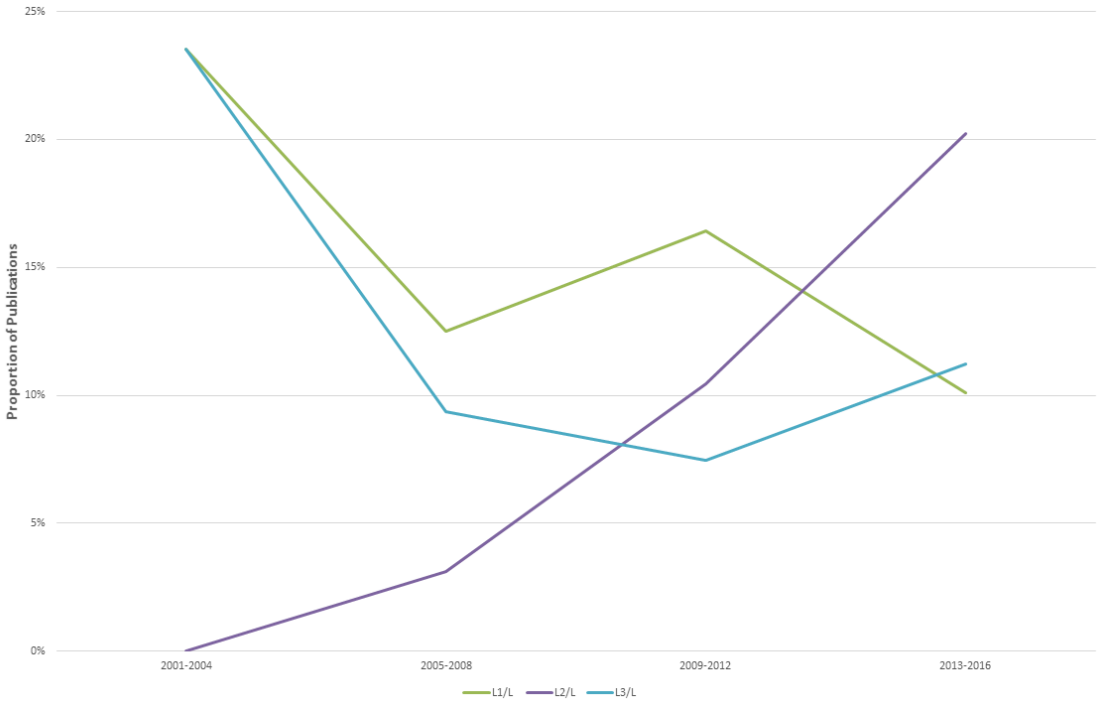
99. Arnaud I, Maugat S, Jarlier V, Astagneau P; National Early Warning, Investigation and Surveillance of Healthcare-Associated Infections Network (RAISIN)/multidrug resistance study group. Ongoing increasing temporal and geographical trends of the incidence of extended-spectrum beta-lactamase-producing Enterobacteriaceae infections in France, 2009 to 2013. Euro Surveill 2015; 20:.

100. McDonald LC, Simor AE, Su IJ, *et al.* SARS in healthcare facilities, Toronto and Taiwan. Emerg Infect Dis 2004; 10:777–781.

101. Assiri A, McGeer A, Perl TM, *et al.* Hospital Outbreak of Middle East Respiratory Syndrome Coronavirus. N Engl J Med 2013; 369:407–416.

102. Lee SS, Wong NS. Probable transmission chains of Middle East respiratory syndrome coronavirus and the multiple generations of secondary infection in South Korea. Int J Infect Dis 2015; 38:65–67.

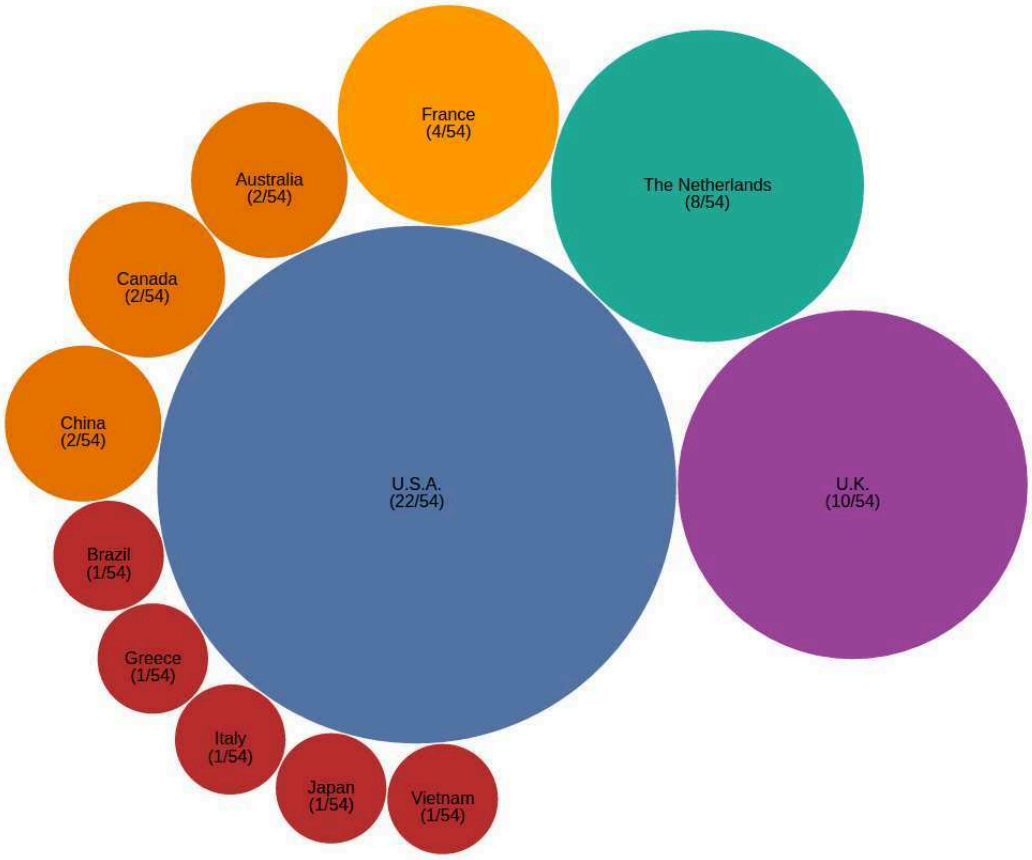
Suppl Figure1



Assab* R, Nekkab* N, Crépey P, Astagneau P, Guillemot D, Opatowski L, Temime L. Using data on network structures to inform transmission dynamics of infections in healthcare settings: a review of mathematical models. Published August 2017 (*Current Opinion in Infectious Diseases*)

* *the first authors contributed equally to the paper*

Suppl Figure2



Assab* R, Nekkab* N, Crépey P, Astagneau P, Guillemot D, Opatowski L, Temime L. Using data on network structures to inform transmission dynamics of infections in healthcare settings: a review of mathematical models. Published August 2017 (*Current Opinion in Infectious Diseases*)

* *the first authors contributed equally to the paper*

Part Three: The French healthcare networks

Chapter 7. The French medico-administrative database (PMSI)

Here we present the exhaustive database of patient discharge summaries that was exploited for construction of the French healthcare networks of patient transfers. In addition, a literature review assessing the use of this database for epidemiological research on HAIs was conducted and the results are described.

7.1 Description of the database

The Programme de Médicalisation des Systèmes d'information (PMSI) database or Program for the Medicalization of Information Systems database in English (also known as the French hospital discharge database) is a comprehensive French medico-administrative database of hospital activity and patient discharge information.(178-180) Originally the PMSI was created for calculating hospital activity for budget regulation and attribution; however, it has also been used for epidemiological and medical research not related to cost analysis.(181) The database is composed of patient discharge summaries in the form of the French equivalent of diagnosis-related groups (DRGs): one primary diagnosis, one related diagnosis and up to 30 associated diagnoses coded according to the International Classification of Diseases (ICD) tenth revision (ICD-10).(180, 182) Each patient discharge summary contains information on the age, gender, place of residence in the form of municipality zip code, the month and year of admission and discharge, and the healthcare facility of stay in addition to the diagnoses. Each healthcare facility is identified by a unique FINESS number (Fichier National des Etablissements Sanitaires et Sociaux). Each patient stay is also numbered by order of stay across different healthcare facilities. The PMSI is the largest French database on medical information in terms of healthcare facility coverage.(179)

Since 2003, the PMSI has been part of a larger warehouse of data along with French health insurance databases that together compose the National Health Insurance Information System (Système national d'information inter-régime de l'assurance maladie (SNIIR-AM)).(180) The SNIIR-AM includes inter-scheme consumption data (données de consommation inter-régimes (DCIR)) that include all outpatient reimbursed health expenditures. These data are managed by the National Health Insurance Fund for Salaried Workers (Caisse nationale de l'assurance Maladie des travailleurs salariés (CNAMTS)).(180)

In 2004, the financing of health facilities was drastically changed to an activity-based pricing system (T2A) where hospital budgets depended on the number of treated patients based on their disease.(182) This applied only to activities under the medical, surgical, and obstetric (MCO)

departments while the psychology and postoperative care and rehabilitation (SSR) were still financed by an annual grant. As a result, the PMSI database is split into three parts: MCO, SSR patient discharge summaries, and home hospitalisation (hospitalisation à domicile (HAD)). Hospitals send their discharge summaries to the regional health agencies which in turn send the data to the French Technical Agency on Hospitalisation Information (Agence Technique de l'Information sur l'Hospitalisation (ATIH)) which owns the data and makes it accessible for research purposes.(180)

7.2 HAI detection in the database: a literature review

Because the PMSI database includes ICD-10 codes for the discharge diagnosis of patients, it could be expected that identifying patients who were diagnosed with an HAI will be possible from this database; however, the specific Y95 code for nosocomial infections is not systematically used and other ICD-10 codes may be either not sensible or not specific enough. Hence, a literature review was conducted to evaluate the validity of the PMSI database for epidemiology research on HAIs and to assess appropriate inclusion criteria for HAI-related diagnoses. The inclusion criteria for the review required studies to have evaluated the PMSI database for general HAIs or specific HAIs in terms of sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), or used a test statistic comparing the database to another source of HAI data. The review identified a total of 34 publications (as of April 1st 2018) in the PubMed search engine of the MEDLINE database using the following search terms:

(PMSI[Title/Abstract] OR "Programme de Médicalisation des Systèmes d'information"[Title/Abstract] OR HDD[Title/Abstract] OR "hospital discharge database"[Title/Abstract] OR "hospital information system"[Title/Abstract] OR "medico-administrative database"[Title/Abstract])

AND

(nosocomial[Title/Abstract] OR infect*[Title/Abstract] OR pathogen[Title/Abstract] OR pathogens[Title/Abstract] OR "hospital-acquired"[Title/Abstract] OR "hospital acquired"[Title/Abstract] OR "hospital-associated"[Title/Abstract])*

AND

("French"[Title/Abstract] OR "France"[Title/Abstract])

A total of seven publications met the inclusion criteria.(18, 182-187) The results of these studies are summarized in Table 3.

Table 3. Summary of publications assessing HAI detection in the PMSI database.

Reference	Objective	Study population	Validity of PMSI database	Evaluation measure
Legras et al. 1993(183)	Cross-sectional study comparing biological lab results of ICU to PMSI database of HAIs	8821 patients, 13692 strains in CHU Nancy from 1989-1991	Lab data and PMSI not sufficient for extensive understanding of HAIs; PMSI not specific and only 50% coverage	Urology unit: Se 65%, Sp 99 %
Bouzbid et al. 2011(182)	Cross-sectional study evaluating seven strategies for detection of HAIs in ICU	1499 patients in 9 adults ICUs in Hospices Civils de Lyon between 2000-2006	PMSI had lowest sensitivity and specificity for all HAIs among strategies; additional use of PMSI did not improve sensitivity overall but had highest specificity for UTI	PMSI only: Se 49% and Sp of 78% Combined with other strategies: Se 99%, Sp 54%
Gerbier et al. 2011(18)	Cross-sectional study evaluating PMSI to detect HAIs	General surgery (446 patients, 2007), ICU (1499 patients, 2000-2006), obstetrics (1081 patients, 2006) in Hospices Civils de Lyon	PMSI not sufficiently efficient method in terms of sensitivity and specificity of surveillance of HAIs	SSIs: Se 26%, Sp 99.5% Expanded SSI codes: Se 79%, Sp 66% ICU: Se 48%, Sp 78% PPI: Se 43%, Sp 87%
Gerbier-Colomban et al. 2012(184)	Cross-sectional study assessing different strategies for detecting SSI using a combination of different sources	446 patients in 2007 in Hospices Civils de Lyon surgical unit	Addition of discharge diagnosis codes was a complement to detect SSI when the patient was already discharged	PMSI only: Se 26.3%, Sp 99.5% Combined: Se 87%, Sp 86%
Nuemi et al. 2013(185)	Cross-sectional study evaluating PMSI data on MRSA incidence compared to national surveillance data (InVS)	Average of 22.7 million hospital stays and 11.6 million patients in France 2006-2009; 1417 MRSA infections	Recommend use of the PMSI data as an additional source of information in the hospital MRSA surveillance process	Incidence density measured by InVS higher than MRSA density using PMSI, difference appeared to decrease over time; year of study significantly associated with incidence density in PMSI (P = 0.01)
Grammatico-Guillon et al. 2016(186)	Case-control study assessing the efficacy of PMSI as a routine surveillance system for detecting hip or knee arthroplasty–related infections	1010 hospital stays of 989 patients (530 cases, 480 controls) from 2008-2010 in 23 French hospitals	Potential of PMSI as tool for routine SSI surveillance after low-risk surgery, under conditions of appropriate algorithm for selecting infections	Definition A: Se 97%, Sp 95%, PPV 87%, NPV 98% Definition A, B: Se 98%, Sp 83%, PPV 72%, NPV 99% Definition A, B, C: Se 98%, Sp 71%, PPV 63%, NPV 99%

Sahli et al. 2016(187)	Cross-sectional study evaluating the positive predictive values of infection codes as reason for hospitalization as coded in the PMSI	4545 hospital stays in CHU Toulouse in 2014	Positive predictive values of overall infections leading to hospitalization in general population very high in PMSI; details of the codes were closely consistent with type of infection that occurred	Primary diagnosis: PPV 97% Related diagnosis: PPV 70%

Publications were identified using the MEDLINE database (up to April 1st 2018)
Se: sensitivity
Sp: specificity
PPV: positive predictive value
NPV: negative predictive value
InVS: Institut de Veille Sanitaire (French Institute for Public Health Surveillance)
PPI: postpartum infections

In 1993, Legras et al. found that the PMSI was not specific enough in identifying HAIs.(183) Bouzbid et al. confirmed these observations arguing that the use of PMSI in HAI studies did not improve sensitivity but they highlighted the advantage of using multiple existing data systems to improve detection and monitoring of HAIs.(182)

In 2001, Gerbier et al. conducted an extensive sensitivity and specificity analysis using various diagnostic codes for SSI and ICU-HAIs in the database and comparing them to lab confirmed surveillance data.(18) Gerbier et al. used patient discharge summaries from the PMSI to detect nosocomial infections in the University Hospital of Lyon in 2006 and 2007 for the identification in surgery, intensive care and obstetric units. The authors compared the PMSI data to a gold standard by doing a systematic review of patient files for those classified under surgery, the Cpias/CClin Southwest surveillance network for ICU patients, and a combination of surveillance data from Cpias /CClin and patient information data for obstetrics. The ICD-10 codes that were evaluated were those recommended by the surveillance network. For SSIs, the authors also tested the impact of expanding the codes of potential SSIs in the database to assess the performance of these codes on the sensitivity and specificity. Additional codes to those recommended by the surveillance system for the ICU and obstetric units were not evaluated. A summary of the results from the Gerbier et al. study (18) are shown in Table 4.

Table 4. Performance of HAI case detection in PMSI database by Gerbier et al.

Unit type and HAI sites	Sensitivity (95% CI)	Specificity (95% CI)
Surgical unit with limited ICD-10 codes for SSIs	26.3% (13.2-42.1)	99.5% (98.8-100)
Surgical unit with additional ICD-10 codes for SSIs	78.9% (65.8-92.1)	78.4% (76.1-80.1)
ICU, all sites	48.8% (42.6-55.0)	78.4% (76.1-80.1)
Obstetrics, all sites	42.9% (25.0-60.7)	87.3% (85.2-89.3)

The additional codes for SSI detection were able to improve the sensitivity from 26.3% to 78.9%; however, the specificity reduced from 99.5% to 78.4%. The authors found in another study that the sensitivity and specificity of SSIs improved to a sensitivity of 87% and specificity of 86% when the PMSI was used in combination with other strategies.(184)

Other authors evaluated the ability of the PMSI to detect MRSA in comparison to data from national surveillance network and found that the incidence density measured by the national surveillance network was higher than the MRSA density calculated using the PMSI databases.(185) A case-control study assessed the efficacy of PMSI as a routine surveillance system for detecting hip or knee arthroplasty–related infections and found that the PMSI was useful for low-risk surgery and under certain conditions when an appropriate algorithm for selecting infections was used.(186) A cross-sectional study evaluating the positive predictive values of infection codes as reason for hospitalization in 2016 found a 97% PPV for primary diagnoses and PPV of 70% for related diagnosis for all infections, including some of those classified as sensitive and specific to HAIs in the Gerbier et al. study.(18, 187)

Although they did not meet the criteria for the review, Fourquet et al. conducted a feasibility study of the PMSI on HAI detection in 2003.(179) Fourquet et al. argued that the PMSI was not suitable for epidemiological studies on HAIs due to the fact that the two criteria for HAIs (an infection absent during admission that appears at least 48 hours after admission) and date of the medical procedures were absent in the database. In addition, the ICD-10 Y95 code was hardly used at the time and did not have any impact on the DRGs. However, they found that it was a useful tool in that it has existed for many years, was readily available, and provided exhaustive data. The authors also compiled a list of HAI-related codes in addition to the ICD-10 Y95 code recommended for better detection of HAIs in the database that was complemented by the Gerbier et al. study.(18, 179) Other studies also evaluated the PMSI for its diagnostic accuracy included studies on tuberculosis (188) and encephalitis (189) but were not included in the review.

Several conclusions can be made about the ability of the PMSI to detect HAIs from this review: although the PMSI lacks precise medical information on HAIs such as the recorded date of HAI events and source of infection, it does have the advantage of providing immediately available data from the course of many years of reporting and provides a detailed account of hospital stays; the performance of the PMSI (in terms of sensitivity, specificity, PPV, and NPV) to detect either all infections or hip or knee arthroplasty–related infections has been shown to be robust when using the studies' algorithms (186, 187, 190); however, the performance of the PMSI to

detect HAIs was sufficient only when investigators compiled a more comprehensive list of diagnoses since the Y95 code did not have robust sensitivity and specificity alone.(18, 179) In addition, in some studies the added value of the PMSI was only observed when it was combined with other strategies of HAI detection such as the national surveillance system data.(182, 184) It should be noted that the PMSI does not necessarily cover asymptomatic carriage of common HAI pathogens. In addition, discharged patients initially diagnosed with an HAI may have been cleared of their infection before being transferred. Therefore, the PMSI was found to provide a large set of data that could contribute to the better understanding of the epidemiology of certain HAIs; however, for the most robust assessment of HAIs in the database, a more comprehensive algorithm of HAI-sensitive and specific diagnoses should be applied and limitations concerning undetected pathogen carriage should be considered.

Chapter 8. Second article – Analysis of the French healthcare networks

8.1 Summary

In recent years, the utility of reconstructing national healthcare networks in order to better inform healthcare pathogens and MDRO spread dynamics and infection control practices has been demonstrated.(17, 110) These studies have pointed to the advantages using healthcare networks to identify healthcare facilities in order to develop more effective infection control policy. Some examples include developing more effective sentinel surveillance systems, improving coordinated regional control efforts, and identifying hospital communities and hospitals that connect patients and MDROs that are geographically distant from one another. However, studies of national healthcare networks have relied on the general patient population in order to inform their mathematical models of pathogen spread. These models may have overlooked the specificities of patients who may be at higher risk of acquiring healthcare setting pathogens and whose transfer patterns may differ from that of the general patient population. Therefore, our objective was to assess the potential differences between a general patient healthcare network and a healthcare network of HAI-diagnosed patients in France. Thus, to better understand how to reduce the potential scale of HAI epidemic spread, we explored patient transfer patterns using the PMSI database of all hospital discharge summaries in France for the year 2014.(191) We constructed and analysed three patient transfer networks: transfers of patients with a HAI; of patients with a suspected HAI; and of all patients.

We found that all three networks had heterogeneous patient flow and demonstrate small-world and scale-free characteristics, meaning that a small number of university hospitals had very high connectivity to other healthcare facilities in France and that theoretically patients required on two to three transfers in order to be admitted to any hospital in France. Patient populations that comprised these networks were also heterogeneous in their movement patterns. Ranking of hospitals by centrality measures, comparing community clustering using community detection algorithms, and comparing the networks to a "null model" of random healthcare networks showed that despite the differences in patient population, the HAI-specific and suspected-HAI networks relied on the same underlying structure as that of the general network. We were able to identify transfer patterns at both the French regional and sub-regional levels that could be important in the identification of key healthcare facility clusters that may serve as a basis for novel wide-scale infection control strategies. In conclusion, our study found that the general network was potentially more reliable in studying potential spread of pathogens since the structure of the network did not differ significantly from the HAI networks.

PLOS COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

Spread of hospital-acquired infections: A comparison of healthcare networks

Narimane Nekkab[1,2,3]*, Pascal Astagneau[3,4,5], Laura Temime[1,2], Pascal Crépey[3,6,7]

1 Laboratoire MESuRS, Conservatoire National des Arts et Métiers, 292 Rue Saint-Martin, Paris, France, **2** Institut Pasteur, Cnam, Unité PACRI, 25–28, rue du Docteur Roux, Paris, France, **3** Ecole des Hautes Etudes en Santé Publique, Département d'Epidémiologie et de Biostatistiques, 15 Avenue du Professeur-Léon-Bernard, Rennes, France, **4** Centre de prévention des infections associées aux soins (C-CLIN), APHP, Paris, France, **5** Faculté de médecine Pierre et Marie Curie, Sorbonne Universités, Paris, France, **6** UMR190, Emergence des Pathologies Virales, Marseille, France, **7** UPRES EA 7449 Reperes, Rennes, France

* narimane.nekkab@cnam.fr

Citation: Nekkab N, Astagneau P, Temime L, Crépey P (2017) Spread of hospital-acquired infections: A comparison of healthcare networks. PLoS Comput Biol 13(8): e1005666. https://doi.org/10.1371/journal.pcbi.1005666

Editor: Mark M. Tanaka, University of New South Wales, AUSTRALIA

Received: January 6, 2017

Accepted: July 3, 2017

Published: August 24, 2017

Data Availability Statement: We were given access to the PMSI database owned by the French Agence Technique de l'Information sur l'Hospitalisation (ATIH) after meeting the criteria for access determined by the Commission nationale de l'informatique et des libertés (CNIL). We do not legally own the data and it cannot be made publicly available for legal reasons as public availability would compromise the privacy of the 2.3 million patients. For those wishing to have access to the database, the request can be sent to demande_base@atih.sante.fr and more

Abstract

Hospital-acquired infections (HAIs), including emerging multi-drug resistant organisms, threaten healthcare systems worldwide. Efficient containment measures of HAIs must mobilize the entire healthcare network. Thus, to best understand how to reduce the potential scale of HAI epidemic spread, we explore patient transfer patterns in the French healthcare system. Using an exhaustive database of all hospital discharge summaries in France in 2014, we construct and analyze three patient networks based on the following: transfers of patients with HAI (HAI-specific network); patients with suspected HAI (suspected-HAI network); and all patients (general network). All three networks have heterogeneous patient flow and demonstrate small-world and scale-free characteristics. Patient populations that comprise these networks are also heterogeneous in their movement patterns. Ranking of hospitals by centrality measures and comparing community clustering using community detection algorithms shows that despite the differences in patient population, the HAI-specific and suspected-HAI networks rely on the same underlying structure as that of the general network. As a result, the general network may be more reliable in studying potential spread of HAIs. Finally, we identify transfer patterns at both the French regional and departmental (county) levels that are important in the identification of key hospital centers, patient flow trajectories, and regional clusters that may serve as a basis for novel wide-scale infection control strategies.

Author summary

Hospital-acquired infections (HAIs), including emerging multi-drug resistant organisms, threaten healthcare systems worldwide. Efficient containment measures of HAIs must mobilize the entire healthcare network. Thus, to best understand how to reduce the scale of potential HAI epidemic spread, we explore patient transfer patterns in the French healthcare system. We construct and compare the characteristics of three different patient

information on the ATIH can be found on their website www.atih.sante.fr

Funding: This work was supported by the Interdisciplinary research program on health crisis and health protection (PRINCEPS) of Sorbonne Paris Cité University, within the program "Investissements d'Avenir" launched by the French State (http://www.ehesp.fr/recherche/domaines-de-recherche/securite-sanitaire/princeps-programme-de-recherche-interdisciplinaire-sur-les-crises-et-la-protection-sanitaire/). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

transfer networks based on data on transfers of patients with diagnosed HAIs, suspected HAIs, or of all patients. Our analyses show that these healthcare networks, the patient populations that comprise them and the patient movement patterns are heterogeneous and centralized. Despite the differences in patient populations, the HAI-specific and suspected-HAI healthcare networks have the same underlying structure as that of the general healthcare network. We identify key hospital centers, patient flow trajectories, at both the regional and department (county) level that may serve as a basis for proposing novel wide-scale infection control strategies.

Introduction

The emergence and spread of multi-drug resistant organisms threatens healthcare systems worldwide.[1] This is particularly true concerning methicillin-resistant *Staphylococcus aureus*, vancomycin-resistant enterococci, and multi-resistant gram-negative bacteria such as carbapenemase-producing Enterobacteriaceae (CPE). Spread of CPE is now a global public health problem associated with patient transfers between healthcare facilities within the same country as well as across national borders, as shown in many recent studies.[2–7]

In recent years, patient transfer or referral data has been used to construct "healthcare networks" to propose innovative approaches for hospital infection prevention and control. Healthcare networks are cooperative healthcare systems where hospitals and other healthcare centers are linked by shared patients through secondary transfers or referral.[8, 9] Rather than being exclusive to one sole hospital, as Ciccolini et al. argue, the extent of hospital-acquired infection (HAI) spread is dependent on the healthcare network connected by inter-institutional patient transfers.[8] Heterogeneous hospital patient populations and the interactions that occur between them and with the community are important in the understanding of the spatial spread of HAI between hospitals across geographic regions.[9]

As early as 2007, studies applied more complex social network analysis approaches to reconstructed healthcare networks in order to demonstrate that infection control measures that take into account network properties can decrease the risk for outbreaks.[8, 10] Lee et al. consider network properties to assess the individual influence of different hospitals and the impact of hospital proximities on HAI spread on a regional scale.[11] Many studies show that healthcare networks display a community structure.[8, 12–14] Network analysis is especially effective in the identification of sensor hospitals for surveillance of HAIs.[15, 16] In addition, mathematical models of healthcare networks may serve to inform decision-makers on enhanced coordinated regional and national approaches to infection control strategies, in a context where increasingly centralized healthcare systems favor the spread of HAIs.[8, 15, 17]

Although national healthcare networks are informative regarding novel HAI control strategies, the impact of reconstructing these networks based on a general patient population rather than a HAI-diagnosed patient population has rarely been addressed. In this study, we assess and compare French healthcare networks based on either patients diagnosed with HAIs or the general patient population, in order to better understand the potential implications in terms of HAI spread predictions. To that aim, we perform social network analyses to describe the different patient flow patterns, network topology characteristics, and community clustering structure.

Results

We analyzed and compared three different networks built using transfer data from an exhaustive database of all hospital discharge summaries in France in 2014: (1) a network based on

transfers of patients with diagnosed HAI (HAI-specific network); (2) a network based on transfers of patients with suspected HAI (suspected-HAI network); and (3) the network of all patient transfers (general network).

Characteristics of healthcare networks

More than 10 million hospital transfers were recorded in France in 2014, for a total of 2.3 million transferred patients, creating a hospital network of 2063 hospitals (nodes) and 50026 patient trajectories (edges) linking them (Table 1). Patients with a HAI-specific diagnosis created a healthcare network of 1266 hospitals and 3722 connections for 13627 patient transfers. A larger population of patients suspected to have an HAI infection formed a healthcare network of 1975 hospitals and 18812 connections for a total of 128681 patient transfers. With the increasing number of patient transfers, the networks increased from an average 5.88, 19.05, and 48.05 average connections per hospital (average degree $\bar{k}$) and an average 2.31, 4.92 to 14.02 transfers per connection (average strength $\bar{s}$) for the HAI-specific, suspected-HAI, and general healthcare networks respectively (Table 1).

Overall, the three networks displayed "scale-free" and "small-world" characteristics that indicated the presence of a small number of very highly connected hospitals with high degrees, referred to as "hubs." Analyses of the degree, strength, and shortest path length distributions in addition to the small-world characteristics of the healthcare networks are discussed in S1–S3 Texts and S1–S7 Figs. Compared to random networks, we also showed the general network was more clustered and efficient in transferring patients (S4 Text, S1 Table). We identified several high degree hospitals in all three networks with a consistent outlier–the Assistance Publique—Hôpitaux de Paris (AP-HP)–a conglomerate of 39 hospitals predominately in Paris and the Ile-de-France region represented as one hospital code in our database.[18] AP-HP also acted as the most important intermediary hospital system in the networks due to having the highest betweenness centrality measure.

Table 1. Networks characteristics of the French healthcare networks.

Network Characteristics	General Network	Suspected-HAI Network	HAI-Specific Network
Patients	2300728	394859	21279
Patient Transfers	1033239	128681	13627
Hospitals	2063	1975	1266
Patient Trajectories*	50026	18812	3722
Average Edge Weight**	14.02	4.92	2.31
Average Degree***	48.50	19.05	5.88
Average In-Degree	24.25	9.53	2.94
Average Out-Degree	24.25	9.53	2.94
Average Betweenness***	5292.32	6338.81	3824.91
Average Edge Betweenness	301.27	852.23	1556.94
Average Closeness***	0.00016	0.000074	0.000032
Diameter	30	64	47
Average Path Length	2.99	3.63	5.23
Global Clustering Coefficient	0.23	0.16	0.08
Density	0.012	0.005	0.002

* Also referred to as edges, they represent the sum number of connections between the hospitals

** The average number of patients per trajectory

*** Measures of node (or hospital) centrality

https://doi.org/10.1371/journal.pcbi.1005666.t001

Table 2. Healthcare facility types among the general, suspected-HAI, and HAI-specific networks and their hub hospitals.

Health Facility Type		General		Suspected-HAI		HAI-Specific	
		All facilities	Among hubs* (N = 103)	All facilities	Among hubs* (N = 99)	All facilities	Among hubs* (N = 63)
Type 1**	SSR***	38.00%	0.97%	37.22%	2.02%	35.31%	1.69%
	MCO**** & SSR	36.74%	66.99%	38.13%	64.65%	44.63%	88.14%
	MCO	25.25%	32.04%	24.66%	33.33%	20.06%	10.17%
Type 2†	Private hospitals authorized to provide SSR services	30.63%	0	29.90%	1.03%	30.76%	1.72%
	Acute-care hospitals or clinics	28.72%	30.69%	28.83%	31.96%	25.04%	5.17%
	Hospital centers	23.50%	32.67%	24.44%	27.84%	30.29%	31.03%
	Local hospitals	10.75%	0	10.77%	0	6.76%	0
	University hospital centers††	1.47%	27.72%	1.53%	29.90%	2.38%	48.28%
	Nursing home	1.27%	0.99%	1.22%	0	1.35%	0
	Cancer centers	0.93%	3.96%	0.97%	4.12%	1.27%	3.45%
	Other health facilities practicing under the healthcare law	0.59%	0	0.56%	2.06%	0.48%	1.72%
	Armed forces hospitals	0.44%	2.97%	0.46%	3.09%	0.72%	8.62%
	Long-term care hospitals	0.39%	0.99%	0.41%	0	0.24%	0
	Other facilities for mental health	0.39%	0	0.26%	0	0.08%	0
	Medical homes for handicapped adults	0.34%	0	0.26%	0	0.32%	0
	Hospital centers specialized in mental health	0.24%	0	0.10%	0	0.08%	0
	Home care facilities	0.20%	0	0.15%	0	0.16%	0
	Outpatient dialysis centers	0.10%	0	0.10%	0	0	0
	Home care or outpatient care for the handicapped	0.05%	0	0.05%	0	0.08%	0

The percent of different health facilities in the networks by Type 1 and Type 2 classification are given.

* Hubs are defined as facilities that comprise the top 5% of facilities by degree

** Type 1 refers to categorization of the general activities performed in the facility

*** SSR = postoperative and rehabilitation care (soins de suite et de réadaptation)

**** MCO = medical, surgery, and/or obstetrics care (médecine—chirurgie—obstétrique)

† Type 2 refers to the categorization of the facilities by hospital type or specialized services

†† Often referred to as regional hospital centers (centre hospitalier régionale)

https://doi.org/10.1371/journal.pcbi.1005666.t002

The hospitals involved in the patient transfers recorded in the three networks were of various types, including private rehabilitation and postoperative care facilities, acute-care hospitals or clinics, and hospital centers (Table 2). However, the majority of hubs, defined as the top 5% of hospitals by their degree, were large hospitals providing both acute and postoperative or rehabilitation care (67%, 65%, and 88% in the general, suspected-HAI, and HAI-specific networks respectively). In addition, in the general and suspected-HAI networks, hubs were mostly acute-care hospitals or clinics, hospital centers, or university hospitals centers, with many concentrated in the Ile-de-France, Marseille, and Lyon metropoles (31%, 33%, 28%, and 32%, 28%, 30% respectively). In contrast, university hospital centers rather than acute-care facilities dominated the hub hospitals of the HAI-specific network, representing 48% of hubs (Table 2). The hub university healthcare centers, which provided highly specialized services, included the AP-HP, Hospices Civils de Lyon, and the Assistance Publique—Hôpitaux de Marseille (AP-HM); among them there were also university hospitals of other major cities in France.

To better understand the role of hub hospitals across the networks, the shared hospitals between the networks were ranked based on their degree, closeness, and betweenness (Fig 1).

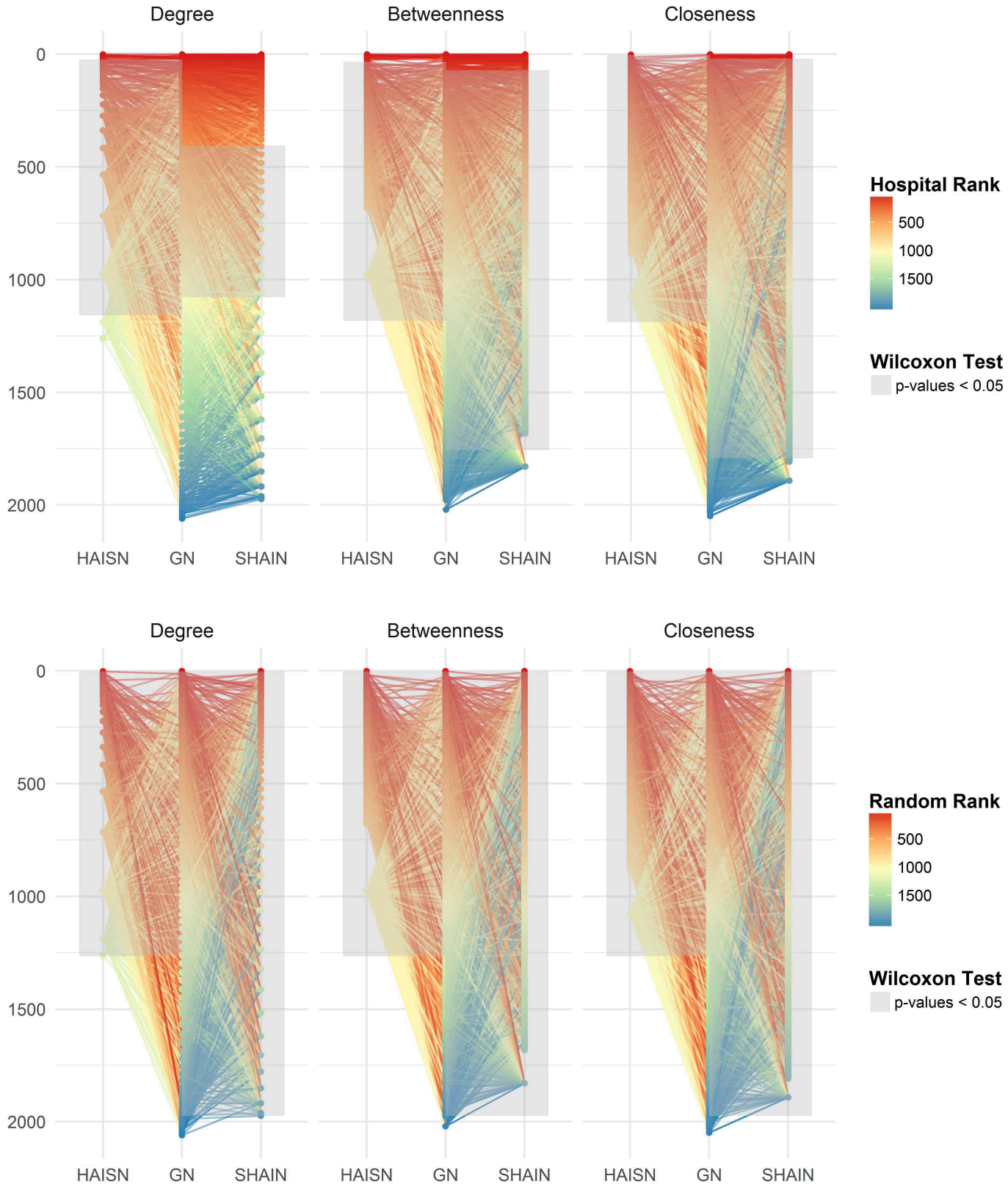


Fig 1. Hospital rankings by degree, betweenness, and closeness across the networks. Hospitals in the HAI-specific network (HAISN) (n = 1266), suspected-HAI network (SHAIN) (n = 1975), and general network (GN) (n = 2063) are displayed vertically and plotted against their ranking by degree, betweenness, and closeness centrality measures (top row). Only the hospitals shared between the HAISN and GN or the SHAIN

and GN were linked. The color gradient refers to the hospital ranking for each centrality measure and the line colors correspond to the rankings of the hospitals in the GN. We tested the differences in rankings by Wilcoxon rank sum test of an increasing subset of hospital degrees starting from the highest rank and adding each consecutive rank and retesting. The grey area represents the range where the HAISN or SHAIN differed from the general network hospital rankings. We chose rankings at random for the hospital degrees, betweenness, and closeness centrality measures for comparison (bottom row). All random rankings were statistically different across the centrality measures between the HAISN and GN and the SHAIN and GN shared hospitals.

https://doi.org/10.1371/journal.pcbi.1005666.g001

Overall, when comparing the degree, betweenness, and closeness, the hospital rankings did not differ between the complete set of 1266 HAI-specific network hospitals and these same hospitals in the general network ($p = 0.81$, $p = 1$, $p = 0.99$ respectively, Wilcoxon rank sum test), or between the 1975 suspected-HAI network hospitals and the same hospitals in the general network ($p = 0.99$, $p = 1$, $p = 0.99$, Wilcoxon rank sum test). For comparison and illustration purposes, we showed that random rankings for degree, betweenness, and closeness of all hospitals differed significantly between patient specific networks and the general network ($p < 0.05$ respectively, Wilcoxon rank sum test) (Fig 1).

Suspecting that the differences between rankings might exist between subsets of hospitals, we tested the differences between rankings on an increasing subset of shared hospitals, starting with the highest rank, adding the next ranked hospital, and testing for significant differences. As a result, we determined the range of hospital rankings across the networks where the rankings significantly differed. We defined significant differences as Wilcoxon rank sum test p-values under the 5% alpha risk which we represent as a grey area in Fig 1. Distributions of these p-values are provided in S8 and S9 Figs. For the HAI-specific network, the range of statistically significant degree ranking differences were observed between the 24th ranked hospital to the 1159th ranking hospital. For the suspected-HAI network, statistically significant degree ranking differences were observed between the 405th ranked hospital to the 1078th ranked hospital.

For hospital rankings based on betweenness and closeness centrality measures, the hospitals ranked with highest and lowest centralities in the general network were also the hospitals ranked with highest and lowest centralities ranking in the HAI-specific and suspected-HAI networks. Even though hospital rankings of all hospitals did not differ, the majority did differ for betweenness ranks between the 33rd highest ranking to the 1183rd ranking in the HAI-specific network and the 71st highest ranking to the 1757th ranking in the suspected-HAI network ($p < 0.05$, Wilcoxon rank sum test). Closeness rankings differences were observed for almost all rankings after the first 3 rankings in the HAI-specific network and after the first 6 in the suspected-HAI network. The lack of statistically significant differences for the highest rankings may have been only due to insufficient power and for lowest hospital rankings due to a series of repeating small closeness values. With this method, we highlight that differences do exist for subsets of hospitals, but we also observe that the most highly connected hub hospitals were consistently highly connected across the networks, irrespective of the different patient population that connected them.

What community structures in patient sharing are observed?

To further assess patient movement patterns in the networks, we investigated how our healthcare networks displayed "community" or hospital clustering structure. We compared hospital communities detected with two different community clustering algorithms: 1) the Greedy algorithm [19] that selected members of the communities to maximize the density of links between vertices as it reconstructed the network one vertex at a time and 2) the Map Equation algorithm [20], based on network structure-induced movement using a flow-based and information-theoretic method, detecting communities by measuring probability flows by taking into consideration the directionality and weight of the edges. In general, we detected fewer

Table 3. Community clustering distance.

	General Network	Suspected-HAI Network	HAI-Specific Network
Map Equation algorithm			
Modularity	0.764	0.716	0.698
Number of communities	132	160	193
Average community size	15.63	12.34	6.56
Average community clustering distance (km)	30.51	23.63	22.86
Greedy algorithm			
Modularity	0.863	0.847	0.830
Number of communities	18	21	36
Average community size	114.61	94.05	35.17
Average community clustering distance (km)	39.01	41.60	31.40

Two community detection algorithms were used to assess community clustering for each network, both of which take into account weighted graphs. The Greedy algorithm, developed by Clauset et al.[19] optimized modularity; however, it applied only to non-directed graphs. The Map Equation[20] algorithm applied to directed graphs and detects communities based network structure-induced movement using a flow-based and information-theoretic method. Average community size refers to the average number of hospitals within a detected community. For each community, the clustering distance in kilometers was calculated as the average geographic distance between pairs of hospitals of the same community.

https://doi.org/10.1371/journal.pcbi.1005666.t003

communities with the Greedy algorithm given that it seeks to maximize modularity–a value that measures the density of links inside communities by comparing the fraction of edges within the communities to the fraction in a random network; a maximum value of 1 corresponds to a network structure with the highest strength possible–as a result, the algorithm searched for the repartitions that maximized the density of the edges.[21–23] The Greedy algorithm considered pairwise interactions and the formation of the network whereas the Map Equation considered the interdependence of links and the dynamics of an already formed network.

For each network, we calculated the modularity, the number of communities, community size, and average community clustering distance using the Greedy and Map Equation community detection algorithms (Table 3). For each community, the pairwise clustering distance was calculated as the average geographic distance between all pairs of hospitals of the same community in kilometers. Compared to the general healthcare network, the patient-specific networks had more communities. In the HAI-specific network, there were on average 35.17 hospitals per community (SD = 44.31) and 31.40 kilometers between pairs of hospitals in the same community (SD = 25.60 km). In the larger networks, the larger community sizes resulted in a higher average distance between community hospitals (41.60 km (SD = 34.71) and 39.01 km (SD = 45.63) for the suspected-HAI and general healthcare network respectively). For the Map Equation-based communities, as the number of communities decreased from the HAI-specific to suspected-HAI to the general healthcare network, the average community size and average community distance between hospitals of the same community increased (Table 3). Overall, the suspected-HAI network was more similar to the general network than the HAI-specific network in terms of community structure (S5 Text).

The regional community clustering using the Greedy algorithm in the three networks are represented in Fig 2. The hospitals in communities were geo-localized, color-coded, and labelled across the networks according to the administrative region(s) they encompassed. We observed that the Greedy-based communities accurately reflected the French regional administrative structure (Fig 2). The identified community clusters formed hospitals communities in which most of the patients were shared between hospitals of the same region frequently centralized towards the hub acute-care centers, university hospital centers, and general hospital

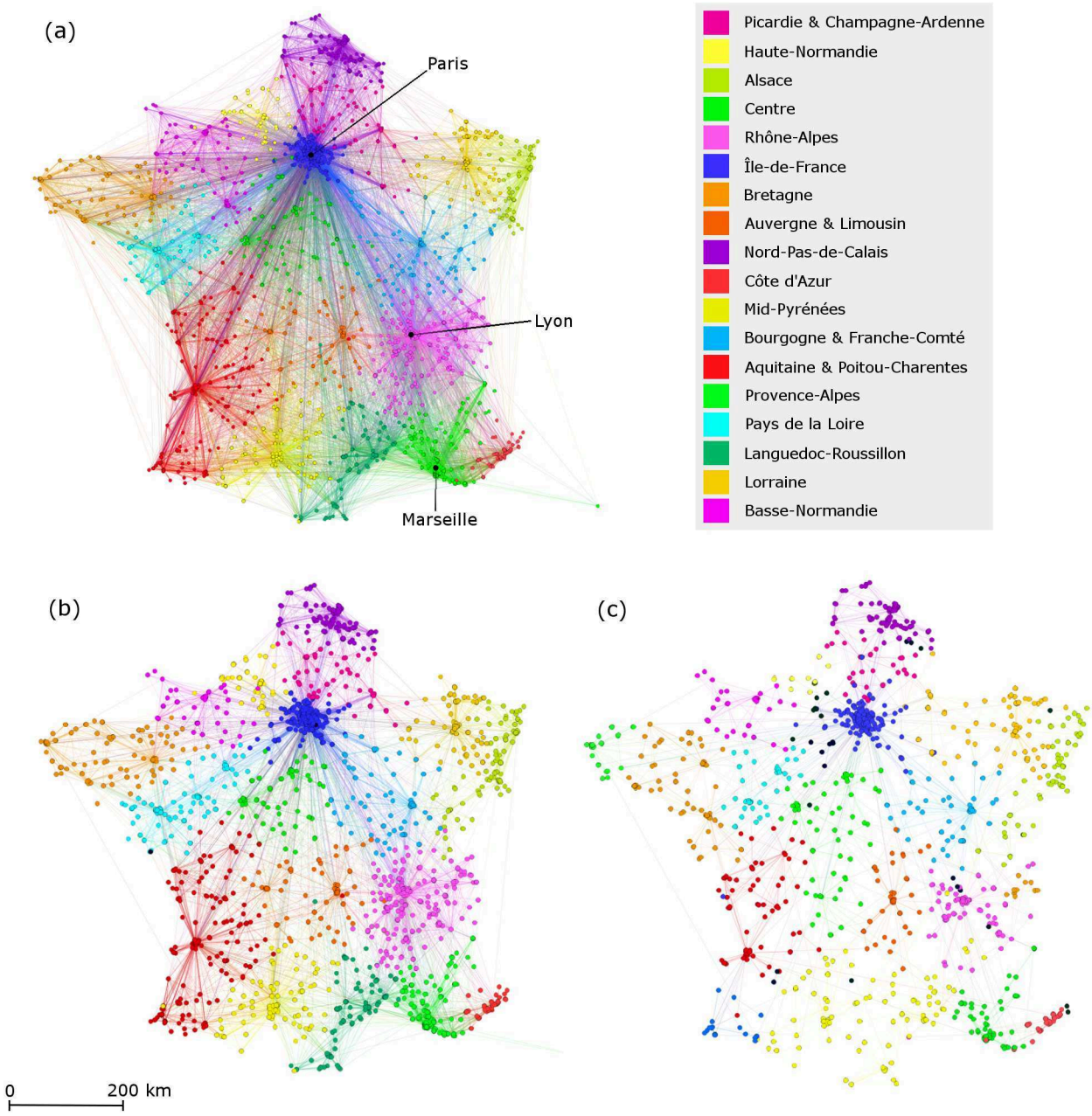


Fig 2. Regional clustering of communities detected with greedy algorithm. Network hospitals and patient trajectories of the healthcare network in France of (a) the general healthcare network, (b) the suspected-HAI healthcare network, and (c) the HAI-specific healthcare network. In the general healthcare network, 18 communities were detected by the community clustering algorithm. Four of the 18 communities identified by the algorithm combine hospitals from two regions each, such that the 22 geographical regions are mapped into 18 communities. The original 22 French metropolitan regions before they were reformed to 13 regions implemented in 2016 are shown to correspond to the 2014 data. For the HAI-specific and suspected-HAI networks, the algorithm detected a higher number of communities (36 and 21 communities respectively). The communities, which overlapped the same regional communities in the general network, were given the same color and the newly detected communities were given different colors.

https://doi.org/10.1371/journal.pcbi.1005666.g002

centers. On the other hand, the Map Equation-based communities displayed geographic community clustering at the French "departmental" or county level–the administrative division between the administrative region and the municipalities, similar to "counties" or "districts";

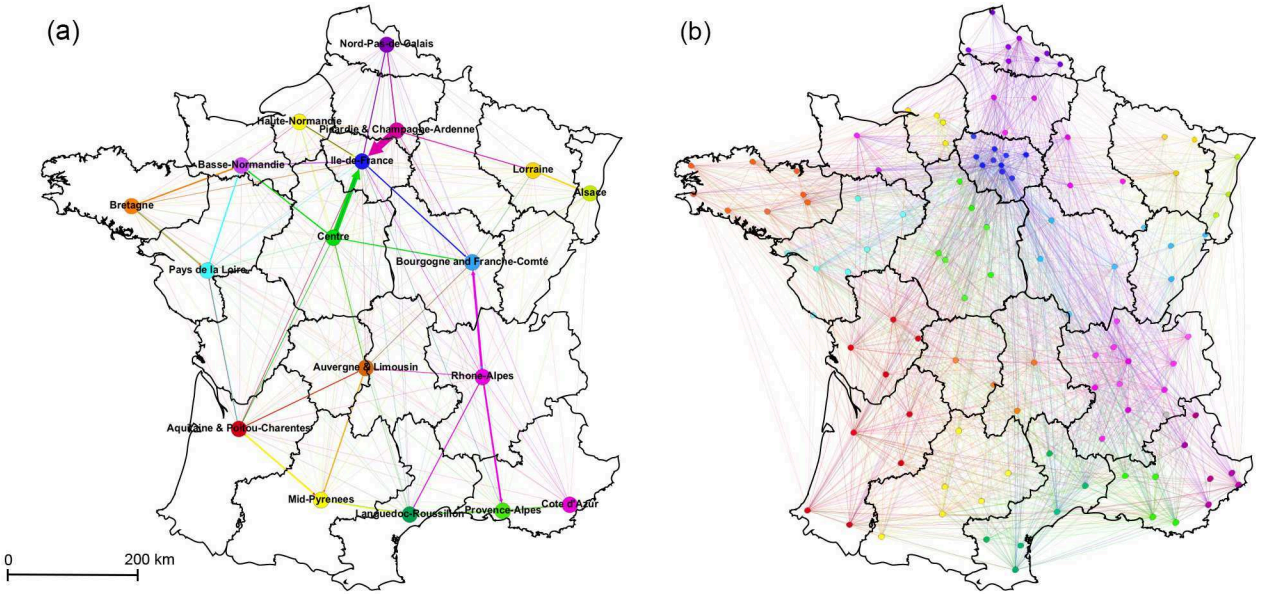


Fig 3. The intercommunity networks of patient transfers. (a) The intercommunity network from the 18 detected general patient network Greedy-based communities named based on the French metropolitan regions they encompass. Edge size and color indicate the source community and number of patients discharged. (b) The intercommunity network from 113 Map Equation communities detected in the general network. The nodes of the networks represent the geographical center of hospitals within the shared community.

https://doi.org/10.1371/journal.pcbi.1005666.g003

of which 96 departmental divisions are present in continental France. The vast majority of these departmental-level community clusters were acute-care centers followed by university hospitals centers and long-term care facilities.

What are the patterns of patient transfer between communities?

To further understand the community structure, we constructed intercommunity networks by combining patient flows between hospitals of the same community and across communities. The Greedy-based intercommunity network was composed of 18 nodes representing the sum of all patient transfers that occurred between hospitals of each community with 306 regional transfer trajectories (Fig 3A). Out of the 22 French metropolitan regions in 2014, 4 pairs of 8 metropolitan regions were combined in this intercommunity network (Picardie and Champagne-Ardenne, Auvergne and Limousin, Aquitaine and Poitou-Charentes, and Bourgogne and Franche-Comté). The network was completely connected. All regional communities were connected to one another with an average of 4590 patients moving within these intercommunity trajectories over the year. Certain trajectories played a larger role in patient movement, notably Ile-de-France which admitted the largest number of patients from neighboring regions Picardie and Champagne-Ardenne (4772 transfers) and Centre (3205 transfers) where healthcare hubs were most concentrated. The subsequent largest traffic came from the Rhone-Alpes, the second largest regional center around the city of Lyon, which discharged patients to its neighboring regions (1482 transfers to neighboring Bourgogne and Franche-Comté and 1342 transfers to neighboring Provence-Alps respectively). Nonetheless, the greatest amount of transfers (93%) occurred within the communities themselves on average with up to 98% of transfers occurring within Ile-De France for instance. Although most of these transfers occurred within the communities, the regions remained highly interconnected and certain trajectories played an important role in the interregional and nation-wide movement of patients in France.

Building the intercommunity network where community affiliation was determined by the Map Equation algorithm allowed us to consider communities based on the directionality of patient flow, which was overlooked by the Greedy algorithm. The intercommunity network was composed of 113 community nodes with 3215 trajectories with an average degree of 57 and an average of 2597 patients moving between these connections (Fig 3B). Map Equation-based intercommunity communities demonstrated more comprehensive department-level patient flow. Communities were composed of hospitals from many different departments within and across regions; however, the majority of communities were of hospitals within the same department where most of the patient exchange occurred. Concerning the most important transfer routes with the highest traffic, discharged patients coming from many neighboring departments were preferentially going to hospitals in one or a few number of departments, indicating that there was interdepartmental centralization of patient movement. For example, a community composed of 200 hospitals from 9 Ile-de-France departments sent the largest number of transfers (3137 patients) to 28 hospitals of which 24 were from one department (Val-d'Oise). In exchange, this 28-hospital community sent back 2772 patients to the larger community. Overall, patient transfers in Map-Equation communities displayed departmental clustering, but also demonstrated asymmetric movements of patients, concentrating towards small communities of hospitals usually in one department, illustrating the different nature of the communities.

Patient sharing patterns and community clustering in the networks were also analyzed based on patient age groups in which new communities were identified (S6 Text, S10–S12 Figs). Moreover, analysis of monthly temporal dynamics of the networks showed that monthly communities may be less clustered and patients may not visit all of the hospitals each month but they still retained the same regional patient sharing patterns seen in the annual network (S7 Text).

Do HAI-diagnosed patients have specific transfer flows in the healthcare network?

Having assessed the role of hospitals, hospital communities, and patient trajectories in each network, we considered if the differences in the patient-specific networks and the general networks are due to the number of patient transfers that could lead to structural differences between the specific patient population flows. We first compared the general patient network to two sets of 1000 networks built from a subset of randomly chosen patients: in the first set we selected the same number of patients as the HAI-specific network (21276 patients) at random and in the second set the same number as in the suspected-HAI network (394859 patients) at random 1000 times and reconstructed each network. Overall, both sets of random patient networks (RP) were smaller in size compared the general network in terms of the number of nodes, edges, edge weight, and as a result average degree (Table 4). In addition, most of the diameters and all average path lengths were larger in the RP networks. The diameters and path lengths of the RP networks are skewed and not normally distributed ($p < 0.001$, Shapiro-Wilk normality test). As a result, the number of patients used to reconstruct the networks did have an impact of network characteristics.

We then compared the characteristics of the HAI-specific and suspected-HAI networks to the RP networks with the same number of patients to assess if HAI patients modified network structure differently than other patients. Overall, the RP networks were larger than their HAI-specific and suspected-HAI healthcare networks analogues meaning that HAI patients were transferred to fewer hospitals than other patients (Table 4). However despite these differences, for some networks measures such as diameter, average path length, and global clustering

Table 4. Network characteristics of the random patient networks.

Network Topology Measures	General Network (GN)	Suspected-HAI Network (SHAIN)	1000 Suspected-HAI-like RP networks			HAI-Specific Network (HAISN)	1000 HAI-Specific-like RP networks		
			Mean	% ≤ GN	% ≤ SHAIN		Mean	% ≤ GN	% ≤ HAISN
Nodes	2063	1975	2032	100%	0%	1266	1583	100%	0%
Edges	50026	18812	22139	100%	0%	3722	3882	100%	0.3%
Average Edge Weight	14.02	4.92	5.43	100%	0%	2.31	1.62	100%	100%
Average Degree	48.50	19.05	21.79	100%	0%	5.88	4.91	100%	100%
Diameter	30	64	61.59	1.9%	63.0%	47	36.27	7.0%	98.8%
Average Path Length	2.99	3.63	3.78	0%	0%	5.23	8.24	0%	0%
Global Clustering Coefficient	0.23	0.16	0.19	100%	0%	0.08	0.09	100%	2.7%
Density	0.012	0.005	0.005	100%	0%	0.002	0.0016	100%	100%
Average Edge Betweenness	301	852	796	0%	100%	1557	2384	0%	0%
Average Total Closeness	1.6E-4	7.4E-5	1E-3	100%	11.5%	3.2E-5	1.7E-5	100%	100%

Comparison of the healthcare network topology measures with the average measures of 1000 simulated random patient (RP) networks that were composed of the same number of patients as the patient-specific healthcare network. The proportion of network measures equal to and less than the general network and the proportion equal to and less than the patient-specific network measures are shown in percent %. Note: "E" refers to the E-notation for the scientific notation of "×10" followed by the power.

https://doi.org/10.1371/journal.pcbi.1005666.t004

coefficient, there was less of a difference between the RP networks and the HAI networks than the RP networks and the general network. For example, 63% of suspected-HAI-like RP networks had a diameter equal to or less than that of the suspected-HAI network (64) while 1.9% of these networks had a diameter equal to or less than that of the general network (30). The average diameter (61.59) and the average path lengths (3.78) of these RP networks approached more of that of the suspected-HAI network than the general network. Thus having controlled for the number of patients and thus the size of the network, the differences observed between the suspected-HAI and the general network diameter and average path length may have been due to the suspected-HAI network being a subset of the healthcare network rather than due to differences between HAI patient transfer patterns and non-HAI patient transfers.

Discussion

In this study we show that the French healthcare networks have heterogeneous patient flows, demonstrate characteristics of small-world and scale-free networks, and are characterized with highly centralized movement of patients towards hub hospital centers. Hub hospitals are characterized as university hospitals and private hospitals in the large metropoles that dominate patient flow. The healthcare networks displayed two-level community clustering: regional community clustering reflecting the French administrative structure, and department or county-level clustering. Certain patient transfer trajectories play a more important role in transferring patients between the regional and departmental communities. Despite differences in the patient population and size, both the HAI-specific and suspected-HAI specific healthcare networks seem to rely on the same underlying structure as that of the general healthcare network.

Due to weak sensitivity and specificity of the PMSI database to detect nosocomial infections with the sole ICD-10 Y95 diagnostic, the HAI-specific network is not reliable in demonstrating

the real patient movement patterns for those infected with an HAI.[24–28] There was no confirmation if an infection was absent during admission and if an infection appeared during the first 48 hours of their stay. We suspect that the degree ranking differences and the low percent of acute-care facilities, notably private hospitals, in the HAI-specific network may be due to differences in coding practicing among hospitals rather than the epidemiology of HAIs. The suspected-HAI network reflects a standardized list of diagnoses used by the French HAI surveillance network which has been shown to be more specific and sensitive at detecting patients with HAIs.[28, 29] Having considered the network size differences in the patient-specific networks and the general network, we show that despite the differences in size of the patient population, both the HAI-specific and suspected-HAI specific healthcare networks seem to rely on the same underlying structure as that of the general healthcare network. Indeed, patient-specific networks are a subset of the general patient network and are subject to the same network dynamics.

Public university hospital centers and private hospitals in the main metropoles of France dominate patient flow. A study conducted in the Bourgogne region of France has shown that patient flow was centered towards the university hospital that admitted patients from the entire region and based on the regional proximity of the patients' residence and patients also sought care in two of the closest main healthcare hubs for specialized care (Paris or Lyon).[30] Highly connected hospitals may harbor more MRSA and MRSA bacteremia cases and may have the most potential to transmit HAIs in the entire network.[12, 13, 31, 32] HAIs may spread at a higher rate than expected at random due to the centralization of patient movement and due to the small average number of transfers required for patients to move throughout the network. A 2012 point prevalence study has shown that HAIs are most prevalent in cancers centers, university hospitals, and armed forces.[33] HAI prevalence was high in the Ile-de-France region which has many hubs, and the north-eastern regions which were not reflected by a higher number of transfers in the patient specific.[33] Albeit some difference in prevalence and patient transfer patterns, hubs should be proposed as targets for sentinel surveillance in addition to priority targets of HAI control strategies where HAI is most prevalent to achieve the most effective reduction in transmission across the country.[15]

Regional community clustering patterns as a form of network connectedness are also important in the development of strategies for coordinated HAI control.[8, 13] Our regional community clustering findings are consistent with that of the healthcare network of England in which communities tend to share more patients among clusters of hospitals in addition to patient flows centered towards a university hospital within the community.[13] Important intermediary trajectories may play a key role in the spread of HAI between hub hospitals and between communities. A study has shown that modifying the number of patients moving between communities may reduce the spread of MRSA.[34] Furthermore, we demonstrated that a two-tier hospital community exists. Depending on the clustering algorithm used, we identified clustering of healthcare communities at the regional level, consistent with the French administrative regions, and department-level communities and inter-departmental hospital clusters that took into account the directionality of patient flow. Coordinated department-level control such as screening of patients based on the identification of key department-level cluster admissions may be the first line of defense against HAI spread within the regions before spread reaches the hub university hospitals. We identified differences between department-level communities of the suspected-HAI and the general network that were overlooked at the regional community level. This may be important in distinguishing hospitals with higher potential to harbor HAI patients, with possible consequences in terms of HAI spread prediction.

Studies have proposed reducing hospital connectedness in order to reduce the risk of epidemic spread of HAI in networks.[13, 35] Decentralization of the healthcare system and more

specifically human resource and specialized health services towards the regional and department levels may help reduce the high connectedness of hubs in the metropole centers and redirect patient transfers. France has moved towards regionalization strategies with the creation of regional hospital agencies, albeit not very effective.[36, 37] In addition, the number of university hospitals may be insufficient, below that of the UK, a country with a similar population size. We recommend increasing the number facilities providing specialized services and distributing them at the local level to help redirect patient flow and potentially avoid large-scale HAI dispersal.

We considered temporal dynamics, masked in a network constructed with data for the entire year of 2014, in which observed that monthly healthcare networks were smaller and displayed less centralized patient flow; hence, infection control strategies–for short-term control–should rely more on the local department-level dynamics to minimize hospital-level outbreaks and transmission to neighboring hospitals. In the long term, regional community dynamics may give us clues regarding the gradual propagation of specific HAI pathogens over time assuming HAI carriage patterns follow that of patient flow patterns in the healthcare networks. Further studies are required to assess the temporal dynamics of HAI spread in networks to identify any potential seasonality patterns of flow and how to prevent emerging multi-drug resistant bacteria from becoming endemic.

Our study was subject to certain limitations which should be considered. Many of the university hospitals represent more than one public hospital or healthcare facility due to sharing the same identification number. For example, the largest outlier hub in Paris (AP-HP) represented 39 hospitals, 12 hospitals and 2 specialized health facilities constituted Hospices Civils de Lyon, 9 hospitals make up the university hospital of Toulouse, and 4 hospitals make up the APHM of Marseille. Consequently, university hospital centers accommodated a larger patient population than hospital centers or local hospitals, influencing the network characteristics, which may have led us to overestimate the specific patient movement patterns to and from these centers. However, the high concentration of other hospitals especially hub private hospital centers, armed forces hospitals, cancer centers, psychiatric hospitals, and private post-operative and rehabilitation centers within proximity of these public hospital hubs demonstrates that despite this, major cities such as Paris play the most important role in connecting patients in the national network and that the French healthcare network is a highly centralized system.

The healthcare networks did not include patient flow from private nursing homes that have been shown to play an important role in HAI spread.[38–42] Without private nursing homes included in our study, our results only describe the network topology of hospital patient populations which may be both younger, have shorter duration stay, and may spread HAI differently than the complete nursing home population. As a result, transmission dynamics in our networks may over or underestimate average hospital centrality measures, the volume of patient movements, and the speed at which HAI can spread.

By considering all HAIs as a whole, our networks and recommendations reflect action for a broad spectrum of HAIs; however, one should consider that specific HAIs can vary in terms of carriage and transmission patterns. In addition, recommendations based on our networks would overlook the potential exposure to community-acquired infections, although these may later spread in hospital settings, leading to healthcare-associated outbreaks. Future studies should consider all potential components of patient exposure to both community-associated and healthcare-associated infections and account for individual exposure histories to these infections.

Despite these limitations, our study provides a first description and analysis of the healthcare networks in France. The identified characteristics and community structures may greatly improve future inter-hospital HAI control strategies. The general patient network responds

best to informing regional control strategies targeting key patient trajectories and hub hospital centers. We show that the scale-free structure, the number of communities, and their distribution over the country remain qualitatively similar across all networks and that patient-specific networks rely on the underlying structure of the general patient network. Future studies should take into consideration network topology in the prediction of HAI spread and should consider the potential impact of different community definitions for multi-level infection control strategies.

Methods

Materials

The Programme de Médicalisation des Systèmes d'information (PMSI) database, a comprehensive French medico-administrative database of hospital activity and patient discharge information, is used to construct the hospital networks.[24, 25] The PMSI database has been used for epidemiological and medical research regarding HAIs.[24–28, 43] A lack of sufficient specificity and sensitivity of the PMSI to detect HAIs is highlighted in these studies. Comparison between laboratory data and hospital data shows that the PMSI has limited coverage of detecting nosocomial conditions.[25–28]

Hence, Gerbier et al. [28] use patient discharge summaries from the PMSI to detect nosocomial infections in the University Hospital of Lyon in 2006 and 2007 for the identification of HAIs in surgery, intensive care and obstetric units. They compare the PMSI data to a gold standard by systematic review of patient files for those classified under surgery, the Centre de Coordination de la Lutte contre les Infections Nosocomiales (CClin) Southwest surveillance network for ICU patients, and a combination of surveillance data from CClin and patient information data for obstetrics. The list of ICD-10 codes related to nosocomial conditions, which we entitle "suspected-HAIs," can be found in S1 Annex. Gerbier et al. find a sensitivity and specificity for case identification of nosocomial infections to be 26.3% (95% CI 13.2–42.1) and 99.5% (95% 98.8–100.0) for the identification of surgical site infections (78.9% and 65.7% by expanding the number of diagnostic codes) respectively; 48.8% (95% CI 42.6–55.0) and 78.4% (95% CI 76.1–80.1) in intensive care respectively, and 42.9% (95% CI 25.0–60.7) and 87.3% (95% CI 85.2–89.3) for identification of postpartum infections respectively.[28]

Inclusion and exclusion criteria

Using patient transfer data from 2014, three healthcare networks are reconstructed based on the following criteria:

- All patient transfers (non-specific diagnoses)
- Patient transfers with the ICD-10 code of Y95 (for nosocomial conditions or HAI) as their principal, related, or associated diagnosis in the medical, surgery, obstetric hospitals (MCO) and postoperative and rehabilitation centers (SSR)
- Patient transfers identified with all possible and suspected cases of HAIs in the surgical, intensive care, and obstetric wards in 2014, by referencing the diagnoses with known specificities and sensitivities listed in Gerbier et al. publication[28] with supplementary information from other publications.[24–26]

Only direct transfers of patients who are discharged from a hospital and sent to another in another jurisdiction ("transfer") or those who are discharged from one medical unit and move to another in the same hospital jurisdiction ("mutation") are included. The hospital discharge summaries reflected the overall hospital stay of patients and a single diagnosis made them

eligible without specification if it occurred during admission or at discharge. Patients who are discharged to their residence or deceased in a hospital are excluded. Patients hospitalized in non-continental European departments are also excluded.

Construction of patient transfer network matrices

First, the networks of hospitals and healthcare centers are re-built *in silico* using patient transfer data to model the potential movement of patients with HAIs from one hospital to another. In the PMSI database, each patient discharge summary contains information on the hospital facility of stay. Each hospital facility is identified by its unique FINESS number (Fichier National des Etablissements Sanitaires et Sociaux).[44] For this study, the FINESS and stay number of each patient discharge summary are used to merge two PMSI databases: one for acute-care and one for long-term care hospitals. Each patient stay is also numbered by order of stay across different hospitals. To create the logical sequence of patient movement, we sort each discharge summary: by patient ID and patient stay number for all observations.

The adjacency matrix [21], a graph of N nodes and E edges can be described by its' $N \times N$ adjacency matrix A defined as:

$$A_{ij} = \begin{cases} = 1 \text{ if } i \text{ and } j \text{ are connected} \\ = 0 \text{ otherwise} \end{cases}$$

In our patient transfer network, nodes (N) are defined as hospitals and edges (E) as the patient trajectories that connect hospitals. We computed origin *i* and target *j* hospitals for each patient stay by assessing if for each discharge the patient entered the hospital *i* as a transfer or mutation and left hospital *i* as a transfer or mutation. The same is computed for each *j* hospital. Using the iGraph package for R statistical software, we create the adjacency matrix of all *i* and *j* hospitals, including *i* and *j* if *i* did not transfer out any patients but received them and vice versa for *j*.[45]

We also compute the number of patients moving between hospitals *i* and *j*, as w_{ij}. The sum of the edge weights of the adjacent edges, the weight strength, is given by:

$$s_i^w = \sum_{j \in \Gamma(i)} w_{ij}$$

in which $\Gamma(i)$ is the set of neighbor hospitals of *i*.[21] Edge weights represent the number of patients within the trajectories between two healthcare facilities.

To identify the most important hospitals of a network, a series of centrality measures are calculated. The degree of a hospital, *k*, is the number of hospitals one hospital is connected to through its patient trajectories [21] defined as:

$$k_i = \sum_j A_{ij}$$

The average degree of a network[21] is given by:

$$\langle k \rangle = \frac{1}{N} \sum_i k_i = \frac{2E}{N}$$

In addition, A_{ij} is a directed graph in which the directionality of patient transfers from one hospital to another is taken into account. Consequently, we can calculate the indegree (deg^-) and outdegree (deg^+) of any given node in which the degree sum formula is given by:

$$\sum_{n \in N} deg^+(n) = \sum_{n \in N} deg^-(n) = |E|$$

Betweenness centrality measures the importance of hospital acting as an intermediary between other hospitals defined as:

$$g(i) = \sum_{s \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}$$

Where betweenness centrality $g(i)$ is equal to the sum of the σ_{st} the number of shortest paths going from s to t through hospital i measuring the importance of hospital i to the organization of flow in the network.[21] The same measure is calculated for patient trajectories defined as:

$$g(e) = \sum_{e \in E} \frac{\sigma_{st}(e)}{\sigma_{st}}$$

where edge betweenness centrality $g(e)$ is equal to the sum of the σ_{st} the number of shortest paths going from s to t through edge e measuring the importance of edge e to the organization of flow in the network.[21]

Community clustering

Two community detection algorithms are used to assess community clustering for each network, which both take into account weighted graphs.[45] A common measure of the quality of partitions of a network into communities of densely connected nodes is modularity. Modularity is a scalar value between the vales of -1 and 1 that measures the density of links inside communities compared to links between them.[21, 22] The modularity and different communities of our network are defined using a community detection algorithm. The Greedy algorithm developed by Clauset et al.[19] optimizes modularity as the algorithm relies on network formation and as a result, computes a smaller range of communities as modularity approaches 1; however, the Greedy algorithm does not take into account edge directionality and we detect communities for undirected graphs of the healthcare networks. On the other hand, the Map equation algorithm developed by Rosvall et al. detects communities based on patterns of flow and takes into account edge directionality and the directed graphs are assessed.[20] This algorithm detects communities based on network structure and how it influences the system's behavior.

Based on the community partitioning for each network, the mean geographic distance between hospitals of the same community is measured. To geo-localize hospitals, we used public government data on French hospital facilities and postal code addresses (https://www.data.gouv.fr/). Using an online batch geocoding server (http://www.findlatitudeandlongitude.com/), the hospitals' addresses were converted to latitude and longitude coordinates. A distance matrix was calculated using the haversine formula to measure great-circle distances between all hospitals.[46]

Two intercommunity matrices were developed to assess patient sharing between different communities 1) Greedy algorithm-based communities 2) Map Equation-based communities. Based on the algorithm, each hospital node is assigned a community number. A matrix summing the individual hospitals transfers for hospitals that share the same community is constructed and converted into a directed graph. In addition, the mean latitude and longitude are calculated for each community from individual geocodes of the member hospitals. For the Map Equation intercommunity network, the Greedy algorithm is applied to identify the number of communities present when modularity is maximized.

Ranking of hospitals

Hospitals were ranked by their degree, betweenness, and closeness centrality measures for each network. When the centrality measures were equal, we replaced the rankings by the mean

rankings. We tested the differences between rankings on an increasing subset of shared hospitals with the Wilcoxon rank sum test. The test was conducted as follows: starting with the highest ranked hospital in the general network, adding the next ranked general network hospital, and testing for significant differences between the general network rank and either the HAI-specific or suspected-HAI network rank of the same hospital until we compared all shared hospitals. As a result, we determined the thresholds where hospital rankings across the networks start to significantly differ which was defined as Wilcoxon rank sum test p-values under the 5% alpha risk.

Random patient networks

To compare the networks between each other, we built 1000 random patients networks from the general network. We selected the same number of patients as either the HAI (21276 patients) or suspected HAI networks (394859 patients) from the general patient network at random and reconstructed these networks using their hospital discharge summaries. We calculated various network measures and the proportion of random patient networks that had values greater than, equal to, or less than the general patient network and the respective patient-specific networks.

Supporting information

S1 Annex. All transfer patients considered as suspected to have a hospital-acquired infection.
(PDF)

S1 Text. Power-law behavior: average strength s(k) as a function of degree k.
(PDF)

S2 Text. Power-law, log-normal, and Poisson distribution goodness-of-fit tests
(PDF)

S3 Text. "Small-world" network characteristics
(PDF)

S4 Text. Comparison with Erdos-Renyi random networks.
(PDF)

S5 Text. How do the communities vary across the networks?
(PDF)

S6 Text. Does the general healthcare network change with the age of the patients?
(PDF)

S7 Text. What are the temporal dynamics of the general healthcare network?.
(PDF)

S1 Fig. Average strength and degree distribution of the general network. The degree k represents the number of hospital connections of each hospital in the general network and the average strength s(k) stands for the number of patient transfers as a function of degree. The number of patient transfers and number of hospital connections were highly positively correlated (r = 0.91). The best-fitting power law model was s(k) = k1.51 (dashed line). The curves for s(k) = k (dotted line) and s(k) = 10*k (dash-dot line) are shown for comparison.
(PDF)

S2 Fig. Average strength and degree distribution of the suspected-HAI network. Distribution of hospital connections k of each hospital in the suspected-HAI network and the average strength s(k) or number of patient transfers as a function of degree. The number of patient

transfers and number of hospital connections were highly positively correlated (r = 0.95). The best-fitting power law model was s(k) = k1.36 (dashed line). The curves for s(k) = k (dotted line) and s(k) = 10*k (dash-dot line) are shown for comparison.
(PDF)

S3 Fig. Average strength and degree distribution of the HAI-specific network. Distribution of hospital connections k of each hospital in the HAI-specific network and the average strength s(k) or number of patient transfers as a function of degree. The number of patient transfers and number of hospital connections were highly positively correlated (r = 0.99). The best-fitting power law model was s(k) = k1.26 (dashed line). The curves for s(k) = k (dotted line) and s(k) = 10*k (dash-dot line) are shown for comparison.
(PDF)

S4 Fig. Cumulative distribution functions and fit for degree and strength distribution of the general, suspected-HAI, and HAI-specific network. Cumulative distribution functions of degree k (top left) and strength s (bottom left) for the general network, suspected-HAI network (top center, bottom center), and HAI-specific network (top right, bottom right). Fitted power-law (red), log-normal (green), and Poisson (blue) distributions are shown when: x-min for degree = 77 and strength = 1191 in the general network; x-min for degree = 20 and strength = 119 in the suspected-HAI network; and x-min for degree = 7 and strength = 32 in the HAI-specific network.
(PDF)

S5 Fig. Cumulative distribution functions and fit for indegree and instrength distribution of the general, suspected-HAI, and HAI-specific network. The cumulative distribution functions of k- indegree for the general network (top left) and s- instrength (bottom left), suspected-HAI networks (top center, bottom center), and HAI-specific network (top right, bottom right). Fitted power-law (red), log-normal (green), and Poisson (blue) distributions are shown when: x-min for indegree = 36 and instrength = 698 in the general network; x-min for indegree = 13 and instrength = 131 in the suspected-HAI network; and x-min for indegree = 5 and instrength = 18 in the HAI-specific network. Power-law and log-normal had good fit for indegree and instrength in the three networks (KS-statistic p-values > 0.15) with the exception of log-normal distribution of indegree in the general and suspected-HAI network (KS-statistic p-value < 0.04). Poisson distribution was not a good fit for indegree and instrength in all networks (KS-statistic p-value < 0.0001).
(PDF)

S6 Fig. Cumulative distribution functions and fit for outdegree and outstrength distribution of the general, suspected-HAI, and HAI-specific network. The cumulative distribution functions of k+ outdegree for the general network (top left) and s+ outstrength (bottom left), suspected-HAI networks (top center, bottom center), and HAI-specific network (top right, bottom right). Fitted power-law (red), log-normal (green), and Poisson (blue) distributions are shown when: x-min for outdegree = 101 and outstrength = 1102 in the general network; x-min for outdegree = 27 and outstrength = 70 in the suspected-HAI network; and x-min for outdegree = 7 and outstrength = 3 in the HAI-specific network. Only power-law distribution had a good fit for both outdegree and outstrength (KS-statistic p-values > 0.41) while log-normal distribution was only a good fit for the HAI-specific network (KS-statistic p-value = 0.15).
(PDF)

S7 Fig. Shortest path length distributions in the networks. The length of the shortest paths or steps between any two nodes in the networks are calculated and plotted by their frequency.
(PDF)

S8 Fig. Distributions of p-values of hospital rank subsets using the Wilcoxon rank sum test in the HAI-specific network compared to the general network hospital ranks.
(PDF)

S9 Fig. Distributions of p-values of hospital rank subsets using the Wilcoxon rank sum test in the suspected-HAI network compared to the general network hospital ranks.
(PDF)

S10 Fig. Healthcare networks by age for all patients. Healthcare networks for all transferred patients (a) aged 18 and younger (b) aged 18 to 60 (c) older than 60 years old. Greedy communities were colored by the corresponding general network community regions with additional community color if not present for (a) 23 communities in which 3 were black (<5 hospitals) (b) 17 communities in which 1 was black (<5 hospitals) (c) 29 communities in which 14 were black (<5 hospitals).
(PDF)

S11 Fig. Healthcare networks by age for suspected-HAI patients. Healthcare networks for all transferred patients (a) aged 18 and younger (b) aged 18 to 60 (c) older than 60 years old. We detected a total number of Greedy-based communities for each age network (a) 22 total and 19 with over 2 hospitals from a network of 1894 hospitals and 11234 edges (b) 30 total with 17 with over 2 hospitals from a network of 1559 hospitals and 5423 edges (c) 29 total with 14 with over 2 hospitals per community from a network of 218 hospitals and 318 trajectories.
(PDF)

S12 Fig. Healthcare networks by age for HAI-specific patients. Healthcare networks for all transferred patients (a) aged 18 and younger (b) aged 18 to 60 (c) older than 60 years old. Network communities are detected using the Greedy algorithm and colored according to community membership for (a) 1143 hospitals and 2260 edges with 50 total communities with only 28 composed of more than 5 hospitals (b) 603 hospitals and 593 edges with 41 total communities and 19 with over 5 hospitals (c) 44 hospitals and 33 edges, 18 total communities, 2 communities with more than 5 hospitals, and 9 communities with more than 2 hospitals.
(PDF)

S1 Table. Network characteristics of the Erdos-Renyi random networks. Comparison of the healthcare network topology measures with the average measures of 100 simulated Erdos-Renyi (ER) networks that are parameterized with same number of nodes, edges, and Poisson-distributed average edge weight. For each measure, a t-test is conducted to compare the difference between the health network value and the average values of the ER networks with given 95% confidence intervals and p-values.
(PDF)

Acknowledgments

The authors thank Fei Gao for technical assistance.

Author Contributions

Conceptualization: Narimane Nekkab, Pascal Astagneau, Laura Temime, Pascal Crépey.

Data curation: Pascal Crépey.

Formal analysis: Narimane Nekkab, Pascal Crépey.

Funding acquisition: Pascal Astagneau.

Investigation: Narimane Nekkab.

Methodology: Narimane Nekkab, Pascal Crépey.

Project administration: Pascal Astagneau, Laura Temime, Pascal Crépey.

Resources: Narimane Nekkab, Laura Temime, Pascal Crépey.

Software: Narimane Nekkab, Pascal Crépey.

Supervision: Pascal Astagneau, Laura Temime, Pascal Crépey.

Validation: Pascal Astagneau, Laura Temime, Pascal Crépey.

Visualization: Narimane Nekkab.

Writing – original draft: Narimane Nekkab, Pascal Crépey.

Writing – review & editing: Narimane Nekkab, Laura Temime, Pascal Crépey.

References

1. O'Neill J. Antimicrobial resistance: tackling a crisis for the health and wealth of nations. London: Review on Antimicrobial Resistance 2014.
2. Poirel L, Lienhard R, Potron A, Malinverni R, Siegrist HH, Nordmann P. Plasmid-mediated carbapenem-hydrolysing beta-lactamase KPC-2 in a Klebsiella pneumoniae isolate from Switzerland. Journal of Antimicrobial Chemotherapy. 2011; 66(3):675–6. https://doi.org/10.1093/jac/dkq499 PMID: 21193473
3. Kassis-Chikhani N, Decre D, Gautier V, Burghoffer B, Saliba F, Mathieu D, et al. First outbreak of multi-drug-resistant Klebsiella pneumoniae carrying bla(VIM-1) and bla(SHV-5) in a French university hospital. Journal of Antimicrobial Chemotherapy. 2006; 57(1):142–5. https://doi.org/10.1093/jac/dki389 PMID: 16284103
4. Navon-Venezia S, Leavitt A, Schwaber MJ, Rasheed JK, Srinivasan A, Patel JB, et al. First Report on a Hyperepidemic Clone of KPC-3-Producing Klebsiella pneumoniae in Israel Genetically Related to a Strain Causing Outbreaks in the United States. Antimicrobial Agents and Chemotherapy. 2009; 53 (2):818–20. https://doi.org/10.1128/AAC.00987-08 PMID: 19029323
5. Kassis-Chikhani N, Decre D, Ichai P, Sengelin C, Geneste D, Mihaila L, et al. Outbreak of Klebsiella pneumoniae producing KPC-2 and SHV-12 in a French hospital. Journal of Antimicrobial Chemotherapy. 2010; 65(7):1539–40. https://doi.org/10.1093/jac/dkq132 PMID: 20460399
6. Poirel L, Fortineau N, Nordmann P. International Transfer of NDM-1-Producing Klebsiella pneumoniae from Iraq to France. Antimicrobial Agents and Chemotherapy. 2011; 55(4):1821–2. https://doi.org/10.1128/AAC.01761-10 PMID: 21245442
7. Arnold RS, Thom KA, Sharma S, Phillips M, Johnson JK, Morgan DJ. Emergence of Klebsiella pneumoniae Carbapenemase-Producing Bacteria. Southern Medical Journal. 2011; 104(1):40–5. https://doi.org/10.1097/SMJ.0b013e3181fd7d5a PMID: 21119555
8. Ciccolini M, Donker T, Kock R, Mielke M, Hendrix R, Jurke A, et al. Infection prevention in a connected world: The case for a regional approach. International Journal of Medical Microbiology. 2013; 303(6–7):380–7. https://doi.org/10.1016/j.ijmm.2013.02.003 PMID: 23499307
9. Smith DL, Dushoff J, Perencevich EN, Harris AD, Levin SA. Persistent colonization and the spread of antibiotic resistance in nosocomial pathogens: resistance is a regional problem. Proceedings of the National Academy of Sciences of the United States of America. 2004; 101(10):3709–14. https://doi.org/10.1073/pnas.0400456101 PMID: 14985511
10. Liljeros F, Edling C, Amaral L, Stanley H, Aberg Y. The web of human sexual contacts. Nature. 2001; 411(6840):907–8. https://doi.org/10.1038/35082140 PMID: 11418846
11. Lee BY, McGlone SM, Song Y, Avery TR, Eubank S, Chang C-C, et al. Social Network Analysis of Patient Sharing Among Hospitals in Orange County, California. American Journal of Public Health. 2011; 101(4):707–13. https://doi.org/10.2105/AJPH.2010.202754 PMID: 21330578
12. Donker T, Wallinga J, Grundmann H. Patient referral patterns and the spread of hospital-acquired infections through national health care networks. Plos Computational Biology. 2010; 6(3). https://doi.org/10.1371/journal.pcbi.1000715 PMID: 20333236
13. Donker T, Wallinga J, Slack R, Grundmann H. Hospital networks and the dispersal of hospital-acquired pathogens by patient transfer. Plos One. 2012; 7(4):8. https://doi.org/10.1371/journal.pone.0035002 PMID: 22558106

14. Fernández Gracia J, Onnela J-P, Barnett ML, Eguíluz VM, Christakis NA. Spread of pathogens in the patient transfer network of US hospitals. Physics and Society. 2015.
15. Ciccolini M, Donker T, Grundmann H, Bonten MJM, Woolhouse MEJ. Efficient surveillance for healthcare-associated infections spreading between hospitals. Proceedings of the National Academy of Sciences of the United States of America. 2014; 111(6):2271–6. https://doi.org/10.1073/pnas.1308062111 PMID: 24469791
16. Smith DL, Levin SA, Laxminarayan R. Strategic interactions in multi-institutional epidemics of antibiotic resistance. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102(8):3153–8. https://doi.org/10.1073/pnas.0409523102 PMID: 15677330
17. Opatowski L, Guillemot D, Boelle P, Temime L. Contribution of mathematical modeling to the fight against bacterial antibiotic resistance. Current Opinion in Infectious Diseases. 2011; 24(3):279–87. https://doi.org/10.1097/QCO.0b013e3283462362 PMID: 21467930
18. APHP. Nous connaître: Assistance Publique—Hôpitaux de Paris; [updated April 14 2016; cited 2016 July 18]. Available from: http://www.aphp.fr/nous-connaitre.
19. Clauset A. Finding local community structure in networks. Physical Review E. 2005; 72(2). https://doi.org/10.1103/PhysRevE.72.026132 PMID: 16196669
20. Rosvall M, Bergstrom CT. Maps of random walks on complex networks reveal community structure. Proceedings of the National Academy of Sciences of the United States of America. 2008; 105(4):1118–23. 10.1073/pnas.0706851105. https://doi.org/10.1073/pnas.0706851105 PMID: 18216267
21. Barthélemy M. Spatial networks. Physics Reports. 2011; 499(1–3):1–101. http://dx.doi.org/10.1016/j.physrep.2010.11.002.
22. Newman MEJ, Girvan M. Finding and evaluating community structure in networks. Physical Review E. 2004; 69(2):026113. https://doi.org/10.1103/PhysRevE.69.026113 PMID: 14995526
23. Chen M, Kuzmin K, Szymanski BK. Community Detection via Maximization of Modularity and Its Variants. IEEE Transactions on Computational Social Systems. 2014; 1(1):46–65.
24. Lamarsalle L, Hunt B, Schauf M, Szwarcensztein K, Valentine WJ. Evaluating the clinical and economic burden of healthcare-associated infections during hospitalization for surgery in France. Epidemiology and Infection. 2013; 141(12):2473–82. https://doi.org/10.1017/S0950268813000253 PMID: 23445665
25. Fourquet F, Demont F, Lecuyer AI, Rogers MA, Bloc DH. French medical hospital information system and cross infection surveillance: theory and feasibility. Médecine Mal Infect. 2003; 33(2):110–3. https://doi.org/10.1016/s0399-077x(02)00005-7
26. Bouzbid S, Gicquel Q, Gerbier S, Chomarat M, Pradat E, Fabry J, et al. Automated detection of nosocomial infections: evaluation of different strategies in an intensive care unit 2000–2006. Journal of Hospital Infection. 2011; 79(1):38–43. https://doi.org/10.1016/j.jhin.2011.05.006 PMID: 21742413
27. Legras B, Feldmann L, Burdin J, Weber M, Hartemann P. Evaluation des infections nosocomiales à partir des données du laboratoire et des résumés d'hospitalisation. Médecine Mal Infect. 1993; 23(7):307–15. https://doi.org/10.1016/S0399-077X(05)80551-7
28. Gerbier S, Bouzbid S, Pradat E, Baulieux J, Lepape A, Berland M, et al. Use of the French medico-administrative database (PMSI) to detect nosocomial infections in the University Hospital of Lyon. Revue d'Epidémiologie et de Santé Publique. 2011; 59(1):3–14. https://doi.org/10.1016/j.respe.2010.08.003 PMID: 21237594
29. Apport du système d'information médicalisé dans la surveillance en hygiène hospitalière. CClin Sud-Ouest, 2009.
30. Nuemi G, Afonso F, Roussot A, Billard L, Cottenet J, Combier E, et al. Classification of hospital pathways in the management of cancer: application to lung cancer in the region of burgundy. Cancer epidemiology. 2013; 37(5):688–96. Epub 07/16. https://doi.org/10.1016/j.canep.2013.06.007 10.1016/j.canep.2013.06.007. Epub 2013 Jul 10. PMID: 23850083
31. Chase-Topping CLG, Bram ADvB, Oliver B, Chris R, Thibaud P, Laura I, et al. Not just a matter of size: a hospital-level risk factor analysis of MRSA bacteraemia in Scotland. BMC Infectious Diseases. 2016; 16(1):222. https://doi.org/10.1186/s12879-016-1563-6 PMID: 27209082
32. Lee BY, McGlone SM, Wong KF, Yilmaz SL, Avery TR, Song Y, et al. Modeling the spread of methicillin-resistant *Staphylococcus aureus* (MRSA) outbreaks throughout the hospitals in Orange County, California. Infection Control and Hospital Epidemiology. 2011; 32(6):562–72. https://doi.org/10.1086/660014 PMID: 21558768
33. Réseau d'alerte d'investigation et de surveillance des infections nosocomiales (Raisin). Enquête nationale de prévalence des infections nosocomiales et des traitements anti-infectieux en établissements de santé, France, mai-juin 2012. Résultats. Saint-Maurice: Institut de veille sanitaire, 2013.

34. Donker T, Wallinga J, Grundmann H. Dispersal of antibiotic-resistant high-risk clones by hospital networks: changing the patient direction can make all the difference. Journal of Hospital Infection. 2014; 86 (1):34–41. https://doi.org/10.1016/j.jhin.2013.06.021 PMID: 24075292

35. Newman MEJ. The spread of epidemic disease on networks. 2002. 10.1103/PhysRevE.66.016128.

36. Decentralization in health care. Strategies and outcomes. 2007.

37. Polton D. Décentralisation des systèmes de santé: Quelques réflexions à partir d'expériences étrangères. Questions d'économie de la santé. 2003.

38. Lee BY, Song Y, Bartsch SM, Kim DS, Singh A, Avery TR, et al. Long-term care facilities: important participants of the acute care facility social network? PLoS One. 2012; 6(12):e29342. Epub 01/05. https://doi.org/10.1371/journal.pone.0029342 10.1371/journal.pone.0029342. Epub 2011 Dec 27. PubMed PMID: PMID: 22216255.

39. Lee BY, Bartsch SM, Wong KF, Singh A, Avery TR, Kim DS, et al. The importance of nursing homes in the spread of methicillin-resistant *Staphylococcus aureus* (MRSA) among hospitals. Medical Care. 2013; 51(3):205–15. https://doi.org/10.1097/MLR.0b013e3182836dc2 PMID: 23358388

40. Lee BY, Singh A, Bartsch SM, Wong KF, Kim DS, Avery TR, et al. The potential regional impact of contact precaution use in nursing homes to control methicillin-resistant *Staphylococcus aureus*. Infection Control and Hospital Epidemiology. 2013; 34(2):151–60. https://doi.org/10.1086/669091 PMID: 23295561

41. Barnes SL, Harris AD, Golden BL, Wasil EA, Furuno JP. Contribution of interfacility patient movement to overall methicillin-resistant *Staphylococcus aureus* prevalence levels. Infection Control and Hospital Epidemiology. 2011; 32(11):1073–8. https://doi.org/10.1086/662375 PMID: 22011533

42. Lesosky M, McGeer A, Simor A, Green K, Low DE, Raboud J. Effect of patterns of transferring patients among healthcare institutions on rates of nosocomial methicillin-resistant *Staphylococcus aureus* transmission: a Monte Carlo simulation. Infection Control and Hospital Epidemiology. 2011; 32(2):136–47. https://doi.org/10.1086/657945 PMID: 21460468

43. Nuemi G, Astruc K, Aho S, Quantin C. Comparing results of methicillin-resistant Staphylococcus aureus (MRSA) surveillance using the French DRG-based information system (PMSI). Rev Epidemiol Sante Publique. 2013; 61(5):455–61. https://doi.org/10.1016/j.respe.2013.04.008 PMID: 23993689

44. Hogea C, Van Effelterre T, Cassidy A. A model-based analysis: what potential could there be for a S. aureus vaccine in a hospital setting on top of other preventative measures? Bmc Infectious Diseases. 2014; 14:13. https://doi.org/10.1186/1471-2334-14-13 PMID: 24405683

45. Csardi G, Nepusz T. The igraph software package for complex network research. InterJournal, Complex Systems 1695. 2006.

46. Sinnott R. Virtues of the Haversine. Sky and Telescope. 1984; 68(2):158.

S1 Annex. All Transfer Patients Considered as Suspected to have a Hospital-Acquired Infection

Suspected-HAI patients are identified by the presence of at least one of the following International Classification of Diseases, ICD-10 codes at the principal, related, or associated diagnosis surveyed by the PMSI database [24, 28]:

Nosocomial condition: Y95

Surgical site infection: T814, T815, T816, T826, T827, T835, T836, T845, T846, T847, T857, O860

Extensive infection: T813, T818, T888, T889, K316, K603, K604, K605, K632, K823, K833, N360, N823, Z090, Z094, Z097, Z098, Z099, Z480, Z488, Z489, R50, R500, R501, R09, A40, A400, A401, A402, A403, A408, A409, A41, A410, A411, A412, A413, A414, A415, A418, A419, A427, T874

Pneumonia: J10-, J11-, J12-, J13-, J14-, J15-, J16-, J17-, J18-

Urinary infection: N300, N34-, N390, O862, O863, T835

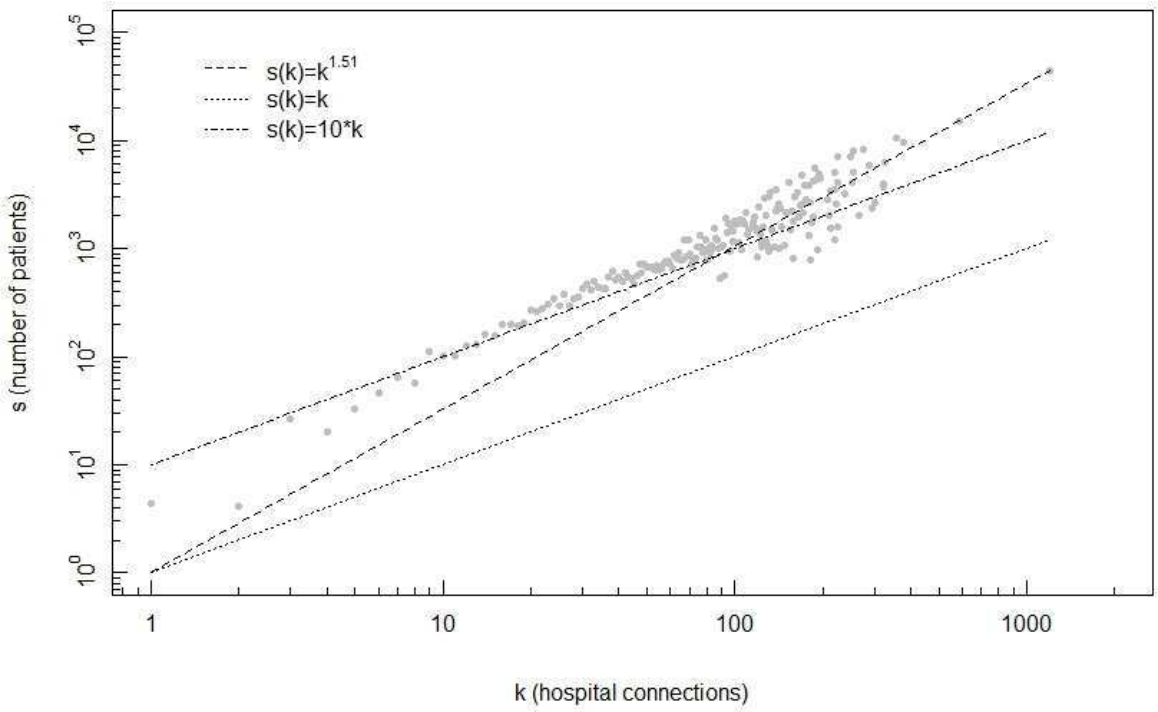
Bacteremia: A021, A207, A217, A227, A241, A267, A280, A327, A392, A393, A394, A40-, A41-, A427, A483, A499, A548, B007, B377, O080, O753, O85, P3600, P3610, P3620, P3630, P3640, P3650, P3680, P3690

Endometritis: N710, N719, N72, O235, O85

Breast infection: O91-

Uncategorized infections: O861, O864, O868

S1 Fig. Average Strength and Degree Distribution of the General Network



S1 Fig. The degree k represents the number of hospital connections of each hospital in the general network and the average strength s(k) stands for the number of patient transfers as a function of degree. The number of patient transfers and number of hospital connections were highly positively correlated (r = 0.91). The best-fitting power law model was $s(k)=k^{1.51}$ (dashed line). The curves for s(k)=k (dotted line) and s(k)=10*k (dash-dot line) are shown for comparison.

S1 Table. Network Characteristics of the Erdos-Renyi Random Networks

	HAI-Specific Network	100 HAI-Specific-like ER networks			Suspected-HAI Network	100 Suspected-HAI-like ER networks			General Network	100 General Network-like networks		
Network Topology Measures	Value	Mean	95% CI	P-value	Value	Mean	95% CI	P-value	Value	Mean	95% CI	P-value
Average Degree	5.88	5.88	5.88-5.88	1	19.05	19.05	19.05-19.05	1	48.50	48.50	48.50-48.50	1
Diameter	47	37.31	36.64-37.98	<0.001	64	27.83	27.43-28.23	<0.001	30	49.87	49.69-50.05	<0.001
Average Path Length	5.23	6.55	6.54-6.56	<0.001	3.63	3.618	3.618-3.619	<0.001	2.99	2.735	2.735-2.735	<0.001
Global Clustering Coefficient	0.08	0.0047	0.0045-0.0048	<0.001	0.16	0.0096	0.0095-0.0097	<0.001	0.23	0.02337	0.02334-0.0234	<0.001
Density	0.002	0.002	0.002-0.002	1	0.005	0.0048	1975-1975	1	0.012	0.012	0.012-0.012	1
Average Edge Betweenness	1556.94	3143.72	3128.89-3158.55	<0.001	852.23	989.39	988.07-990.70	<0.001	301.27	252.81	252.75-252.88	<0.001
Average Total Closeness	3.2e-5	9.3e-5	8.9e-5-9.7e-5	<0.001	7.4e-5	5.8e-5	5.79e-5-5.81e-5	<0.001	1.6e-4	1.89e-5	1.89e-5-1.90e-5	<0.001

S1 Table. Comparison of the healthcare network topology measures with the average measures of 100 simulated Erdos-Renyi (ER) networks that are parameterized with same number of nodes, edges, and Poisson-distributed average edge weight. For each measure, a t-test is conducted to compare the difference between the health network value and the average values of the ER networks with given 95% confidence intervals and p-values.

S1 Text. Power-Law Behavior: Average Strength s(k) as a Function of Degree k

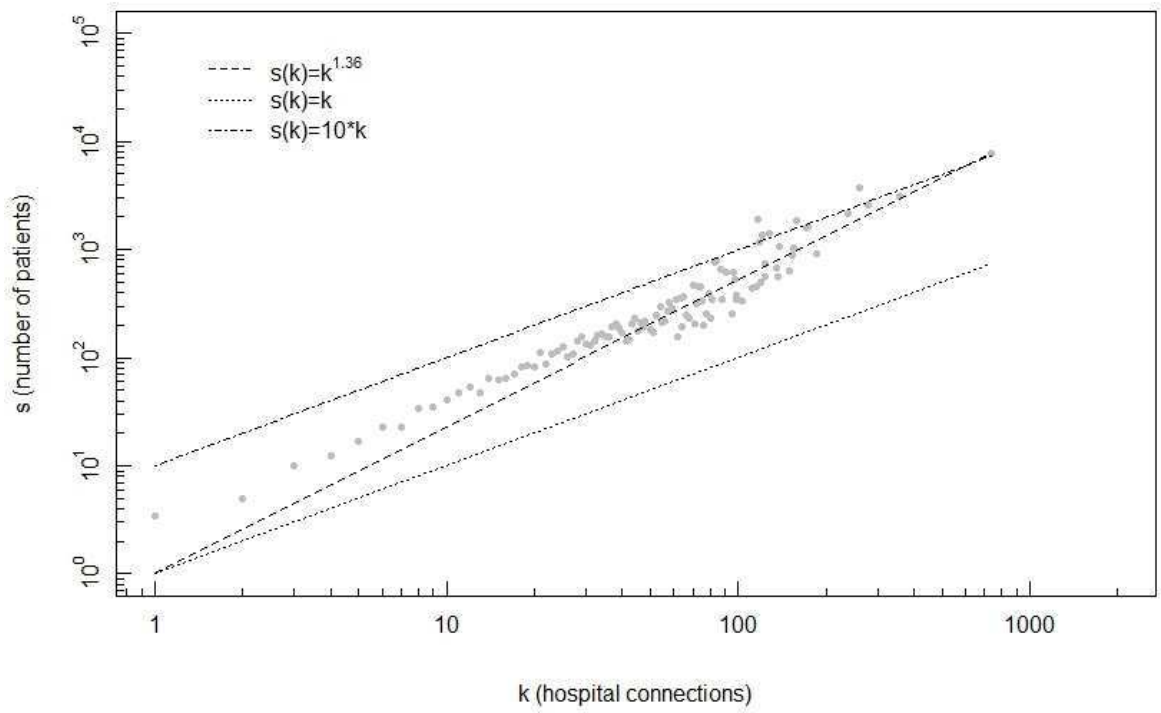
To better understand the heavy tailed behavior in the networks, we plotted average patient transfers and hospital connectedness or degree using the following formula given by Barrat et al.[1]:

$$s(k) \sim k^{\beta}$$

The general healthcare network's average strength given as a function of degree *k, s(k),* varied with a power β of 1.51 (S1 Fig). For the suspected-HAI networks and the HAI-specific network average strength varies by a power of 1.36 and 1.26 as a function of degree respectively (S2 Fig, S3 Fig). Therefore, in these healthcare networks, the number of patients transferred by a hospital increased at a higher rate than that of the hospital's connections and was most high in the general network.

[1] Barrat A, Barthelemy M, Pastor-Satorras R, Vespignani A. The architecture of complex weighted networks. Proc Natl Acad Sci U S A. 2004;101(11):3747-52. doi: 10.1073/pnas.0400087101. PubMed PMID: 15007165; PubMed Central PMCID: PMCPMC374315.

S2 Fig. Average Strength and Degree Distribution of the Suspected-HAI Network



S2 Fig. Distribution of hospital connections *k* of each hospital in the suspected-HAI network and the average strength *s(k)* or number of patient transfers as a function of degree. The number of patient transfers and number of hospital connections were highly positively correlated (r = 0.95). The best-fitting power law model was $s(k)=k^{1.36}$ (dashed line). The curves for s(k)=k (dotted line) and s(k)=10*k (dash-dot line) are shown for comparison.

S2 Text. Power-Law, Log-Normal, and Poisson Distribution Goodness-Of-Fit Tests

Power law distribution of degree is a key characteristic of "scale-free" networks, in which a small number of nodes are highly connected.[1,2] We described the distribution of the degree and strength of our directed healthcare networks and tested the fit of the (1) power-law, (2) log-normal, and (3) Poisson distributions to the data (S4 Fig). Using a goodness of fit test via a bootstrapping procedure, we found that the power law distributions had good fit for degree in all three healthcare networks (Kolmogorov-Smirnoff (KS) statistic p-values > 0.07) evidencing that they indeed displayed scale-free characteristics. The power law distributions for strength had good fit in all three networks as well (KS statistic p-values > 0.2). Log-normal distribution was also a good fit for degree in the three networks (KS statistic p-values > 0.27), as well as for strength in the general and HAI-specific networks, but not in the suspected-HAI network (KS statistic p-values = 0.02). Finally, Poisson distribution was not a good fit for either degree or strength in all three networks, demonstrating that the healthcare networks were heterogeneously distributed (KS statistic p-values < 0.001 for all).

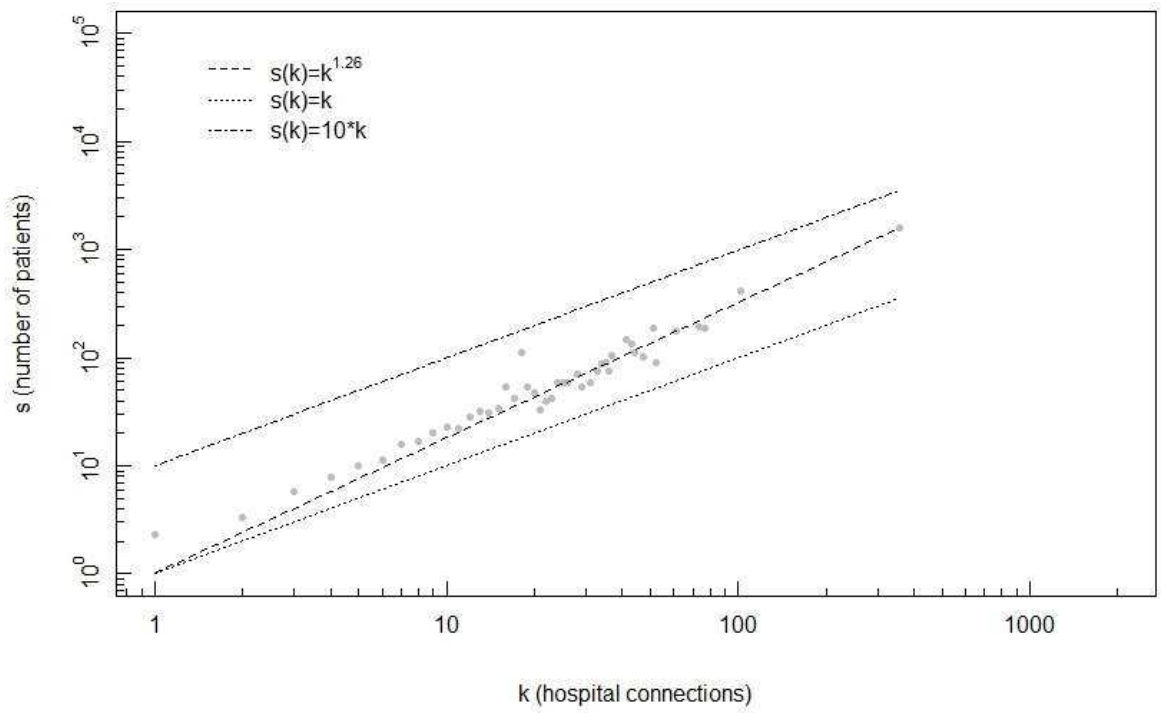
Goodness of fit tests were also performed for distributions for k- indegree, k+ outdegree, s- instrength and s+ outstrength, showing similar results (S5 Fig, S6 Fig). In addition, the average strength as a function of degree also exhibited a power-law behavior, with higher power in the general healthcare network, followed by the suspected-HAI network, and HAI-specific network (S1 Text, S1 Fig, S2 Fig, and S3 Fig). Therefore, all three healthcare networks displayed scale-free properties with a limited number of highly connected "hub" hospitals.

---

[1] Barrat A, Barthelemy M, Pastor-Satorras R, Vespignani A. The architecture of complex weighted networks. Proc Natl Acad Sci U S A. 2004;101(11):3747-52. doi: 10.1073/pnas.0400087101. PubMed PMID: 15007165; PubMed Central PMCID: PMCPMC374315.

[2] Barabási A-L, Albert R. Emergence of Scaling in Random Networks. 1999. doi: 10.1126/science.286.5439.509.

S3 Fig. Average Strength and Degree Distribution of the HAI-Specific Network



S3 Fig. Distribution of hospital connections *k* of each hospital in the HAI-specifc network and the average strength *s(k)* or number of patient transfers as a function of degree. The number of patient transfers and number of hospital connections were highly positively correlated (r = 0.99). The best-fitting power law model was $s(k)=k^{1.26}$ (dashed line). The curves for s(k)=k (dotted line) and s(k)=10*k (dash-dot line) are shown for comparison.

S3 Text. "Small-world" network characteristics

In addition to assessing the scale-free properties of the networks, we assessed the "small-world" characteristics to determine if nodes were closely within reach of all others in the network. Small average path length and large clustering coefficients are indicators of a small-world effect.[1] All three healthcare networks had smaller diameters and smaller average path length given their size compared to Erdos-Renyi random networks (S4 Text, S1 Table), indicating that hospital subpopulations were within close topological proximity to one another and that patients, once admitted to any hospitals, could be more easily sent to all hospitals in the network within a few number of transfers (Table 1). The largest network in size, the general patient network, had a diameter of 30, defined as the longest of the shortest distance between any two nodes in the network, and an average path length of 2.99, given by the average shortest path between all possible pairs of connected nodes in the network. In S7 Fig, the distribution of the shortest path lengths across the networks is shown. The general network has a higher frequency of path lengths between zero and five whereas in the HAI-specific networks, the frequency is reduced and the longer path lengths become more frequent. Therefore, hospitals within the general network were more efficient in transfer patients.

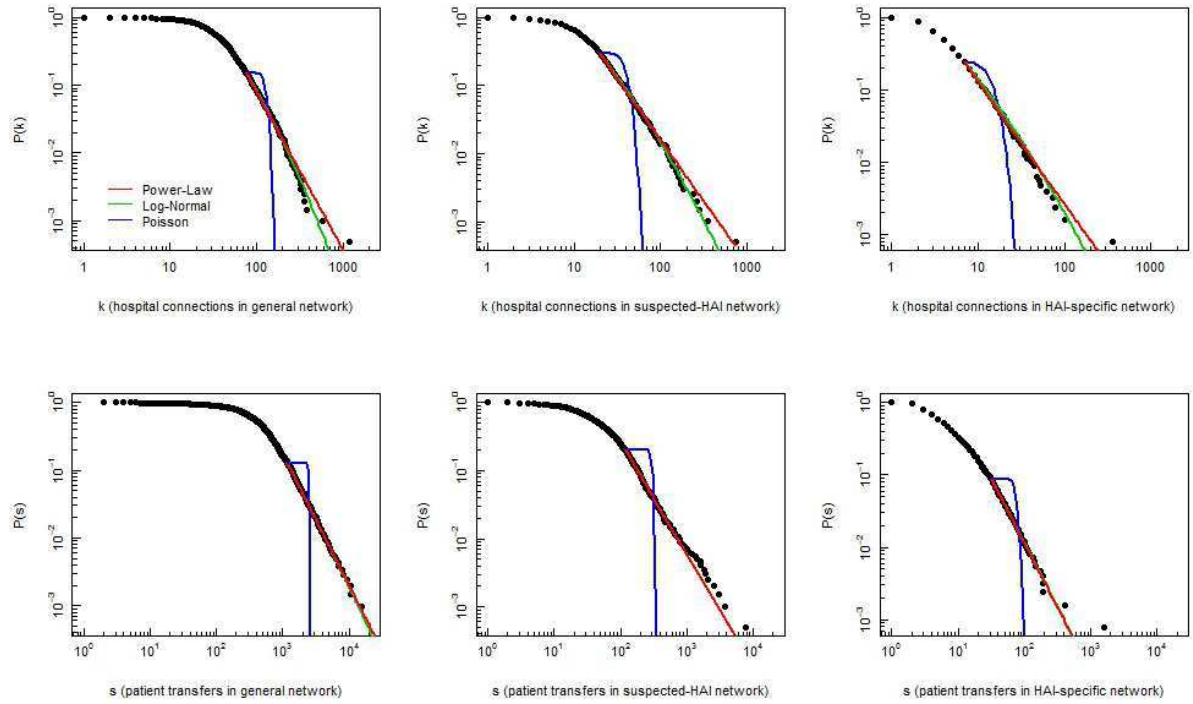
Further analysis of the networks also supported this observed small-world characteristic. Graph density, as observed in the three networks, is the total proportion of existing edges out of the potential edges that can exist to connect all nodes together. Computation of densities indicated that only 0.2%, 0.5%, and 1.2% out of all possible connections exist in the HAI-specific, suspected-HAI, and general patient networks respectively (Table 1). Hospitals shared patients with a limited number of other hospitals in the network. Moreover, the global clustering coefficient (GCC), which gives an overall indication of the clustering or number of triangles

[1] Watts DJ, Strogatz S, H. Collective dynamics of 'small-world' networks. Nature. 1998;393(6684):440-2. doi:10.1038/30918.

(triplets of nodes) existing among the possible connected ones,[2] was high in the three networks compared to a random network of the same size (S4 Text, S2 Table), especially in the general healthcare network. Therefore, hospitals sending patients to the same hospitals were more likely to be linked together by patient sharing.

[2] van der Hofstad R. Random Graphs and Complex Networks. Eindhoven: Cambridge University Press; 2016.

S4 Fig. Cumulative Distribution Functions and Fit for Degree and Strength Distribution of the General, Suspected-HAI, and HAI-Specific Network



S4 Fig. Cumulative distribution functions of degree *k* (top left) and strength *s* (bottom left) for the general network, suspected-HAI network (top center, bottom center), and HAI-specific network (top right, bottom right). Fitted power-law (red), log-normal (green), and Poisson (blue) distributions are shown when: x-min for degree = 77 and strength = 1191 in the general network; x-min for degree = 20 and strength = 119 in the suspected-HAI network; and x-min for degree = 7 and strength = 32 in the HAI-specific network.

S4 Text. Comparison with Erdos-Renyi Random Networks

Given that the healthcare networks differ in size, they cannot be compared directly for all network properties. To further assess the distinct properties of the healthcare networks, we compared each of the three healthcare networks to 100 simulated Erdos-Renyi (ER) random directed networks parameterized with the same number of nodes, edges, and average edge weight. We constructed 100 Erdos and Renyi random networks[1] with the same number of nodes, edges, and a Poisson distributed edge weight corresponding to total transfers of the three networks. In the Erdos-Renyi model, the fixed number of nodes have the same probability of being connected by a fixed number of edges.[8]

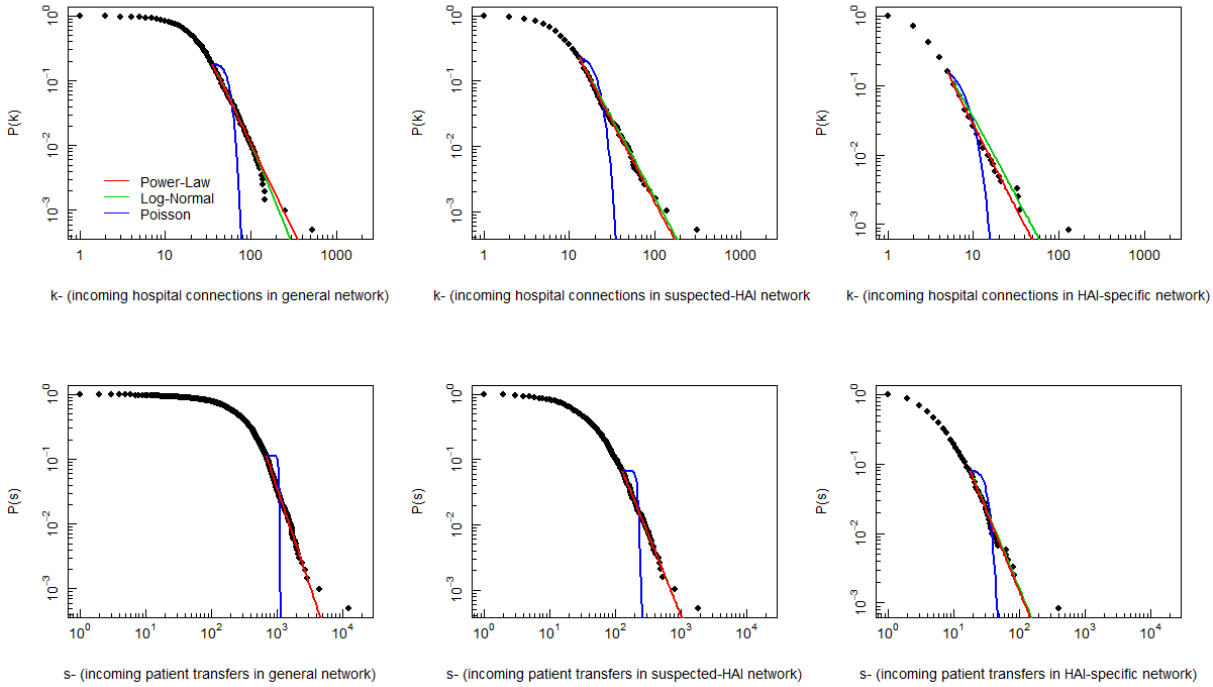
As expected, the random networks displayed less small-world characteristics compared to the healthcare networks. Overall, healthcare networks were more clustered than ER networks (S1 Table). Diameter was smaller in the general healthcare network and larger for the HAI-specific and suspected-HAI networks compared to the average ER network. Average path length was larger in the general healthcare network and suspected-HAI network and smaller in the HAI-specific network indicating that in the HAI-specific network have a closer average distance between any two nodes than that of an ER network. The larger path length in the two largest healthcare networks compared to ER networks may be due to the more heterogeneous distribution of path lengths in the healthcare networks where the distribution of path lengths may vary between highly connected and highly disconnected hospitals. The average total closeness was much smaller in the suspected-HAI network and general network random networks, indicating that hospitals in these networks will be able to disperse their patients in the network quicker, while in the HAI-specific network, patient movement was slower than in an average ER network (S1 Table).

---

[1] Erdos P, Renyi A. On random graphs. Publicationes Mathematicae. 1959;6:290-7.

Heterogeneity in patient trajectories and community clustering also distinguish the healthcare networks and ER networks. Another possible contribution to efficiency of patient movement in the general healthcare network may be due to higher average edge betweenness (301) and a higher maximum edge betweenness (1175) compared to ER networks (252.81 95% CI [252.75-252.88] and 32.02 95% CI [31.74-32.30] in the ER networks respectively). While compared to the average ER network, the average edge betweenness in the HAI-specific and suspected-HAI healthcare networks were lower, maximum edge betweenness of the HAI-specific (73) and suspected-HAI healthcare networks (314) were higher, demonstrating that flow was concentrated in a small number of edges and less evenly distributed in these healthcare networks (S1 Table). ER networks display a smaller number of communities. Hospitals within the same community were geographically further away from one another on average than in the healthcare networks (with the exception of Map Equation-detected communities in the suspected-HAI network). In résumé, the healthcare networks had more centralized and efficient patient flow concentrated in a small number of nodes and edges while the general patient network was more clustered and efficient compared to both the smaller healthcare networks and ER networks.

S5 Fig. Cumulative Distribution Functions and Fit for Indegree and Instrength Distribution of the General, Suspected-HAI, and HAI-Specific Network

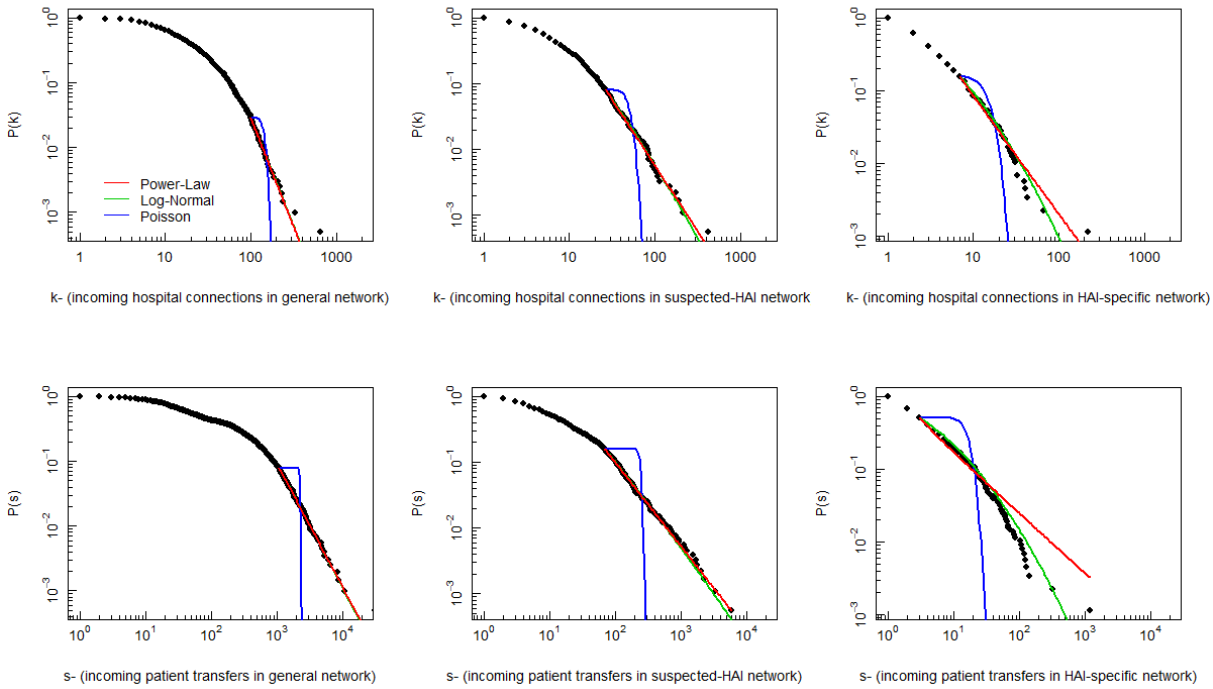


S5 Fig. The cumulative distribution functions of *k*- indegree for the general network (top left) and *s*- instrength (bottom left), suspected-HAI networks (top center, bottom center), and HAI-specific network (top right, bottom right). Fitted power-law (red), log-normal (green), and Poisson (blue) distributions are shown when: x-min for indegree = 36 and instrength = 698 in the general network; x-min for indegree = 13 and instrength = 131 in the suspected-HAI network; and x-min for indegree = 5 and instrength = 18 in the HAI-specific network. Power-law and log-normal had good fit for indegree and instrength in the three networks (KS-statistic p-values > 0.15) with the exception of log-normal distribution of indegree in the general and suspected-HAI network (KS-statistic p-value < 0.04). Poisson distribution was not a good fit for indegree and instrength in all networks (KS-statistic p-value < 0.0001).

S5 Text. How do the communities vary across the networks?

Healthcare networks displayed differences in community clustering of their hospitals which could be important to better understanding activity in the national healthcare system. Visually, community clustering was similar across the three networks (Fig 2). However, on average, each of the 18 Greedy-based community nodes in the general network (Fig 2a) had hospitals belonging to 4.11 different HAI-specific Greedy-based communities (Fig 2c). These 4.11 communities were not evenly distributed and most hospitals in a general network community belonged to only one dominating HAI-specific community. The dominating HAI-specific communities made up for 78% of the general network community hospital composition for shared hospitals between the two networks. Similarly, on average one general network community was made up of hospitals from 2.56 different suspected-HAI communities (Fig 2b). In addition, one suspected-HAI community made up most (92% of the shared hospitals on average) of the general network community composition. As a result, HAI-specific and suspected-HAI healthcare network hospitals shared most of the community composition found in the general network. When taking into consideration only communities containing at least 2 hospitals, the suspected-HAI network had an equal number of communities as the general network for both algorithms (18 Greedy-based communities and 112 Map Equation-based communities). Therefore, the suspected-HAI healthcare network had very similar community patient sharing structure compared to the general network for both algorithms while the HAI-specific network less so, demonstrating that hospitals transfer HAI-specific patients differently than other patients.

S6 Fig. Cumulative Distribution Functions and Fit for Outdegree and Outstrength Distribution of the General, Suspected-HAI, and HAI-Specific Network



S6 Fig. The cumulative distribution functions of $k+$ outdegree for the general network (top left) and $s+$ outstrength (bottom left), suspected-HAI networks (top center, bottom center), and HAI-specific network (top right, bottom right). Fitted power-law (red), log-normal (green), and Poisson (blue) distributions are shown when: x-min for outdegree = 101 and outstrength = 1102 in the general network; x-min for outdegree = 27 and outstrength = 70 in the suspected-HAI network; and x-min for outdegree = 7 and outstrength = 3 in the HAI-specific network. Only power-law distribution had a good fit for both outdegree and outstrength (KS-statistic p-values > 0.41) while log-normal distribution was only a good fit for the HAI-specific network (KS-statistic p-value = 0.15).

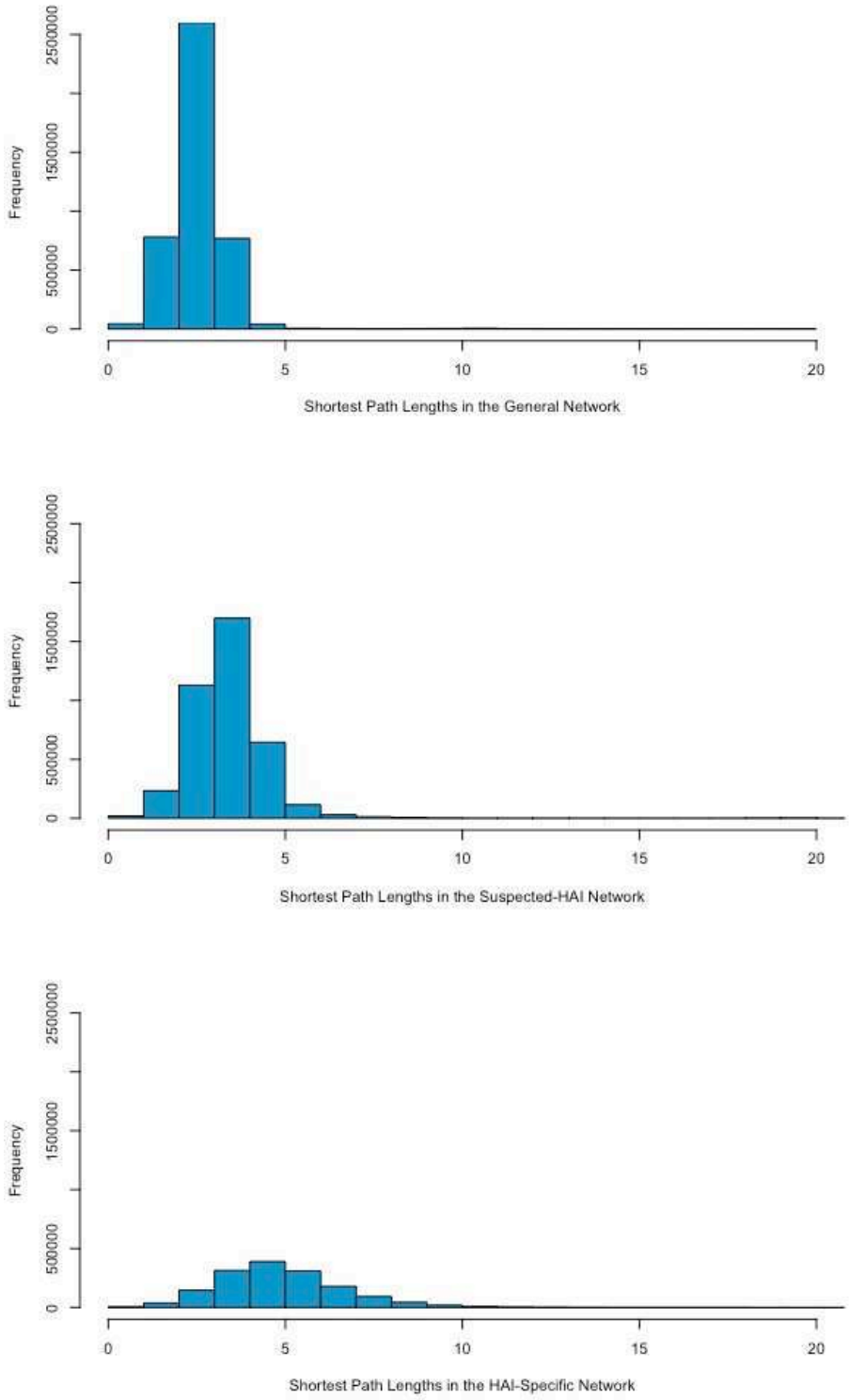
S6 Text. Does the General Healthcare Network Change with the Age of the Patients?

To better categorize the patient profiles in the healthcare networks, we assessed the age distributions of patients since different age populations may display varied pathologies, vulnerabilities, and risks for HAI transmission. Patients across the networks were characterized by older age with an average age for all transferred patients of 61.18 (SD=22.09), an average age of 55.25 (SD=21.65) for patients suspected to have an HAI, and an average age of 64.94 (SD=18.51) for patients diagnosed with an HAI. Women tended to be slightly older than men in the general and HAI-specific networks ($p < 0.05$, t-test). The oldest patient population was in the postoperative and rehabilitation centers for all transferred patients and those suspected to have an HAI (average of 68.60 (SD=18.35) and 67.45 (SD=17.40) respectively).

To assess potential changes in community clustering due to age groups, we reconstructed 3 age-specific networks (1) less than or equal to 18 years old (2) between 18 and 60 years old (3) over 60 years old. Not surprisingly, we found that the largest network was of transfer patients over the age of 60. Results are shown for all transferred patients in S7 Fig and for HAI and suspected-HAI patients in S8 Fig and S9 Fig respectively. As expected, patients over the age of 60 formed a similar network to that of all patient transfers with 1996 nodes, 31427 trajectories, and 23 Greedy communities of which only 21 had more than 5 hospitals (S7a Fig). The middle-aged network was second largest with 1985 nodes, 18304 trajectories, and 17 communities of which 16 had more than 5 hospitals (S7b Fig). On the other hand, children and adolescents consistently made up smaller networks and for all transferred patients of this age group, they created a network of 575 nodes, 1349 trajectories, and 29 communities of which 15 had more than 5 hospitals (S7c Fig). The Greedy-based communities displayed regional geographic clustering of hospitals, differed between age groups, and differed from the general healthcare network. Most importantly, we saw an introduction of three new regional communities in the oldest age group not previously identified (S7a Fig).

To better understand age distribution, which may also play a role in patient movement patterns, we compare the networks obtained for different patient age groups. Although the network is small, we identified that adolescent patient movement is different than that of older patients in terms of relative community size for the largest communities. We identify three new communities unique to patients over the age of 60 that displays some difference in community composition than the general network, which may interest decision makers targeting elderly populations.

S7 Fig. Shortest path length distributions in the networks



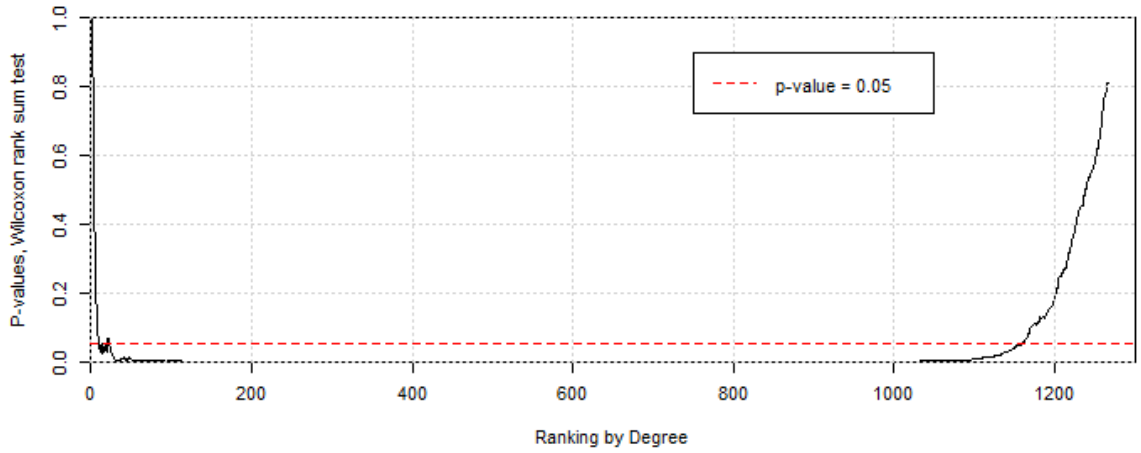
S7 Fig. The length of the shortest paths or steps between any two nodes in the networks are calculated and plotted by their frequency.

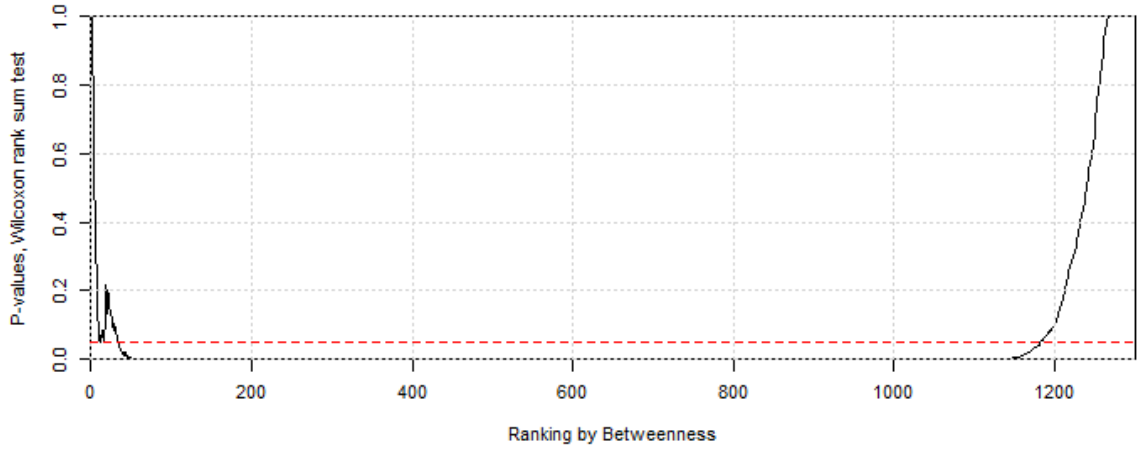
S7 Text. What are the Temporal Dynamics of the General Healthcare Network?

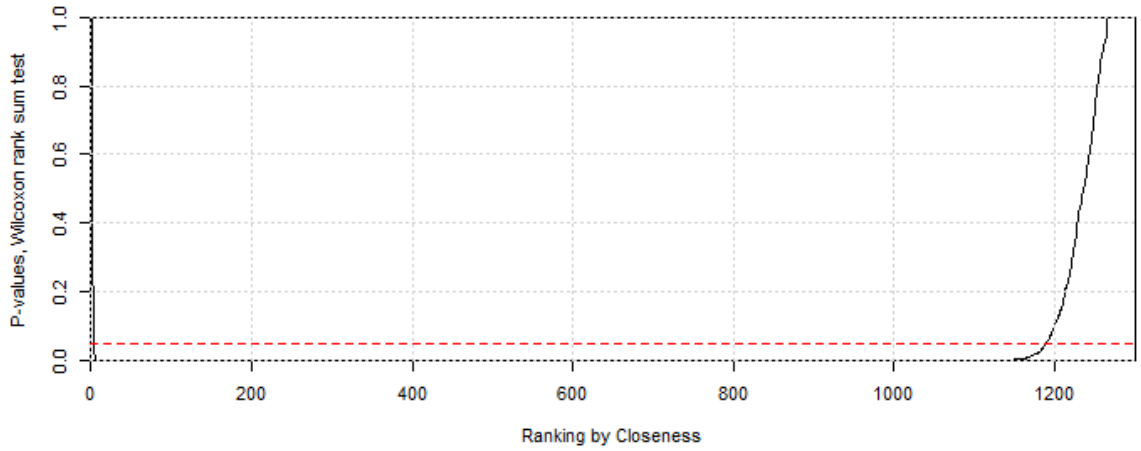
To investigate any existing temporal dynamics of patient admissions, we assessed the number of admissions by month. Patient admissions per month remained stable over the year with a small drop in August during the holidays with an overall monthly average of 979,142 admissions.

We assessed and compared the monthly healthcare networks from the general patient population in order to identify how the healthcare networks changed in size and how patterns of patient flows compared over time. On average, monthly networks were composed of fewer hospitals than the cumulative yearly network (2063 hospitals compared to an average 1218 hospitals 95% CI [1080-1357], p-value < 0.001, t-test). As a result, the number of hospitals connections and the number of patients moving between them was reduced over monthly intervals. We also observed that these networks were slightly less clustered (0.23 GCC in the cumulative network versus an average 0.17 GCC 95% CI [0.17-0.18], p-value < 0.001, t-test) with a larger diameter (30 in the cumulative network versus an average 45.83 95% CI [40.74-50.93], p-value < 0.001, t-test) and path length (2.99 in the cumulative network versus an average 4.86 95% CI [4.77-4.95], p-value < 0.001, t-test) per month on average. Regarding Map Equation communities, monthly networks on average had a smaller number of communities including when considering only communities with more than one hospital ($p < 0.001$, t-test). On the other hand, there was a larger number of Greedy communities in the monthly networks overall but no difference in number and localization of communities when considering only communities with more than one hospital. Monthly communities may be less clustered and patients may not visit all of the hospitals each month but they still retained the same regional patient sharing patterns seen in the annual network.

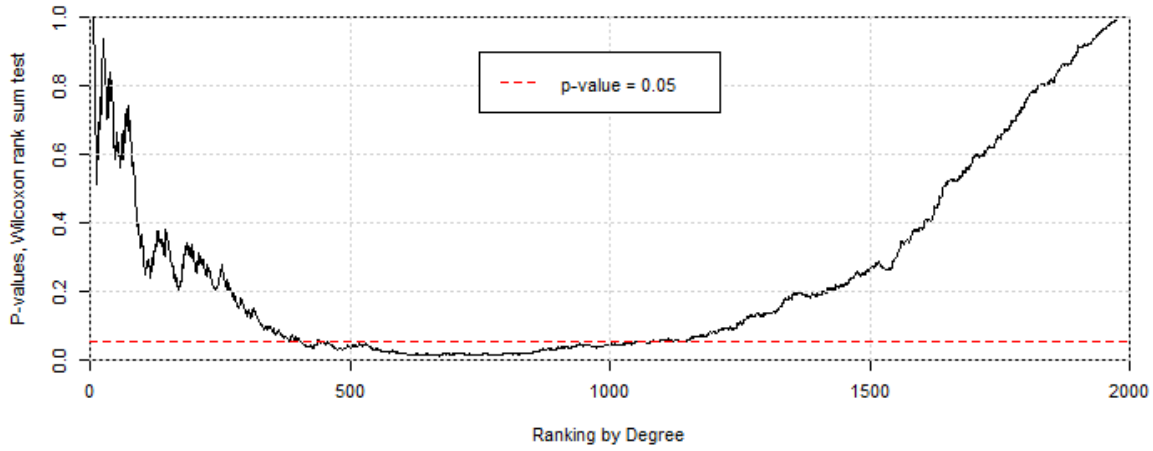
S8 Fig. Distributions of p-values of hospital rank subsets using the Wilcoxon rank sum test in the HAI-specific network compared to the general network hospital ranks

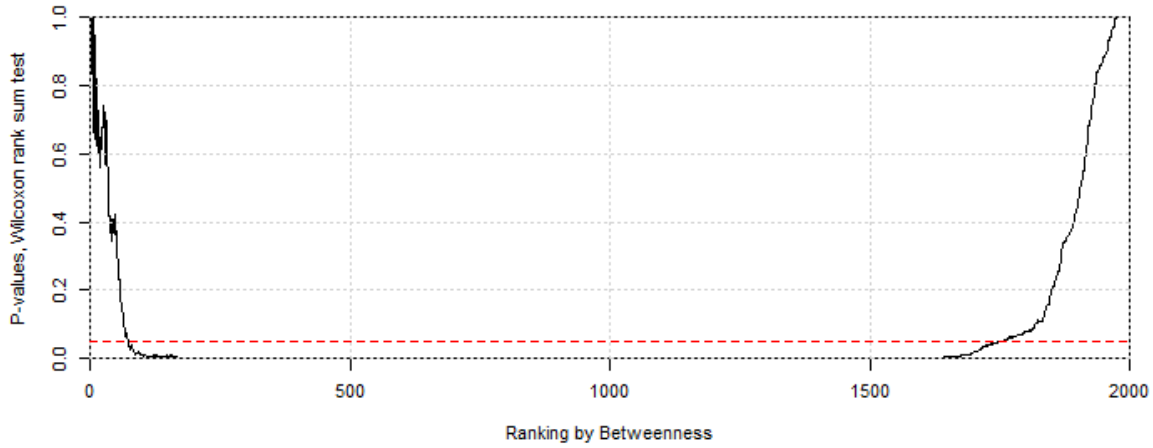


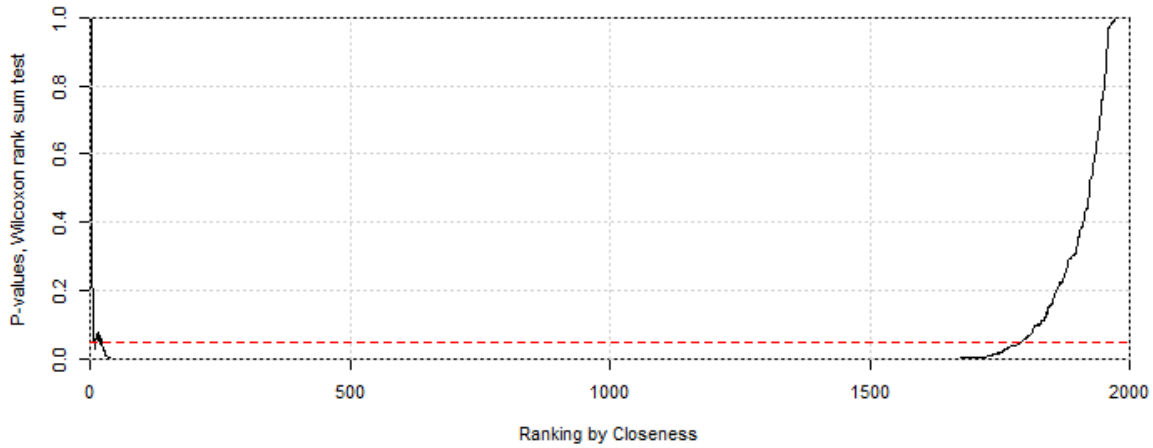




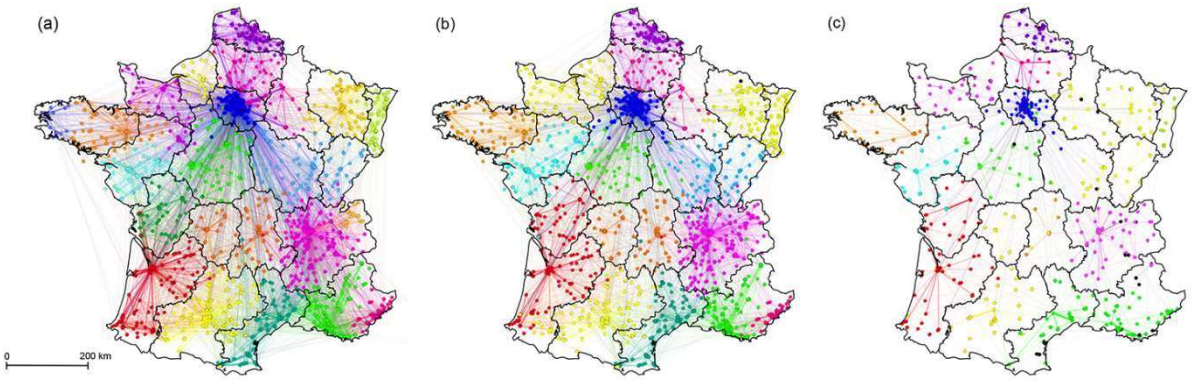
S9 Fig. Distributions of p-values of hospital rank subsets using the Wilcoxon rank sum test in the suspected-HAI network compared to the general network hospital ranks





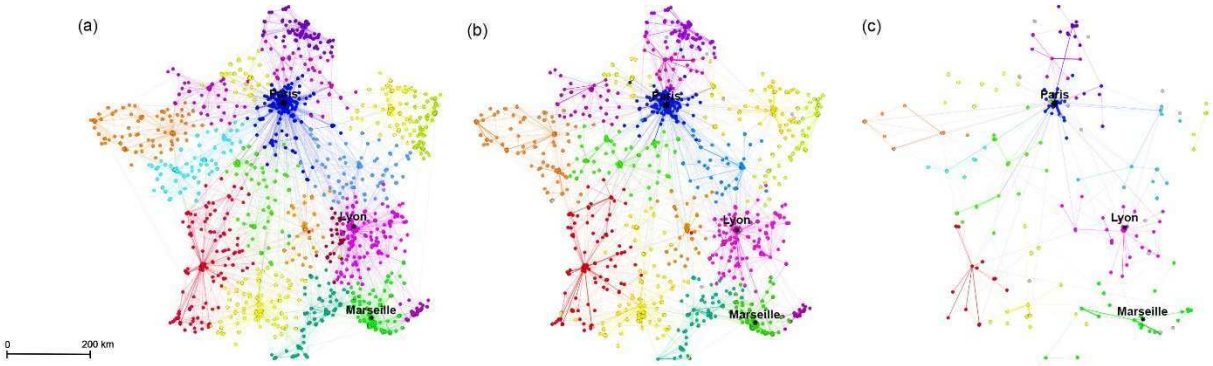


S10 Fig. Healthcare Networks by Age for All Patients



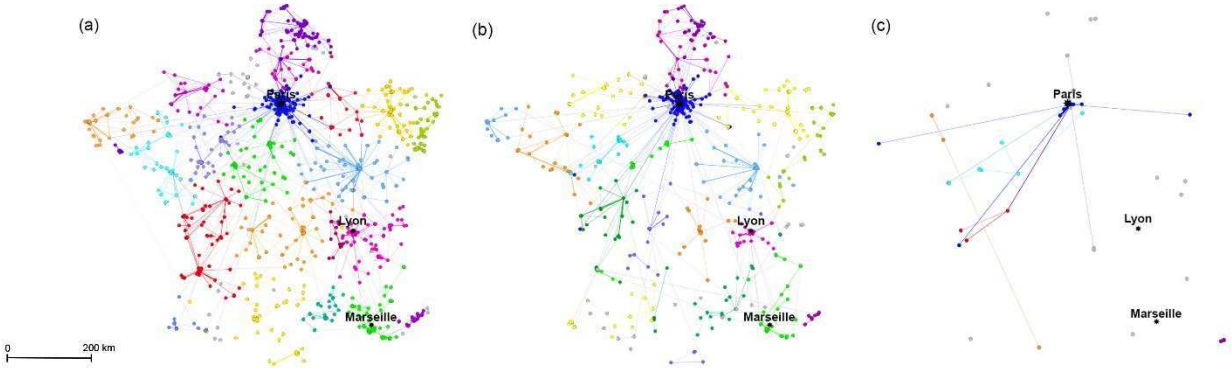
S10 Fig. Healthcare networks for all transferred patients (a) aged 18 and younger (b) aged 18 to 60 (c) older than 60 years old. Greedy communities were colored by the corresponding general network community regions with additional community color if not present for (a) 23 communities in which 3 were black (<5 hospitals) (b) 17 communities in which 1 was black (<5 hospitals) (c) 29 communities in which 14 were black (<5 hospitals).

S11 Fig. Healthcare Networks by Age for Suspected-HAI Patients



S11 Fig. Healthcare networks for all transferred patients (a) aged 18 and younger (b) aged 18 to 60 (c) older than 60 years old. We detected a total number of Greedy-based communities for each age network (a) 22 total and 19 with over 2 hospitals from a network of 1894 hospitals and 11234 edges (b) 30 total with 17 with over 2 hospitals from a network of 1559 hospitals and 5423 edges (c) 29 total with 14 with over 2 hospitals per community from a network of 218 hospitals and 318 trajectories.

S12 Fig. Healthcare Networks by Age for HAI-Specific Patients



S12 Fig. Healthcare networks for all transferred patients (a) aged 18 and younger (b) aged 18 to 60 (c) older than 60 years old. Network communities are detected using the Greedy algorithm and colored according to community membership for (a) 1143 hospitals and 2260 edges with 50 total communities with only 28 composed of more than 5 hospitals (b) 603 hospitals and 593 edges with 41 total communities and 19 with over 5 hospitals (c) 44 hospitals and 33 edges, 18 total communities, 2 communities with more than 5 hospitals, and 9 communities with more than 2 hospitals.

8.2 Additional analyses

In this section, we present preliminary results from simulations of a generic epidemic model over the French healthcare networks. This model was developed using the 2012 transfer data. Hence, we start by describing the 2012 healthcare networks.

8.2.1 The 2012 healthcare networks

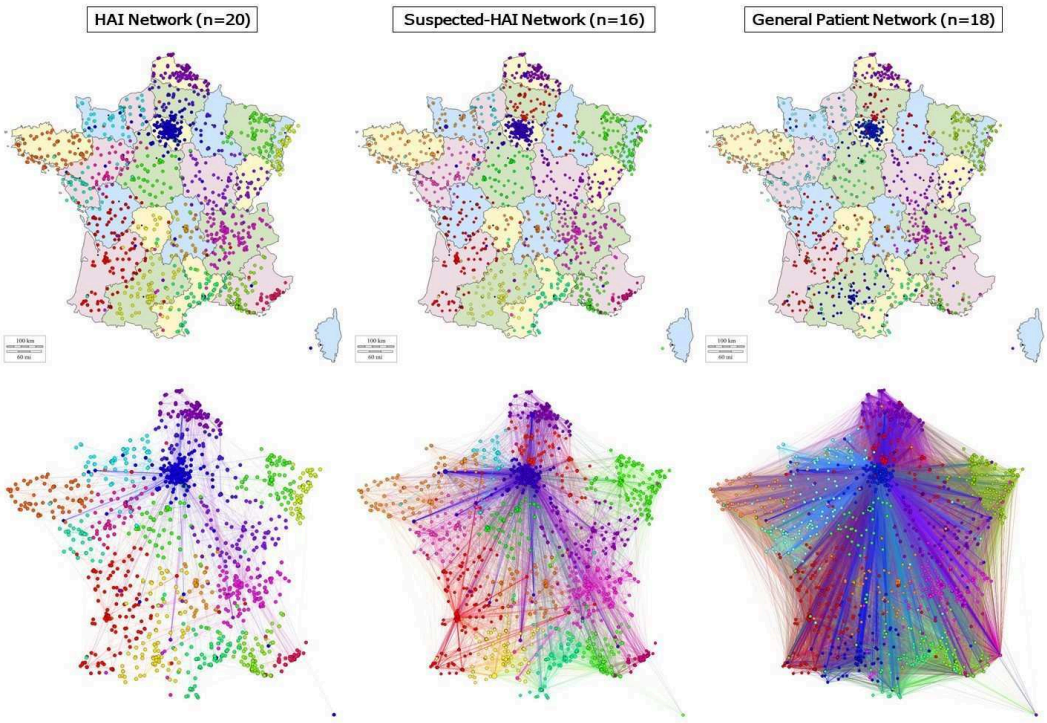
Patient discharge summaries from the PMSI database in 2012 were used to construct three healthcare networks (Table 5, Figure 12). For all 8.1 million transfers of 2.2 million patients in 2012, the general patient network was composed of 2737 different healthcare facilities and 153 665 trajectories. The average degree in the general patient network was 112 and a total of 18 communities were identified using a "Greedy" detection algorithm.(192) The clustered facilities were within an average of 39 kilometres away from one another.(192) The suspected-HAI network was based on the diagnoses identified in the Gerbier et al. publication also used in the 2014 suspected-HAI network.(18, 191) The suspected-HAI patient transfer network was composed of 2 184 facilities with 25 879 different patient trajectories. There was an average weighted degree of 7.8 with a minimum of one and a maximum of 589 patients moving per trajectory. On average healthcare facilities were connected to 23.7 other facilities in the network, connected to an average 11.85 hospitals by their in-degree and an average 11.85 hospitals by their out-degree. Healthcare facilities transferred an average total of 185 patients. There were 16 communities with an average of 20 kilometres between hospitals of the same community. The HAI network was based on any patient with the ICD-10 code of Y95 for nosocomial conditions as their principal, related, or associated diagnosis in the medical, surgery, obstetric hospitals (MCO) and postoperative and rehabilitation centres (SSR). The HAI network was composed of 1 770 healthcare facilities with 5 225 patient transfer trajectories. There average weighted degree was 3.55 and 20 Greedy communities were identified.

Table 5. Characteristics of the 2012 French healthcare networks

	HAI Network	**Suspected-HAI Network**	**General Patient Network**
Patients	21 047	392 537	2 205 799
Patient transfers	18 569	201 869	8 138 254
Hospitals (vertices)	1 770	2 184	2 737
Trajectories (edges)	5 225	25 879	153 665
Average path length	3.75	3.21	2.40
Average degree (SD)	5.90 (12.5)	23.7 (33)	112.27 (132.46)
Average in-degree (SD)	2.95 (8.40)	11.85 (20)	56.14 (75.39)
Average out-degree (SD)	2.95 (5.01)	11.85 (19)	56.14 (83.83)
Communities	20	16	18
Average inter-community distance (km)	39.40	20.00	39.14

SD: standard deviation

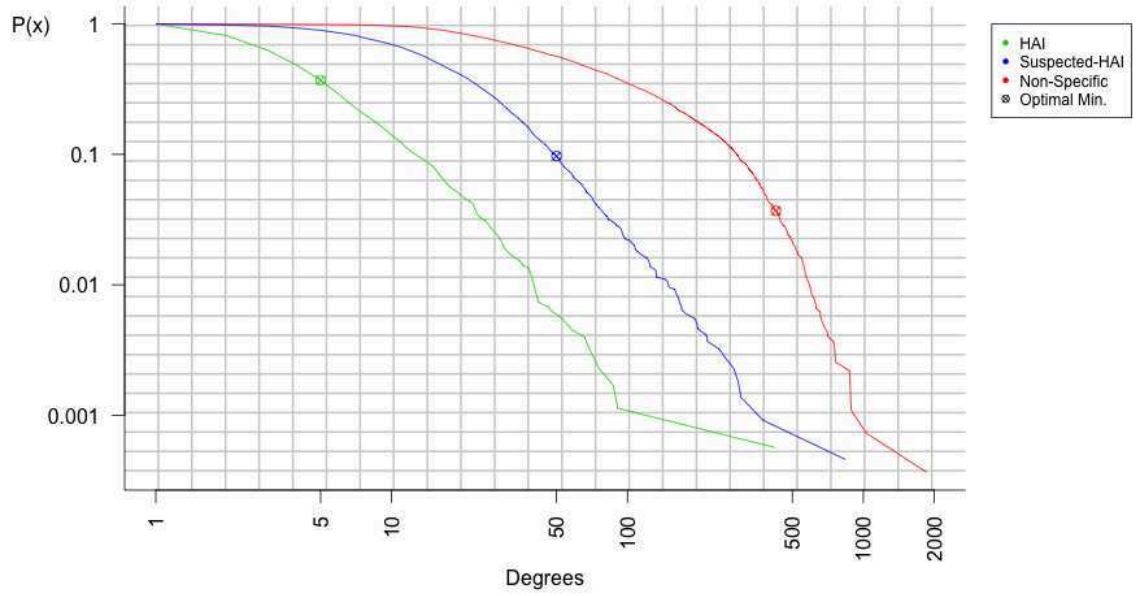
Figure 12. The 2012 healthcare networks



The HAI network (left), the suspected-HAI network (middle), and the general patient network (right) of France with (top) and without (bottom) regions and edges. N refers to the number of Greedy communities detected by the Clauset et al. algorithm.(192)

The degree distribution of the networks' vertices was assessed by the discrete power-law distribution function *P(x)* from Clauset et al.'s algorithm.(19) A Kolmogorov–Smirnov test was used to assess the fit of a power-law distribution to each network's degrees distribution (Figure 13). There was little evidence to show that a power-law distribution was not a good fit for the three networks' degree distributions (p-values > 0.05) when assessed with optimal lower cut-offs using a goodness-of-fit based approach (at 5 degrees for the HAI-defined network, at 50 degrees for the suspected-HAI defined network, and at 426 degrees for the general patient network). However, there was evidence that the power-law was not a good fit for all networks when taking into consideration all degree values (p-values < 0.05).

Figure 13. Degree distributions in the 2012 healthcare networks.



The non-specific network refers to the general patient network.

The difference in the number of healthcare facilities (n=2 063 in 2014 versus n=2 737 in 2012) may be explained by the following: a smaller number of healthcare facilities in the national FINESS database were reported in 2014 compared to 2012 (n=94 843 versus n=104 000) (193) or a fewer number of healthcare facility FINESS to which patients were transferred. Overall, the 2012 healthcare networks were very similar to the 2014 published healthcare networks in terms of structure, hub healthcare facilities, and community clustering structure.(191)

8.2.2 Simulating spread in 2012 healthcare networks

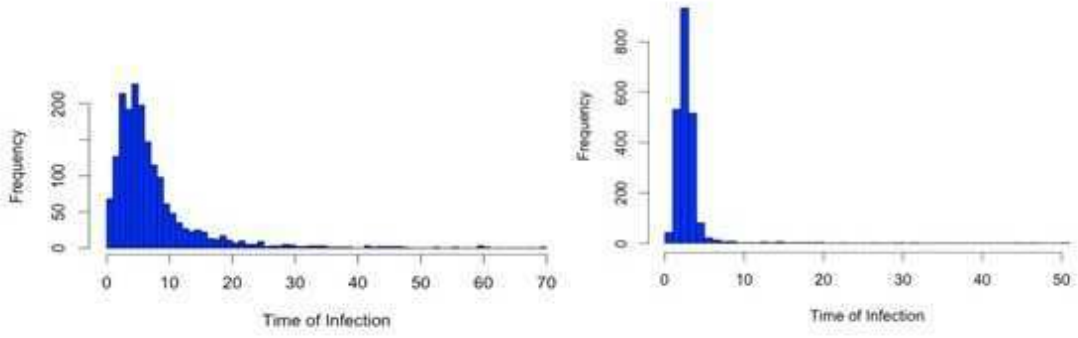
Hypothetical epidemic spread of pathogens was simulated on the 2012 HAI networks. To simulate epidemic spread, the probability of each facility becoming "infected" (meaning that at least one colonized patient stayed in the facility) was calculated. A binomial process was used where the number of infection attempts was the number of transferred patients from an infected facility. The probability of "success" was calculated as the probability of a patient becoming a new source of infection in the new healthcare facility.

In the first analysis, the networks were assessed over a period of 100 time-steps with 10 randomly chosen hospitals being initially infected. Hospital patients in infected facilities had a 10% probability of infecting a hospital where they had been transferred *(p)* in which infection lasted 4 time steps *(t)* in the facility. The simulation followed a stochastic Susceptible-Infected-Susceptible-like (SIS-like) model where hospitals were all susceptible to infection (S), could

become infected (I) and stayed infected for 4 time-steps, and recovered, becoming immediately susceptible (S) again.(194)

The incidence of newly infected hospitals per time step reached a maximum of ~200 hospitals in the HAI network and ~900 hospitals in the larger suspected HAI network (Figure 14). The epidemic died out after 20 time-steps for the HAI network and after 5 time-steps in the suspected HAI network.

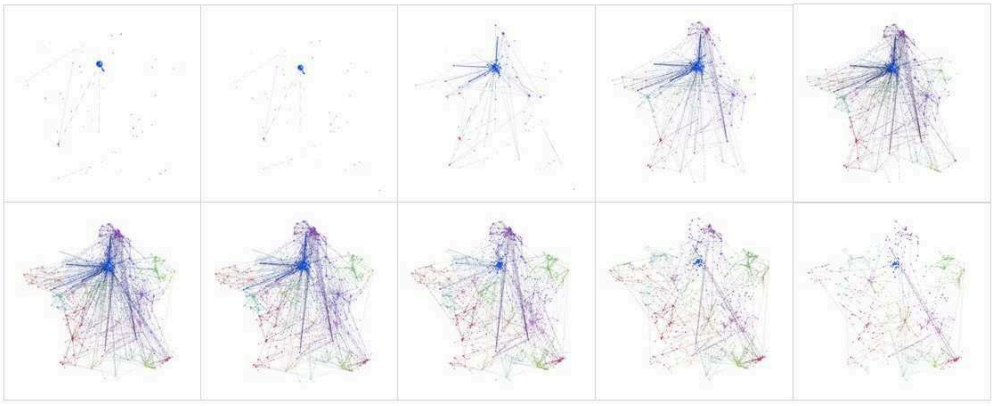
Figure 14. Simulation: histogram of the number of hospitals infected over time



HAI network (left) and suspected HAI network (right).

Figure 15 shows the first steps of a theoretical simulation of epidemic pathogen spread in the geo-localized HAI network. It can be noted that once Assistance Publique – Hôpitaux de Paris (AP-HP) became infected (the largest hub hospital centre), other large hubs became infected soon after. As a result, HAI spread throughout each community cluster's reference centre, leading to a national epidemic. Once infection died out in the regional reference centres, community level spread started to die out as well. These observations could be explained by AP-HP being the most highly connected healthcare centre and the fact that most transfers took place within the same community.

Figure 15. Simulation: epidemic spread in the HAI network

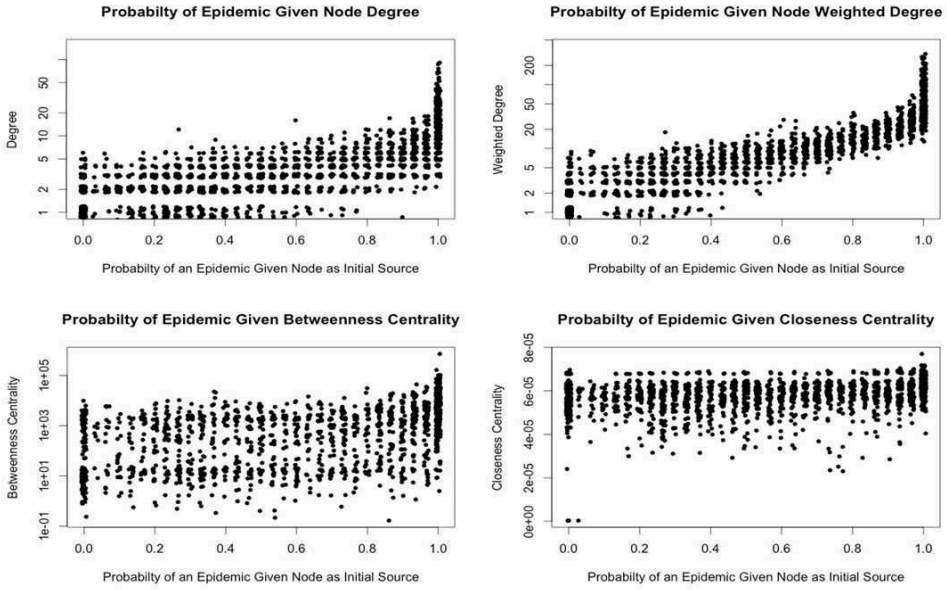


Colour is based on community clustering and node size was dependent on the degree of the facility.

In the second analysis, the vulnerability of hospitals to infection depending on the infection rate was assessed. The probability of a sustained epidemic (at least one infected hospital at the end of the simulation) was evaluated by subjecting each hospital to 30 distinct introductions of infection with 10% infectiousness (p=0.10) with each facility serving as an index case of the epidemic. The probability of hospitals sustaining an epidemic for 30 time-steps was calculated and plotted by the hospitals' characteristics.

The probability of the network sustaining an epidemic varied depending on the characteristics of the index hospital. Figure 16 shows the distribution of the probability that an epidemic occurred in the HAI network given all 1 770 hospitals were set as the index cases 30 times where they had a 10% probability of infecting other hospitals in the network. A probability of 0 meant that no hospitals were infected at the end of the 30 runs and a probability of 1 meant that at least 1 or more hospital in the network was still infected at the end of the runs.

Figure 16. Probability of sustained epidemic in the HAI network



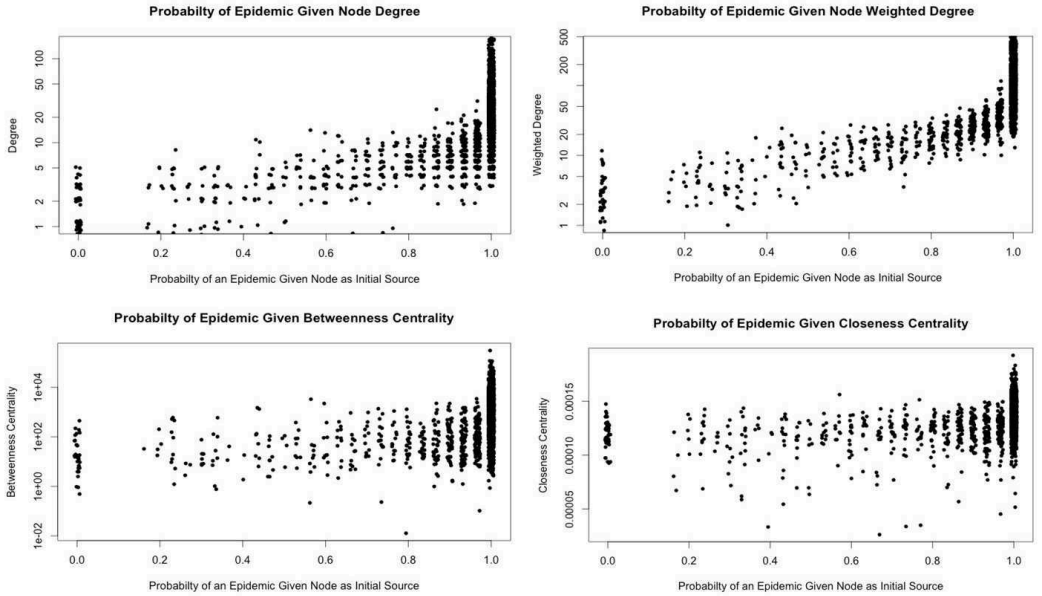
For each node in the network, 30 time-steps of initial infection (index case) were run to determine the probability of a sustained epidemic in the network at the end of the run. Each node's probability is shown in the plot. The probabilities compared to the initial infected node's degree (top left), weighted degree (top right), betweenness centrality (bottom left), and closeness centrality (bottom right). If p=0 no hospital is infected; if p=1 at least 1 or more hospitals are infected at the end of the 30 time-steps.

The initially infected hospital's degree (i.e. connectedness in the network) may explain the probability of a sustained epidemic in the network. Even though any hospitals with a degree greater than two could produce an epidemic in the network (probability=1), hospitals with higher degrees (degree >20) tended to always result in sustained epidemic spread. When taking into account the weighted degree (i.e. the number of patients being transferred per hospital), there was a greater correlation with probability of sustained epidemic spread. Index hospitals that transferred more than 50 patients tended to result in the highest probability of sustained spread in the HAI network. Regarding the initially infected hospital's betweenness centrality (the importance of a hospital acting as an intermediary in the flow of patients) and closeness centrality (the inverse of the average length of the shortest paths between one hospital and all others in the network), there was little to no observed effect of these measures on the probability of epidemic spread on the network.

In comparison, the probability of sustaining an epidemic given the facility's network characteristics was also analysed for the suspected HAI network (Figure 17). In all four plots there were no facilities that sustained an epidemic between p=0 and p=0.2 which distinguished nodes that did not sustain spread and those that had more than a 20% probability of sustaining one. The density of nodes increased as the probability of sustained epidemic increased. In this

network, nodes with a degree greater than 30 or having more than 100 patient transfers were more likely to sustain an epidemic. This was higher than the HAI network and it was expected because the suspected-HAI network is larger with higher average weighted and un-weighted degrees. For betweenness and closeness centrality, there was little effect; however, there was a clearer threshold than in the HAI network in which the hospitals with the highest values has the highest probability of sustaining an epidemic. These analyses may be of interest because they may enlighten novel infection control strategies targeting certain hospitals in the network based on their characteristics and potential impact on epidemic spread of pathogens.

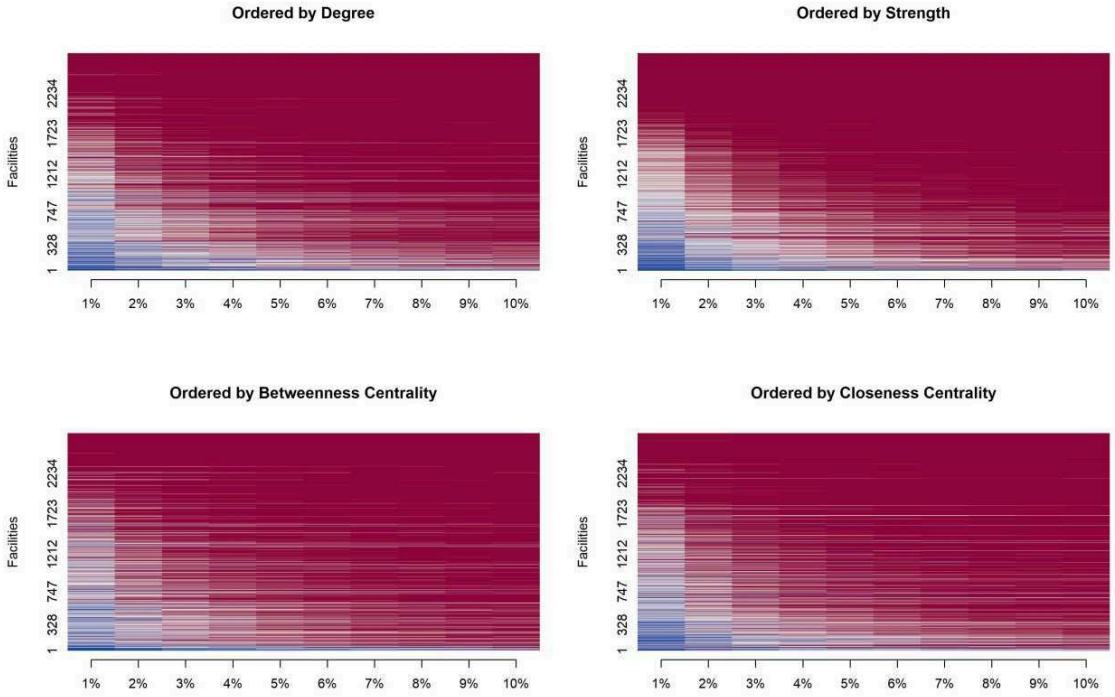
Figure 17. Probability of sustained epidemic in the suspected-HAI network



The probability of hospitals sustaining an epidemic in the network for at least 30 time-steps was calculated and plotted by the hospital's network characteristics. The probabilities are compared to the initial infected node's degree (top left), weighted degree (top right), betweenness centrality (bottom left), and closeness centrality (bottom right). If p=0 no hospital is infected; if p=1 at least 1 or more hospitals are infected at the end of the 30 time-steps.

In a third analysis, each hospital in the general patient network was subjected to 30 epidemic initiations with varying infectiousness (p) from 1% to 10% at increments of 1% with the hospital serving as an index case of the epidemic (Figure 18). For infectiousness greater than 5%, most hospitals in the network had a 100% probability of sustaining an epidemic. For infectiousness at 1%, the probability of sustained epidemic spread was higher for hub hospitals with the highest degree, weighted degree or strength, betweenness centrality, and closeness centrality while hospitals with the lowest of these measures had a very small probability of sustaining an epidemic.

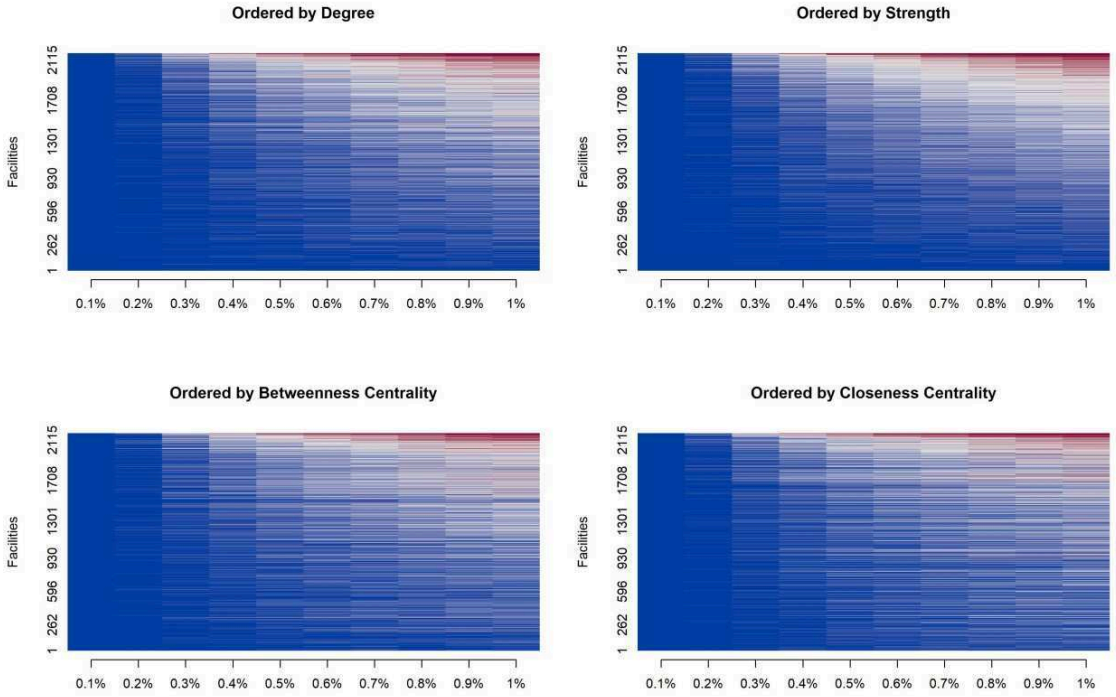
Figure 18. Probability of epidemics varying hospital infectiousness in the general network



Each hospital in the general patient network was subjected to 30 epidemic initiations with varying infectiousness (p) from 1% to 10% at increments of 1% with the hospital serving as an index case of the epidemic. The duration of the epidemic lasted 4 time steps. The probability of hospitals sustaining an epidemic in the network for at least 50 time steps was calculated. The probability of infection ranged from 0% (blue), 50% (white), to 100% (red). The probabilities were ordered by the hospital's network characteristics: degree (top left), strength (top right), betweenness centrality (bottom left), closeness centrality (bottom right).

The suspected-HAI network was also subjected to 30 epidemic initiations with varying infectiousness (p); however, in this case, p ranged from 0.1% to 1% at increments of 0.1% with each hospital serving as an index case of the epidemic. The infectiousness was set lower than the general patient network in order to test if even in a smaller network with a much smaller average degree (Table 5), lower rates of infectiousness would still be able to sustain epidemics or not at all (Figure 19). The epidemic lasted 2 time-steps in this scenario. For infectiousness less than 0.4%, very little hospitals sustained an epidemic at more than 90% probability. At 1%, for a duration of 2 versus 4 time-steps for an infected hospital in the suspected-HAI versus the general network, there were less hospitals that sustained an epidemic at the end of the 50 time-steps.

Figure 19. Probability of epidemics varying hospital infectiousness in the suspected-HAI network



Each hospital in the suspected-HAI network was subjected to 30 epidemic initiations with varying infectiousness (p) from 0.1% to 1% at increments of 0.1% with the hospital serving as an index case of the epidemic. The duration of the epidemic lasted 2 time-steps. The probability of hospitals sustaining an epidemic in the network was calculated for 50 time-steps and the results were ordered by the hospital's network characteristics. The probability of infection ranged from 0% (blue), 50% (white), to 100% (red).

Overall, these results demonstrated that a higher duration of infection and higher infectiousness in any of the networks could lead to a higher probability of a sustained epidemic. Notably, hospital degree and strength were shown to be good predictors of epidemic spread and sustainability. For the highest degree and strength in terms of number of patients transferred annually by a hospital, almost 100% of the epidemic were sustained when these hospitals served as index cases. Betweenness and closeness centrality were also predictors of epidemic spread; however, the gradient was always clear for all ranges of values but more evident for the highest values. These observations may be explained by the fact that the most connected hospitals can disperse infection to a large number of other healthcare facilities (high out-degree and weighted out-degree) which in turn allows other facilities to sustain the epidemic while it dies out in the most connected facilities until they become re-infected again because they also have a high in-degree making them also susceptible to infection. These preliminary observations from a simple SIS-like model further highlight the important role that hub healthcare facilities play in both healthcare network structure and potentially in epidemic spread of pathogens in healthcare settings.

Part Four: Spread of CPE over the French healthcare networks

Chapter 9. Temporal dynamics of French CPE episodes

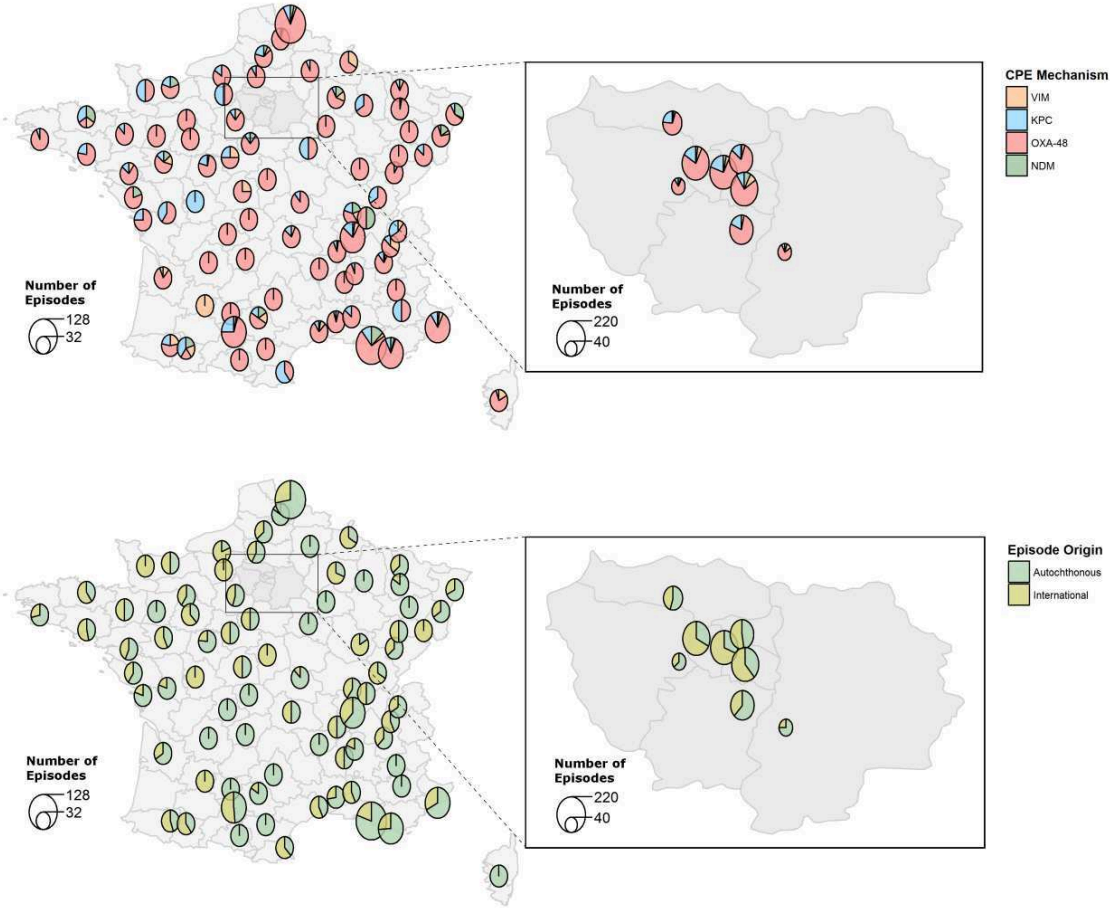
A rising number of CPE episodes have been reported in France over the years.(9, 96) The objectives of this part of the thesis were to describe CPE episode notifications from September 2010 to December 2015 and to predict the number of monthly CPE episodes from the January 2016 to December 2019 using time series models. It should be noted that only data for the department and not the hospital(s) associated with the episode were available; therefore, the results show aggregates of the data at the French department-level.

9.1 The French CPE data from 2010 to 2015

Surveillance data on CPE episodes in France from September 2010 to December 2015 were collected by Santé Publique France through the RAISIN active surveillance system. CPE episodes were defined as a case or group of cases infected with the same strain of CPE known to have been in contact with one another and identified by authorities during the outbreak investigation. Episodes were described by the number of individual cases involved in the chain of transmission, department, episode date (date of the first detected case), mechanisms of CPE resistance, bacterial species, and site of infection or colonization if known. Episodes were classified as "imported" if the index case of the chain of transmission was initially infected or colonized in a foreign country.

A total of 2 346 episodes of CPE were reported in France between September 2010 and December 2015. Out of 2 346 episodes, 2 067 included only one case and 279 included two or more cases per episode. Most CPE episodes were of the class D carbapenemase OXA-48 with a total of 1 747 episodes, followed by NDM with 332 episodes, KPC with 118 episodes, and VIM with 101 episodes. A total of 1 110 episodes were linked to internationally imported cases during the entire period. The episode with the highest number of cases (n=194) occurred in September 2012 of an OXA-48 strain followed by a 149-case outbreak in October 2012 of OXA-48. Both outbreaks were not linked to international importation of OXA-48 strains. The spatial distribution of CPE episodes by mechanism type and importation status are shown in Figure 20.

Figure 20. Spatial distribution of CPE episodes, France 2011-2015



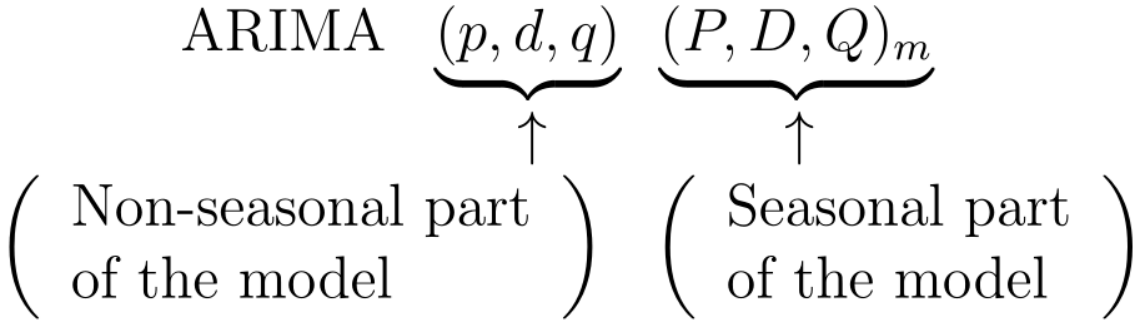
The majority of the episodes (n = 469) occurred in Paris within the Ile-de-France region with the majority (n = 324) being linked to an international case (Figure 11). The second and third most common number of episodes occurred in neighbouring departments Val-de-Marne (n = 83) and Hauts-de-Seine (n = 159). These episodes were dominated by OXA-48 (n = 347 in Paris and n = 208 in Val-de-Marne respectively) followed by NDM (n = 95 in Paris and n = 39 in Hauts-de-Seine respectively). Most cases of KPC are found in the Val-de-Marne (n = 23) followed by Paris (n = 20). VIM cases, the least geographically dispersed, were most common in Val-de-Marne (n = 21) and Marseille (n = 20). OXA-48 cases were the most common and most dispersed covering 87 out of a total 101 departments (95 continental departments including Corsica, Monaco, and 5 overseas departments).

9.2 Modelling the evolution in time of CPE episodes in France

9.2.1 Time series model

In order to predict the number of potential CPE episodes we could expect in the future, we used a simple approach to estimate future trends and shed light on the growing importance of CPE incidence in France. The number of CPE episodes was forecasted using the autoregressive integrated moving average (ARIMA) time series models with seasonality (SARIMA). The (SARIMA(*p,d,q*)(*P,D,Q*)*m*) model combines both the ARIMA forecasting equation ARIMA(*p,d,q*) of *p*, the number of auto-regressive terms, of *d,* the number of non-seasonal differences needed for stationarity, and *q*, the number of lagged forecast errors with seasonal terms *P*, *D*, and *Q* by the number of seasonal units *m* (Figure 21).(195) The seasonal model adds components *P* (the number of seasonal autoregressive terms), *D* (the number of seasonal differences), and *Q* (the number of seasonal moving average terms). The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were used to select the best SARIMA model for each episode type separately. Each selected model was used to forecast 4 years of episodes with calculated 80% and 95% prediction intervals.

Figure 21. SARIMA model by Hyndman & Athanasopoulos.



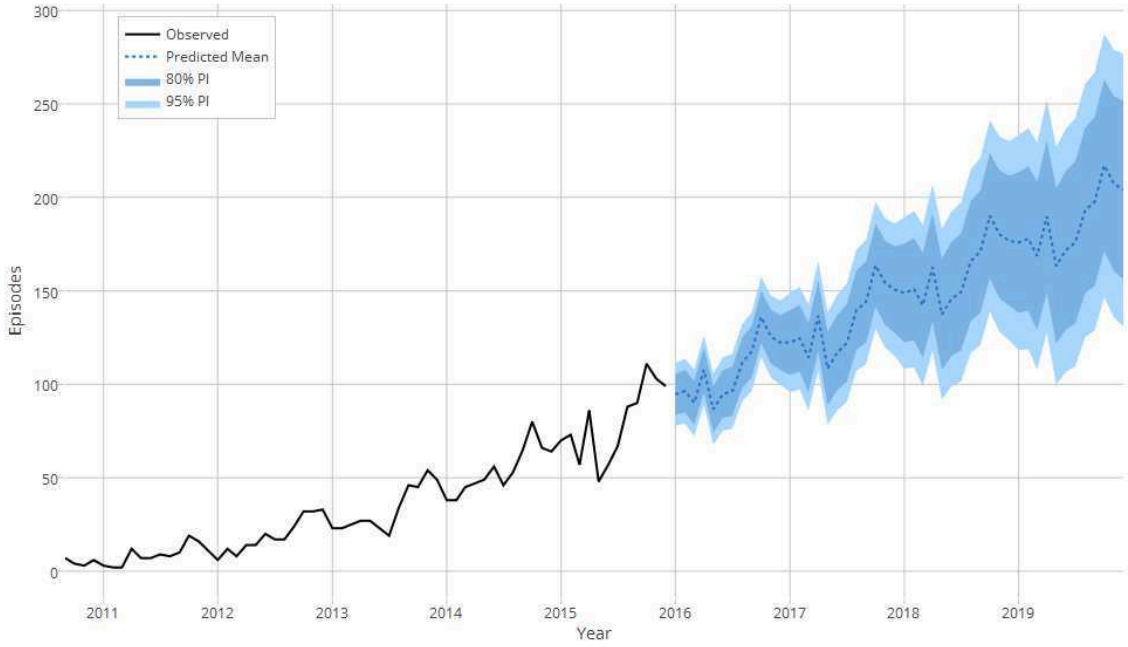
9.2.2 Forecasting results

Decomposition of the data from September 2010 to December 2015 showed an increasing trend (green) and displayed seasonality (orange) with a peak number of episodes during the month of October (Figure 22).

Figure 22. Decomposition of time series of CPE episodes, France 2011-2015

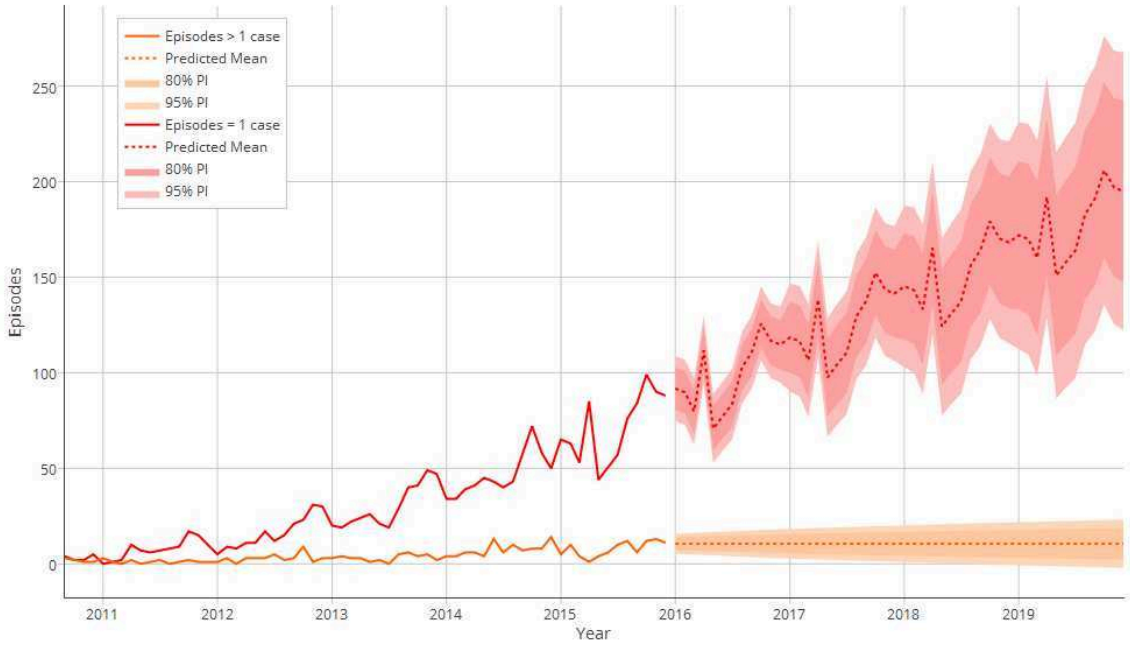
Using the entire dataset, the SARIMA model was used to predict the number of episodes expected from 2016 to 2019. The model predicted an increasing number of episodes over time: an estimated average of 122 CPE episodes per month by the end of 2016 (95% PI [100-149], 80% PI [108-140]), 151 CPE episodes per month by the end of 2017 (95% PI [115-186], 80% PI [128-174]), 177 CPE episodes per month by the end of 2018 (95% PI [124-230], 80% PI [142-212]), and 204 CPE episodes per month by the end of 2019 (95% PI [131-277], 80% PI [156-252]) (Figure 23).

Figure 23. Forecasting CPE episodes in France



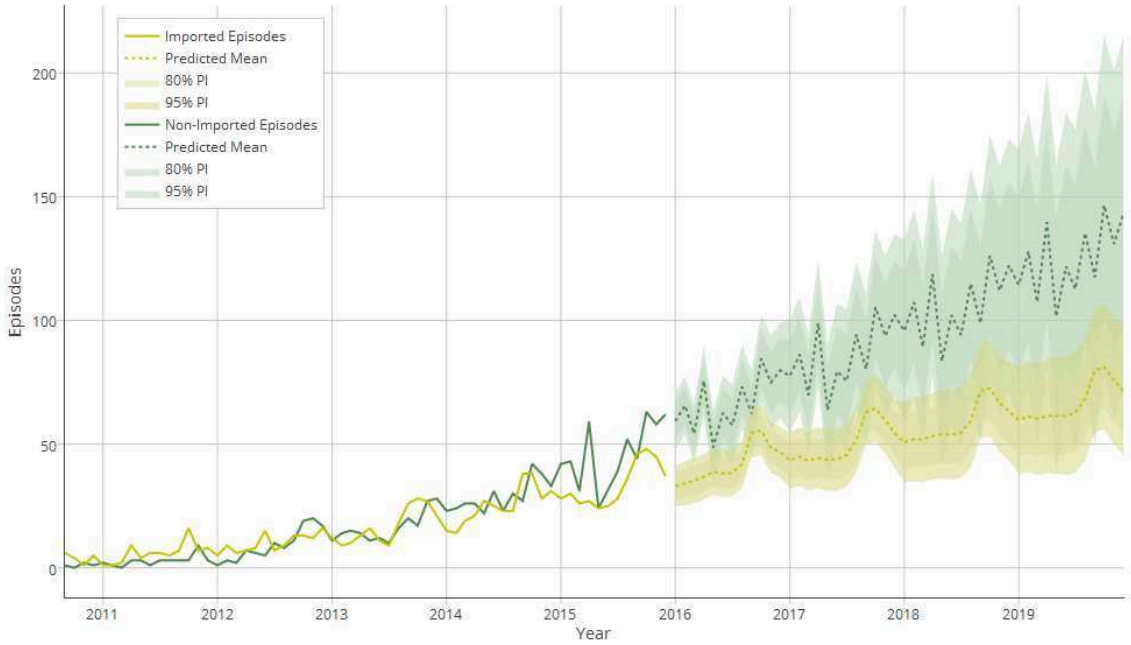
There was a predicted rise in the number of single case episodes (up to 200 single-case episodes per month by the end of 2019) and stabilization of episodes with more than one case (less than 25 multiple-case episodes per month by the end of 2019) (Figure 24). This may suggest that the control measures in place during the 2010-2015 period were able to control transmission between infected individuals that could have resulted in recurring hospital outbreaks. It should be noted that episodes with more than two cases used to forecast episodes represented a range between 2 to 200 CPE cases, therefore representing both small multiple case episodes and larger outbreaks.

Figure 24. Forecasting single-case and multiple-case CPE episodes



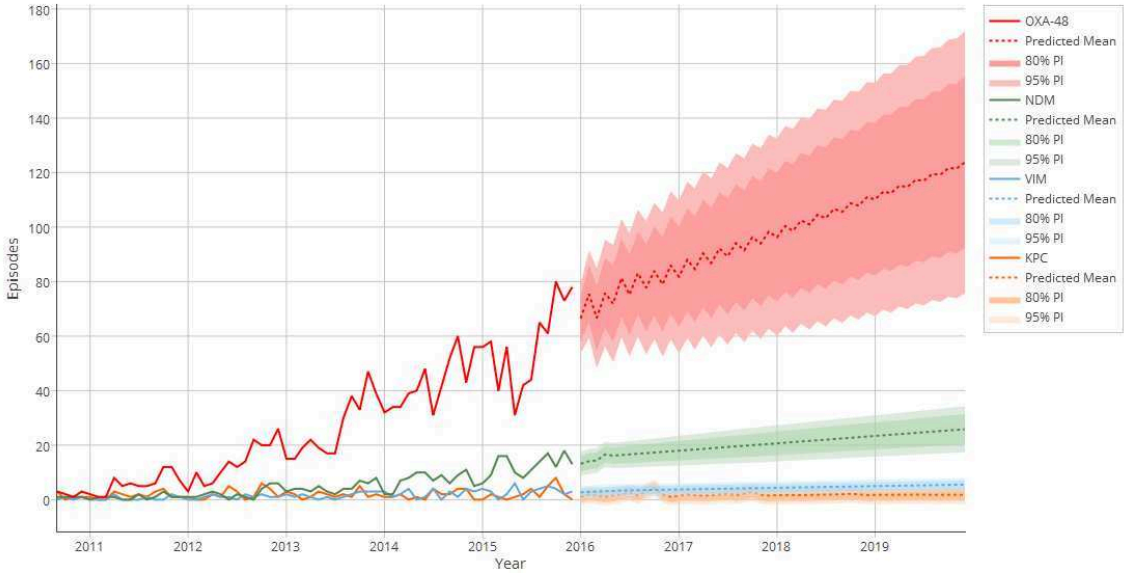
The number of episodes not associated with an internationally imported case was predicted to increase at a higher rate than that of episodes associated with an internationally imported case, suggesting that local spread would sustain the epidemic (Figure 25). An average of almost 150 episodes per month of episodes with no link to international cases were predicted to occur by the end of 2019. Episodes linked to international cases of CPE infection were also predicted to increase but only up to 75 episodes by the end of 2019. There were clear trends of seasonality of imported episodes with a predicted peak number of episodes in October every year.

Figure 25. Forecasting imported and non-imported CPE episodes



In France, OXA-48 producing *Enterobacteriaceae* were the most common CPE infection. OXA-48 episodes were predicted to continue to dominate at a higher rate than that of NDM, KPC, and VIM (Figure 26). In 2014, the number of NDM episodes was higher than KPC and VIM. By the end of 2016, 2017, 2018, and 2019, an average of, respectively, 86 (95% PI [59-113]), 98 (95% PI [63-134]), 111 (95% PI [69-153]), and 124 (95% PI [76-172]) OXA-48 CPE episodes per month were predicted. NDM cases were also forecast to increase with a predicted average of 26 (95% PI [17-34], 80% PI [20-31]) episodes per month by the end of 2019. Monthly episodes of VIM and KPC were predicted to stabilize. The model predicted between 4 and 8 VIM episodes per month and between 0 to 5 episodes of KPC per month by the end of 2019.

Figure 26. Forecasting CPE episodes by mechanism of resistance



In conclusion, the number of CPE episodes in France has increased over the years. The number of episodes were predicted to increase reaching up to twice as many episodes by the end of 2019 compared to the end of 2015. Episodes that were not linked to international importation were predicted to dominate and sustain the epidemic. It is important to note that these results are a reflection of the surveillance data which may have been subject to detection and deceleration bias. However, these alarming trends highlight the need for more precautions to be taken to control the local spread of CPE in the coming years.

Chapter 10. Third article - Contribution of patient transfers on the spread of CPE in France

10.1 Summary

As detailed in chapter 9, we showed that an increasing number of CPE episodes in France have been declared over the years, many of which have been associated with cross-border and local patient transfer between healthcare facilities.(9) Despite France updating its national guidelines for extensively drug-resistant bacteria in 2013, regional and inter-regional spread of CPE in France was still being reported.(82) Therefore, to combat the spread of CPE, the impact of patient transfers on spread must be better understood. Since the impact of patient transfers on spread may vary, our study aimed to assess the contribution of the patient transfer network on CPE spread in France from 2012 to 2015.

Using the healthcare network of patient transfers in France in 2014 (191), we used a previously published statistical method to empirically test the contribution of the network on CPE spread. Out of the total 2 273 CPE episodes reported between 2012 and 2015, we identified the most likely potential infector episodes of each non-imported incident case (n=1 251) by selecting candidate transmitter episodes with the shortest path distance to each incident episode. The distribution of shortest path distances was compared to 500 simulations of permutations of the data. Multiple spreading events and the spatial distribution of potential infectors and incident episodes were also described.

Ninety percent of incident episodes had an identified potential infector episode for the entire study period; however, when stratifying the data by year, only episodes in 2013, 2014, and 2015 had significantly shorter path distances compared to permutations. This suggested that the CPE epidemic in France transitioned from an epidemic sustained importation of episodes from other countries before 2013, to an epidemic sustained by local transmission events supported by patient transfers. In addition, the number of events linking potential infectors to multiple episodes increased over the years suggesting that highly connected metropoles may have led to outbreaks through patient transfer.

Assessing the role of inter-hospital patient transfer in the spread of carbapenemase-producing *Enterobacteriaceae*: the case of France between 2012 and 2015

Narimane Nekkab[1,2,3], Pascal Crépey[3], Pascal Astagneau[4,5], Lulla Opatowski[6,7,8], Laura Temime[1,2]

Affiliations

[1] Laboratoire MESuRS, Conservatoire National des Arts et Métiers, Paris, France
[2] Institut Pasteur, Conservatoire National des Arts et Métiers, Unité PACRI, Paris, France
[3] University Rennes, EHESP, REPERES (Recherche en pharmaco-épidémiologie et recours aux soins) – EA 7449, Rennes, France
[4] Centre régional de prévention des infections associées aux soins (CPias), Paris, France
[5] Département de santé publique, Médecine Sorbonne Université, Paris, France
[6] University Versailles St. Quentin, UMR 1181, B2PHI, Villejuif/Paris, France
[7] Institut Pasteur, UMR 1181, B2PHI, Paris, France
[8] Inserm UMR 1181 « Biostatistics, Biomathematics, Pharmacoepidemiology and Infectious Diseases » (B2PHI), Villejuif/Paris, France

Abstract

The spread of carbapenemase-producing *Enterobacteriaceae* (CPE) healthcare-associated infections is a major public health threat that has been associated with cross-border and local patient transfer between healthcare facilities. Since the impact of transfers on spread may vary, our study aimed to assess the contribution of the patient transfer network on CPE spread in France from 2012 to 2015. **Methods:** Using the French healthcare network of 2.3 million patients, we extended a previously proposed statistical method and tested the ability of this network to support 2273 observed CPE incidence episodes over the study period. We aimed to identify the most likely infector for 1251 non-imported episodes using network-supported paths (NSPs) and compared observed NSP distances to those expected by chance, using random permutations of the data. **Results:** Ninety-percent of non-imported incident CPE episodes were linked to potential infector episodes. NSP distances were significantly shorter in the observed data than expected by chance from 2013 to 2015, but not in 2012. Linked episodes tended to occur within close geographic distances. Multiple spreading events in which potential infectors were linked to multiple secondary incident episodes were identified. The baseline time window of 21 to 28 days between candidate transmitters and the incident episode was supported by the sensitivity analysis. **Conclusions:** We observed a transition in 2013 from an epidemic sustained by importation to being sustained by local transmission events. As a consequence, coordinated prevention and infection control strategies should now focus on transfers of carriers of CPE to reduce regional and inter-regional transmission.

Keywords: carbapenem-producing *Enterobacteriaceae*, healthcare networks, patient transfers

Introduction

The increasing number of carbapenemase-producing *Enterobacteriaceae* (CPE) strains poses a major threat to healthcare systems and jeopardizes patient safety.[1-3] An alarming report by the European Centre for Disease Control (ECDC) in 2015 described regional and inter-regional spread of CPE including four countries in which CPE has become endemic.[3] CPE spread and subsequent outbreaks have been linked to transfers of patients between healthcare facilities within countries and across national borders. [4-6] Due to the unlikeliness that new antibiotics to replace carbapenems will become available in the near future, controlling CPE spread across healthcare systems is essential. This requires better understanding of the impact of transfer patterns on the spread of nosocomial pathogens.

Over recent years, several studies have correlated measures of hospital connectivity in healthcare networks of patient transfers to CPE incidence in healthcare facilities, underlining the importance of coordinated control measures at the regional scale.[7-11] However, many questions have been left unanswered. In particular, despite studies documenting the transmission chains of cross-border transfer of CPE strains and hospital outbreaks, the overall contribution of inter-hospital transfers on spatial dispersal of CPE over time has not yet been assessed.

In this study, we aimed to investigate the contribution of patient transfers on the spread of CPE in France and, in addition to the observations that have been made over the past few years, to further describe the epidemic. Indeed, on the one hand, an increasing number of OXA-48 cases (class D beta-lactamases with oxacillinase activity) coming into France from cross-border transfers have been reported in recent years.[3-5, 12, 13] On the other hand, despite updated national guidelines and strategies, from the 2013 to the 2014-2015 period, France advanced from an epidemiological stage-3 of regional spread to a stage-4 of inter-regional spread of CPE.[14]

In order to assess the extent to which patient transfers may have contributed to the transition from regional to inter-regional spread of CPE in France, we relied on a previously published healthcare network of patient transfers in France.[15] To empirically test the contribution of the network on CPE episode incidence, we extended a previously proposed statistical method that assessed the contribution of a contact network of patients and healthcare workers on *Staphylococcus aureus* spread in a hospital setting.[16] For the 2012 to 2015 period, we evaluated the number of CPE episodes with network-supported paths (NSPs), described the number of CPE episodes linked to patient transfer, and conducted a sensitivity analysis on the transmission window between potential infector episodes and incident episodes.

Methods

CPE episodes data

Surveillance data of 2273 CPE episodes occurring in continental France from January 2012 to December 2015 were used in the analysis. Data were collected by Public Health France through the HAI-EWRS active surveillance system. CPE episodes were defined as a case or group of cases infected with the same strain of CPE known to have been in contact with one another and identified by authorities during the outbreak investigation. Episodes were described by the number of individual cases involved in the chain of transmission, department, episode date (date of the first detected case), mechanism(s) of CPE resistance, bacterial species, and site of infection or colonization if known. Episodes were classified as "imported" if the index case of the chain of transmission was initially infected or colonized in a foreign country. We assumed that for episodes in which there were multiple cases, the cases all occurred in the same department.

Department network

The network of patient transfers between hospitals and healthcare centres in 2014 was built and described in detail in a previous study by the authors.[15] Patient transfer data was collected from the national hospital

discharge database, a comprehensive medico-administrative database of patient discharge summaries. Only direct hospital-to-hospital or medical ward transfers were considered. The hospital network was transformed into an adjacency matrix of nodes representing administrative departments, with the edges representing the connections between departments (French administrative division between the administrative regions and the communes). The department network edge weights were given as the sum of the number of hospital transfers between departments for the entire year of 2014. As shown in the previous study on the network, the number of patient transfers remained stable during the year. In this analysis, we assumed that the number of inter-department transfers in 2014 were comparable over time and therefore, we used this network for the entire study period.

Potential infector identification

In the methodology proposed by Obadia and colleagues [16], the authors used a hospital network of patients and healthcare workers contacts to identify, for incident colonization episodes of *Staphylococcus aureus*, the potential infectors that were best supported by the network in terms of path distance. The distribution of observed path distances between incident cases and their closest potential infector was compared to that obtained using randomly distributed colonization data over the same network.

Here, rather than a contact network, we used the department network of patient transfers to identify the most likely potential infector for CPE incident episodes. We assumed that transmission could have occurred equally among episodes with colonised or infected cases or both and among all genera of *Enterobacteriaceae*. Each non-imported incident CPE episode was investigated using the following algorithm:

1. All episodes involving CPE with the same mechanism(s) of resistance that occurred within a specific time window prior to the incident non-imported episode in any network department were considered "candidate transmitters"
2. All NSPs between the incident episode department and all candidate transmitter departments were compiled from a matrix of total annual transfers of patients between each department
3. For a given incident episode, the candidate transmitter with the shortest NSP length between its department and the incident episode department was considered as the most likely potential infector

In order to statistically assess the patient transfer impact, the distribution of shortest path lengths between each non-imported CPE episode and its potential infector was compared to the distribution expected under the null hypothesis of independence between CPE transmission and the department network of hospital transfers. Expected shortest path lengths under the null hypothesis were determined using a random permutation of the departments of all episodes (described in more detail in Supplement 1). Five-hundred permutations (enough to ensure stability of results) were generated and the algorithm of potential infector selection followed for each permutation. Each incident episode and its new potential infector NSPs from randomly permuted data were averaged, producing a distribution of shortest path distances expected under the null hypothesis. The observed distribution was compared to the distribution of the permutated NSP using a paired Wilcoxon signed-rank test.

Choice of a time window for candidate transmitter selection

Understanding the time it takes for CPE colonization or infection to be detected in one hospital after its contamination via patient transfers from another hospital is essential in being able to appropriately link CPE episodes. The estimation of this delay time relies on data from outbreak investigations.

The median duration of CPE outbreaks in French hospitals from 2004 to 2012 was estimated at 22 days.[17] In addition, a few studies have reported the delay time in the detection of CPE colonization or infection between healthcare facilities as a result of local patient transfers. In a multi-hospital outbreak of carbapenemase-producing *Klebsiella pneumoniae* (KPC), two patient contacts were transferred and detected positive in two other hospital facilities 15 and 29 days after the detection of the hospital index

cases.[18] In another KPC outbreak, following patient transfers out of the hospital in which the outbreak originated, KPC colonization was detected in two other hospitals respectively 19 and 25 days after detection of the index case in the original hospital.[19]

Based on this data, for this study, we chose to look for potential infectors of incident CPE episodes in a 1-week time window ranging from 21 to 28 days ($W_{[21,28]}$) before the incident episode E_i. A sensitivity analysis was conducted to compare this baseline window to a sliding 1-week window starting from $W_{[1,8]}$ to $W_{[30,37]}$ before the notification of E_i.

Distance computation in the weighted department network

In order to identify the closest potential infectors over the weighted department network, we first converted the edge weights w_{ij} to annual transfer rates t_{ij} by dividing the total number w_{ij} of transfers from department i to department j by the sum of all patient admissions in the origin department i:

$$t_{ij} = \frac{w_{ij}}{t_i}$$

We then defined as the distance from department i to j as the negative log of the transfer rate t_{ij}:

$$d_{ij} = -\log(t_{ij})$$

Shortest path distances were computed using Dijkstra's algorithm with the R package "igraph."[20] Another distance definition proposed in a recent paper by Donker et al. [21] was also investigated in a sensitivity analysis (Supplement 2).

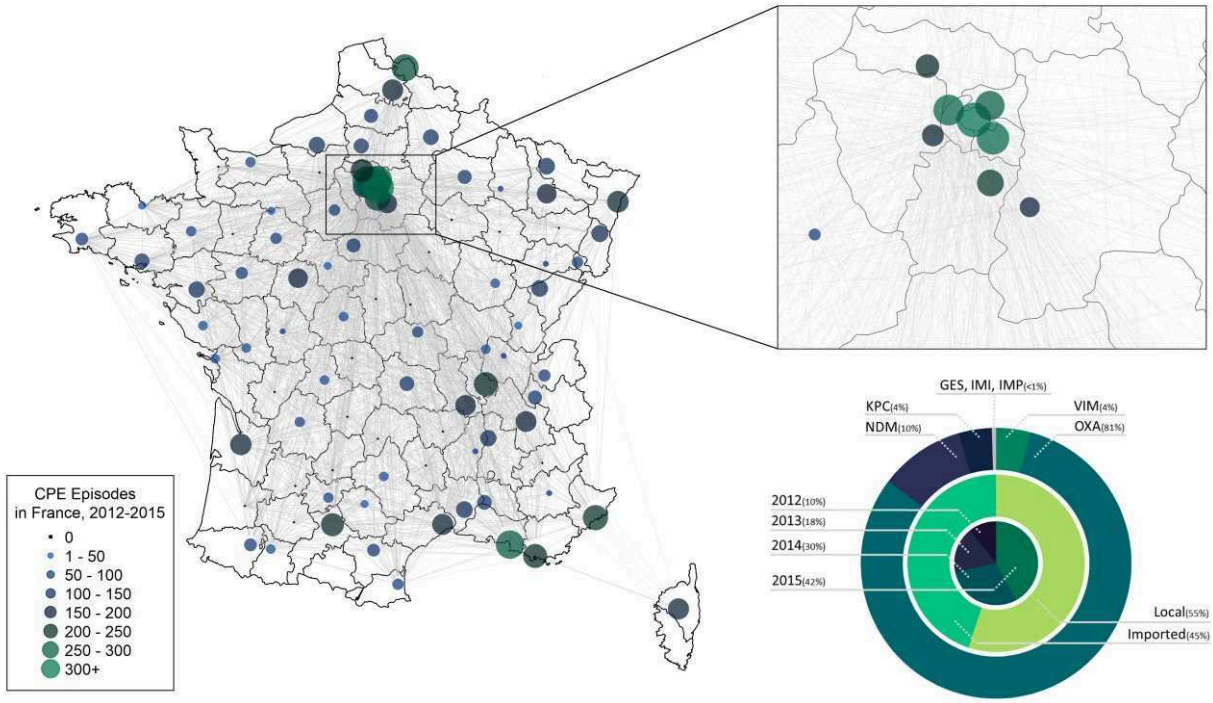
Multiple spreading events

We defined multiple spreading events as events when several incident episodes shared the same most likely potential infector. We assessed the distribution of the size of the multiple spreading events, their evolution over time, and the spatial characteristics of these events.

Results

A total of 2.3 million patients with a total of ten million direct transfers were recorded in 2014. The department network included 93 departments spanning continental France and 3326 connections comprised of 2063 hospitals and healthcare centres. Over the year, a mean of 62 patients were transferred between departments with the maximum of 8742 patient transfers within one connection. Based on the 2014 healthcare network structure, we show the number of annual CPE episodes occurring in each department over the 2012-2015 period in Figure 1.

Figure 1. Network of 2014 patient transfers between French departments and incident CPE episodes occurring between 2012 and 2015. The department network is comprised of 93 departments (the two departments in Corsica (2A and 2B) were merged and two departments (08 and 09) did not have any patient transfers) linked together by over 3000 connections (grey lines). Network components are geo-localized to the department's prefecture. For each department, the cumulated number of incident CPE episodes reported in hospitals of the department from 2012 to 2015 is depicted by a circle. The bottom right multi-level pie chart provides the proportion of incident CPE episodes for each year, along with their importation status and mechanisms of resistance.



Among the 2273 episodes included in the study, all three classes of carbapenemases (A, B and D [2]) were reported, with 81% of class D OXA-enzyme producing *Enterobacteriaceae*. The number of CPE episodes increased almost four-fold during the four-year time span, reaching a total of 956 episodes in 2015. The proportion of episodes not linked to importation from cross-border transfer increased significantly over time from 48% in 2012 to 60% in 2015 (X^2 test for trend in proportions, p-value = 7.44×10^{-5}).

Out of the total 1251 non-imported incident episodes, we aimed to identify a potential infector for each episode over the entire 2012 to 2015 period and for each year independently (Table 1). Our results showed that, when using the baseline time window, 90% or 1122 of the non-imported episodes were linked to a potential infector over the entire study period. Weighted NSP distances between the incident episodes and their closest potential infectors were significantly shorter than the NSP distances of the same incident episodes and their potential infectors from permutations of the departments (Wilcoxon paired rank sum test, p-value = 3.27×10^{-28}).

Table 1. Characteristics of 2273 carbapenemase-producing *Enterobacteriaceae* episodes in France from 2012 to 2015.

Year	2012-2015	2012	2013	2014	2015
Total episodes	2273	242	403	672	956
Imported episodes	1022	125	200	311	386
Non-imported episodes (%)	1251 (55%)	117 (48%)	203 (50%)	361 (54%)	570 (60%)
Non-imported episodes with potential infector (%)	1122 (90%)	97 (83%)	176 (87%)	307 (85%)	496 (87%)
NSP distance of observed data mean [95% CI]	5.46 [5.19-5.72]	7.17 [6.24-8.10]	5.93 [5.27-6.59]	5.91 [5.39-6.43]	4.66 [4.28-5.04]
NSP distance of permutations mean [95% CI]	5.73 [5.49-5.97]	7.07 [6.25-7.88]	6.06 [5.43-6.68]	6.07 [5.60-6.53]	5.17 [4.83-5.50]
P-value*	3.27×10^{-28}	0.28	0.004	6.90×10^{-6}	1.88×10^{-23}

NSP: network-supported path

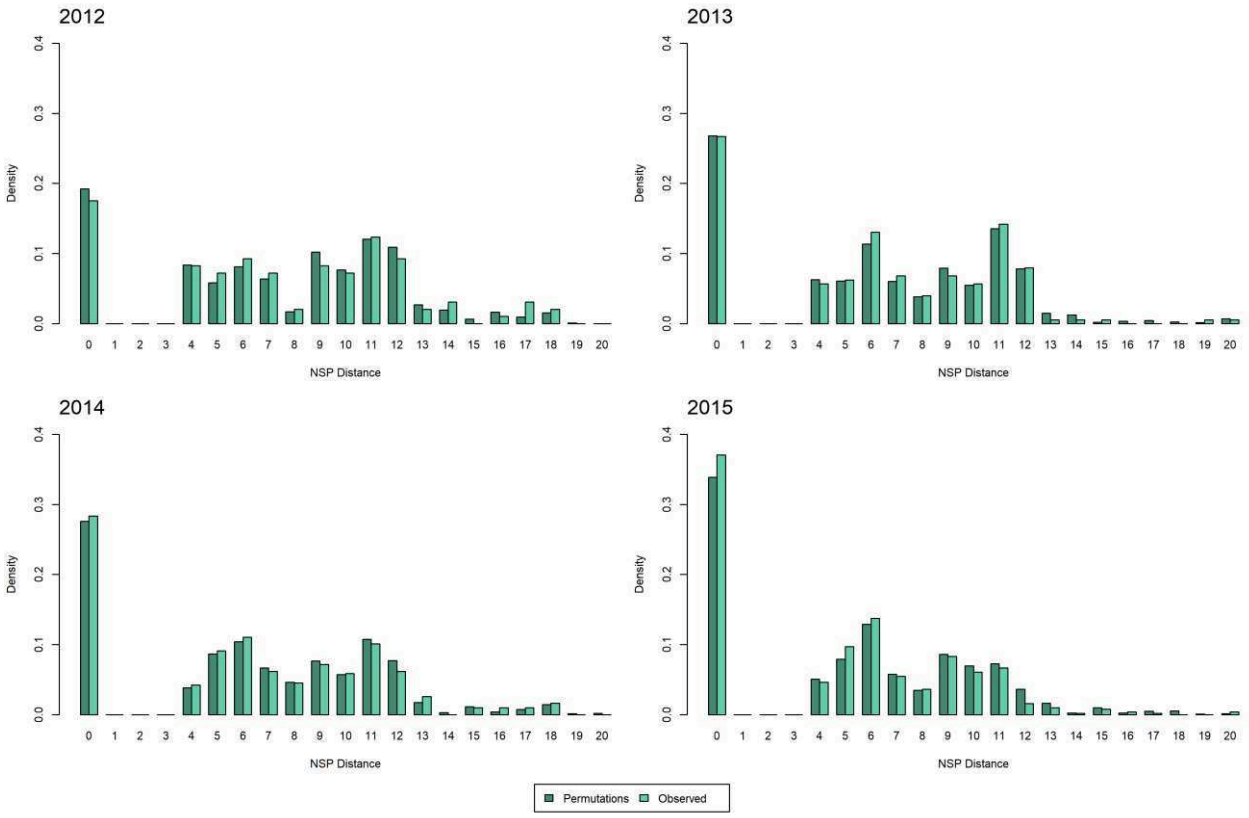
* Wilcoxon paired rank sum test p-value comparing of NSP distances between observed and permuted data

When these results were stratified by year, a potential infector was identified over the department network for 83% (in 2012) to 87% (in 2015) of non-imported episodes. Observed weighted NSP distances were significantly shorter than the NSP distances of the permutations of the data for 2013, 2014 and 2015 episodes (Wilcoxon paired rank sum test, p-value = 0.004, 6.90×10^{-6}, 1.88×10^{-23} respectively). Conversely, in 2012, weighted NSP distances did not differ significantly (Wilcoxon paired rank sum test, p-value=0.28). Similar results were obtained using the distance definition suggested by Donker et al. (Supplement 2).

The percent of multiple spreading events increased significantly over the four years (Supplement 3). From 27.5% in 2012 to 42.5% in 2015 of the identified potential infectors were linked to multiple spreading events. The size of the multiple spreading events also increased: two potential infectors were each linked to four incident episodes in 2012, three infectors to five incident episodes each in 2013, one infector to seven incident episodes in 2014, and one infector to nine incident episodes in 2015 (Supplement 3). However, there was no evidence of an association between the number of cases per potential infector episode and the size of the multiple spreading events (Kendall's rank correlation tau with averaged ties, p-value= 0.37, 0.25, 0.08, 0.06, and 0.34 for 2012 to 2015, 2012, 2013, 2014, and 2015 data respectively). The largest of these events originated in the Paris department and other departments with large metropoles such as Marseille, Lyon, and Toulon (Supplement 3).

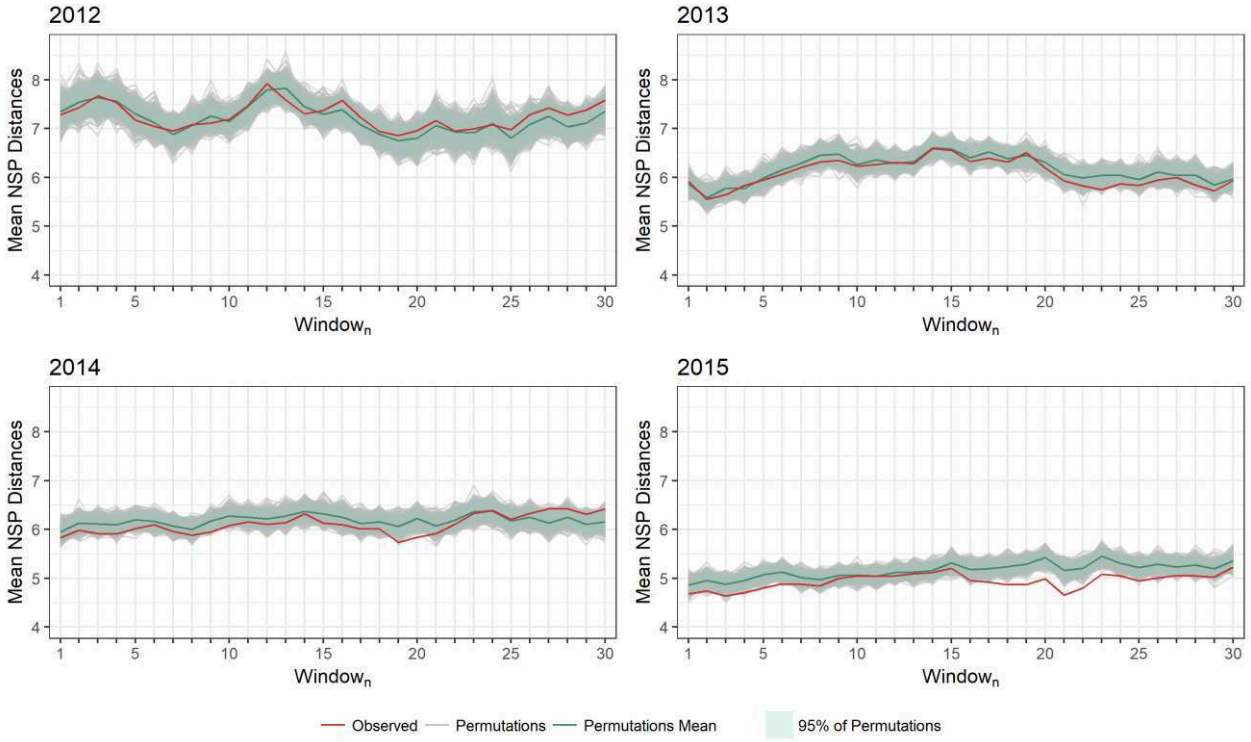
The proportion of potential infectors identified within the same department as the incident case did not differ significantly between the observed and permutated data when all years were combined (Supplement 4). However, by year, the proportion of episodes sharing the same department as their potential infector increased significantly from 18% in 2012 to 37% in 2015 (X^2 test for trend in proportions, p-value = 2.6×10^{-5}) (Figure 2). For NSPs identified between episodes of different departments, the majority were within close geographic distance of one another and the average distance between them reduced over time (from 250 km in 2012 to 170km in 2015) (Supplement 5).

Figure 2. Distribution of network-supported path distances for observed data and permutations by year. A total of 97, 176, 307, and 496 NSP distances are shown for 2012, 2013, 2014, and 2015 respectively. The density distribution is shown for observed data NSP distances (light green) and the mean NSP distances of the 500 permutations (dark green). Distances of zero correspond to paths that occur in the same department. Distances of one correspond to a range of distances between zero and one; the same applies to distances two through 20.



A sensitivity analysis was conducted in order to compare the baseline time window for selecting candidate transmitters to a sliding one-week window starting from 1 to 30 days ($W_{[1,8]}$ to $W_{[30,37]}$) prior to the incident episode starting date. The results for the combined years can be found in Supplement 4. The Wilcoxon paired-rank sum test p-values for all time windows shown in Figure 3 are given in Supplement 6. When comparing all time windows for the combined years, the most significant difference was observed for $W_{[20,27]}$ and the baseline (Supplement 6). In 2012, only four windows showed a significant difference between the data and permutations. Conversely, the test showed a statistically significant difference between the mean NSP distances between incident CPE episodes and their potential infector for all windows in 2015. In 2015, the windows with the lowest mean NSP distance corresponded to $W_{[3,10]}$ and the baseline window $W_{[21,28]}$, while for 2014 it corresponded to the $W_{[1,8]}$ and $W_{[19,26]}$ window. The most significant differences were observed for $W_{[23,30]}$ in 2013, $W_{[20,27]}$ in 2014, and $W_{[21,28]}$ in 2015. However, for all time windows identifying potential infectors 20 to 30 days prior the incident episode date , the results were similar with significantly shorter distances for 2013-2015 but not for 2012 (Figure 3, Supplement 6).

Figure 3. Sensitivity analysis on the impact of the time window chosen to look for candidate transmitters: mean NSP distances between incident episodes and their closest potential infectors obtained for sliding 1-week time windows, 2012-2015. For each year, the mean NSP distance is plotted as a function of the first day of the 1-week time window, $Window_n$, for observed data (in red) and permuted data (in black). For the permutations, 95% confidence bands are also provided.



Discussion

We were able to adapt a previously proposed statistical methodology to quantify the impact of the healthcare network of patient transfers to CPE transmission over a four-year period. Our study suggests that there is an association between the department network of hospital patient transfers and the spread of CPE episodes over time.

Due to a lack of information regarding hospitals in which the CPE episodes were reported, we relied solely on the department of the episodes. We assumed that transfer rates between hospitals of the same department were homogenous. Although this drawback may have reduced the power of the study, we were still able to show statistically significant differences between the NSP distances in the observed CPE episode data and the permutations, due to sufficient heterogeneity in the transfer rates between departments.

Since CPE hospital outbreak duration has been estimated to last approximately three weeks [17], we assumed that potential infector episodes may lead to multiple secondary episodes in other hospitals. Therefore, to reduce the risk of linking incident episodes to other secondary episodes from potentially the same infector episode, we deemed it appropriate to select a baseline window ranging from 21 to 28 days prior to the date of the incident episode for candidate transmitter selection. In addition, the baseline window was supported by the sensitivity analysis results where we observed that, among all windows, it provided the lowest mean distance of NSPs as well as the largest difference with the mean NSP distance from the permutations in 2015. However, it is important to note that other windows around the baseline also gave significantly shorter distances. Therefore, we estimate a delay in notification of CPE episodes between hospitals due to patient transfer could occur within 20 to 30 days after initial notification in an index hospital.

Our results suggest that the dynamics of CPE transmission in France have changed over time. Between 2012 and 2015, we were able to show evidence using NSP distances to evaluate the changing dynamics of CPE spread; evidence that supports potential CPE spatial spread through the carriage of CPE by transferred patients between French departments. CPE episodes from 2012 were not supported by the transfer network and can be explained by other sources of transmission such as previous hospitalizations, stays in a foreign country, or direct cross-border transfers. The proportion of imported CPE episodes decreased over time as

the number of non-imported episodes rose; suggesting a transition in 2013 from an epidemic sustained by importation to local transmission events sustaining the epidemic. Mounting evidence for local spread through transfers emerged in 2014 followed by the strongest evidence for transfer network-supported CPE transmission in 2015. These results suggest that between 2013 and 2014 there was a growing contribution of regional and inter-regional transfers in the spread of CPE in France which is in concordance with reports by the ECDC.[14]

Even though a large percentage of non-imported cases were linked to a potential infector, the reported 87% in 2015 for example corresponded to the maximum number of incident episodes with a potential infector episode, given the data. Since the number and frequency of OXA-48 episodes has increased over time, the chance of an incident episode having a potential infector was high as well. Nonetheless, we found evidence to support that the observed NSP were significantly shorter than what we expect by chance. Both the number of linked episodes occurring in the same department increased over time and the proportion of linked episodes occurring in different departments occurred within shorter geographic proximity over time as well.

While we do not claim that transfers are the sole explanation for the augmentation in observed CPE episodes between 2013 and 2015, our work suggests that they have played an increasingly significant role over time. Episodes from international importation could also have contributed to almost half of the spread of CPE in the country. These results are consistent with the outbreak descriptions we observe in the literature in which both imported and non-imported cases have led to secondary cases of CPE in different hospitals. In addition, the heterogeneity in infection control policies across different types of healthcare facilities in France and limited implementation of specific strategies to control CPE may have led to poor control of CPE and in consequence, dissemination over time.[22]

Due to no observed association between the number of cases per potential infector episode and the number of secondary episodes, we were not able to show evidence to support poor control of hospital CPE outbreaks once health authorities identified and reported a chain of transmission among cases. On one hand this may suggest that control measures have prevented large hospital outbreaks from causing multi-department outbreaks during the 2014-2015 period; on the other hand, most reports are single-case episodes which might suggest a failure of surveillance authorities in identifying single-cases as part of the same chain of transmission of other reported episodes. In addition, potential CPE cases are likely to be occurring in the community which can lead to non-identification and poor control of any potential cross-infections. Guidelines for screening and controlling CPE should include those epidemiological changes and be revised accordingly.

Our results suggest that the number of spreading events involving multiple episodes has increased over time. For example, an imported OXA-48 episode in Paris was linked to nine other episodes in nine different departments in France in 2015. Paris is a large hub for not only CPE episodes linked to importation but also for patient transfer; this underlines the importance for health authorities to improve control efforts in large and highly connected metropoles.

In conclusion, our study has demonstrated that a methodology of identifying potential infectors through NSPs at the patient contact level can also be applied to a national patient transfer network level. Our results suggest that since 2013, patient transfers in France have increasingly contributed to the epidemiological transition of CPE dynamics from regional to inter-regional spread sustained by an increasing number of local spreading events. Systematic screening of at-risk patients, such as hospital contacts of patients transferred from hospitals with previous or current patients infected with CPEs is crucial in identifying carriers of CPE to contain intra-hospital transmission. These efforts rely on regional coordination of hospital control measures targeting patient transfers especially that of university hospital centres that play a large role in connecting patients.[15]

References

1. Falagas ME, Lourida P, Poulikakos P, Rafailidis P, I. , Tansarli GS. Antibiotic Treatment of Infections Due to Carbapenem-Resistant Enterobacteriaceae: Systematic Evaluation of the Available Evidence. 2014. doi: 10.1128/AAC.01222-13.
2. Morrill HJ, Pogue JM, Kaye KS, LaPlante KL. Treatment Options for Carbapenem-Resistant Enterobacteriaceae Infections. Open Forum Infect Dis. 22015.
3. Albiger B, Glasner C, Struelens MJ, Grundmann H, Monnet DL. Carbapenemase-producing Enterobacteriaceae in Europe: assessment by national experts from 38 countries, May 2015. Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin. 2015;20(45). Epub 12/18. doi: 10.2807/1560-7917.es.2015.20.45.30062. PubMed PMID: 26675038.
4. Kassis-Chikhani N, Decre D, Gautier V, Burghoffer B, Saliba F, Mathieu D, et al. First outbreak of multidrug-resistant Klebsiella pneumoniae carrying bla(VIM-1) and bla(SHV-5) in a French university hospital. Journal of Antimicrobial Chemotherapy. 2006;57(1):142-5. doi: 10.1093/jac/dki389. PubMed PMID: WOS:000234436400022.
5. Kassis-Chikhani N, Decre D, Ichai P, Sengelin C, Geneste D, Mihaila L, et al. Outbreak of Klebsiella pneumoniae producing KPC-2 and SHV-12 in a French hospital. Journal of Antimicrobial Chemotherapy. 2010;65(7):1539-40. doi: 10.1093/jac/dkq132. PubMed PMID: WOS:000279926500039.
6. ECDC. Risk assessment on the spread of carbapenemase-producing Enterobacteriaceae (CPE) through patient transfer between healthcare facilities, with special emphasis on cross-border transfer. Stockholm: 2011.
7. Simmering JE, Polgreen LA, Campbell DR, Cavanaugh JE, Polgreen PM. Hospital Transfer Network Structure as a Risk Factor for Clostridium difficile Infection. Infection Control and Hospital Epidemiology. 2015;36(9):1031-7. doi: 10.1017/ice.2015.130. PubMed PMID: WOS:000360917200006.
8. Fernández-Gracia J, Onnela J-P, Barnett ML, Eguíluz VM, Christakis NA. Influence of a patient transfer network of US inpatient facilities on the incidence of nosocomial infections. Scientific Reports. 2017;7(1):2930. doi: 10.1038/s41598-017-02245-7.
9. Ray MJ, Lin MY, Weinstein RA, Trick WE. Spread of Carbapenem-Resistant Enterobacteriaceae Among Illinois Healthcare Facilities: The Role of Patient Sharing. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2016;63(7):889-93. Epub 08/04. doi: 10.1093/cid/ciw461

10.1093/cid/ciw461. Epub 2016 Aug 2. PubMed PMID: 27486116.
10. Lee BY, Bartsch SM, Wong KF, McKinnell JA, Slayton RB, Miller LG, et al. The potential trajectory of carbapenem-resistant *Enterobacteriaceae*, an emerging threat to health-care facilities, and the impact of the centers for disease control and prevention toolkit. American Journal of Epidemiology. 2016;183(5):471-9. doi: 10.1093/aje/kwv299. PubMed PMID: WOS:000371688800023.
11. Ciccolini M, Donker T, Kock R, Mielke M, Hendrix R, Jurke A, et al. Infection prevention in a connected world: the case for a regional approach. International Journal of Medical Microbiology. 2013;303(6-7):380-7. doi: 10.1016/j.ijmm.2013.02.003. PubMed PMID: WOS:000323803700012.
12. Poirel L, Fortineau N, Nordmann P. International Transfer of NDM-1-Producing Klebsiella pneumoniae from Iraq to France. Antimicrobial Agents and Chemotherapy. 2011;55(4):1821-2. doi: 10.1128/aac.01761-10. PubMed PMID: WOS:000288594600073.
13. Vaux S, Carbonne A, Thiolet JM, Jarlier V, Coignard B. Emergence of carbapenemase-producing Enterobacteriaceae in France, 2004 to 2011. Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin. 2011;16(22). Epub 06/15. PubMed PMID: 21663708.
14. ECDC. Rapid risk assessment: Carbapenem-resistant Enterobacteriaceae. Stockholm: 2016 8 April 2016. Report No.
15. Nekkab N, Astagneau P, Temime L, Crepey P. Spread of hospital-acquired infections: A comparison of healthcare networks. PLoS Comput Biol. 2017;13(8):e1005666. Epub 08/25. doi: 10.1371/journal.pcbi.1005666. eCollection 2017 Aug. PubMed PMID: 28837555.

16. Obadia T, Silhol R, Opatowski L, Temime L, Legrand J, Thiebaut ACM, et al. Detailed contact data and the dissemination of *Staphylococcus aureus* in hospitals. Plos Computational Biology. 2015;11(3):16. doi: 10.1371/journal.pcbi.1004170. PubMed PMID: WOS:000352195700048.
17. Fournier S, Monteil C, Lepainteur M, Richard C, Brun-Buisson C, Jarlier V, et al. Long-term control of carbapenemase-producing Enterobacteriaceae at the scale of a large French multihospital institution: a nine-year experience, France, 2004 to 2012. 2014. doi: doi:10.2807/1560-7917.ES2014.19.19.20802.
18. Carbonne A, Thiolet JM, Fournier S, Fortineau N, Kassis-Chikhani N, Boytchev I, et al. Control of a multi-hospital outbreak of KPC-producing Klebsiella pneumoniae type 2 in France, September to October 2009. Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin. 2010;15(48). Epub 12/15. PubMed PMID: 21144448.
19. Delory T, Seringe E, Antoniotti G, Novakova I, Goulenok C, Paysant I, et al. Prolonged delay for controlling KPC-2-producing Klebsiella pneumoniae outbreak: the role of clinical management. Am J Infect Control. 2015;43(10):1070-5. Epub 07/16. doi: 10.1016/j.ajic.2015.05.021

10.1016/j.ajic.2015.05.021. Epub 2015 Jul 11. PubMed PMID: 26174583.
20. Csardi G, Nepusz T. The igraph software package for complex network research. InterJournal, Complex Systems 1695. 2006.
21. Donker T, Smieszek T, Henderson KL, Johnson AP, Walker AS, Robotham JV. Measuring distance through dense weighted networks: The case of hospital-associated pathogens. PLOS Computational Biology. 2017;13(8):e1005622. doi: 10.1371/journal.pcbi.1005622.
22. Lepelletier D, Berthelot P, Lucet JC, Fournier S, Jarlier V, Grandbastien B, et al. French recommendations for the prevention of 'emerging extensively drug-resistant bacteria' (eXDR) cross-transmission. Journal of Hospital Infection. 2015;90(3):186-95. doi: 10.1016/j.jhin.2015.04.002. PubMed PMID: WOS:000355836400002.

Acknowledgements

The authors would like to thank Bruno Coignard, MD, and Valérie Pontiès, PharmD, from Santé Publique France for providing the CPE data, insight discussion, and expertise that greatly assisted the research. The work was supported in part by the PRINCEPS program from the Sorbonne-Paris Cité University.

Supplement 1: Example of potential infector selection in the observed data and permutations

This supplementary material is hosted by *Eurosurveillance* as supporting information alongside the article "Assessing the role of inter-hospital patient transfer in the spread of carbapenemase-producing *Enterobacteriaceae*: the case of France between 2012 and 2015" by N Nekkab et al. on behalf of the authors who remain responsible for the accuracy and appropriateness of the content. The same standards for ethics, copyright, attributions and permissions as for the article apply. *Eurosurveillance* is not responsible for the maintenance of any links or email addresses provided therein.

		Incident Episode *i*	**Window$_{[t-n,t-m]}$**				
	Episode	E_i	**Non-Candidate Episodes**			**Candidate Transmitters *j***	
			E_a	E_b	E_c	E_d	E_e
	Mechanism of Resistance	OXA	VIM	NDM	KPC	OXA	OXA
Observed	Department	75	13	69	54	92	33
	d_{ij}	-	-	-	-	**1**	2
Permutation$_1$	Department	75	54	92	33	69	13
	d_{ij}	-	-	-	-	4	**3**
...							
Permutation$_{500}$	Department	75	69	54	92	33	54
	d_{ij}	-	-	-	-	**2**	3

For a given incident episode E_i during a given one-week sliding window, $W_{[t-n,t-m]}$, in which t corresponds to the E_i date, n corresponds to the first day of the sliding window preceding the date of E_i and m corresponds to n + 7 days preceding E_i, five CPE episodes $E_{a\rightarrow e}$ occurred in five different departments. Among episodes that shared the same mechanism of resistance, the candidate transmitters j, E_d and E_e, the shortest network path distance in the observed data was between $E_{i\rightarrow d}$ ($d_{id} = 1 < d_{ie} = 2$). Therefore, in the observed data, the most likely potential infector of E_i was identified as E_d with a network-supported path (NSP) equal to 1. The departments of $E_{a\rightarrow e}$ were permutated through sampling without replacement five-hundred times and the potential infectors were identified. In permutation$_1$, E_e was identified as the potential infector given that $d_{id} = 4 > d_{ie} = 3$ and, in permutation$_{500}$ E_d was identified as the potential infector ($d_{id} = 2 < d_{ie} = 3$). The distribution of the observed NSP distances was compared to the mean of the NSP distances of the 500 permutations for each non-imported episode E_i for significance using a Wilcoxon paired rank sum test.

Supplement 2: alternative definition of edge weights, a sensitivity analysis

Another distance definition was recently proposed by Donker et al.[1] to measure the distances in a healthcare network. The authors' measure of distances was based on Opsahl et al.'s[2] work in which they proposed a tuning parameter, alpha α, to adjust between a value of 0 corresponding to the same distance measure we expect if we do not consider only the number of "jumps" between hospitals and 1 corresponding to Dijkstra's algorithm in which we measure distance as the number of transfers w_{ij}. Donker et al. simulate the spread of antimicrobial resistant bacteria using mathematical modeling with varying measures of alpha to obtain an optimal α = 0.25 in which the time to infection was greatest. In order to compare our results, we re-tested our baseline and sensitivity analysis (Table S2, Figure S2) using a network of transformed edges weights in which the new distances d_{ij} was equal to:

$$d_{ij} = \frac{1}{w_{ij}^{\alpha}} \quad \text{where } \alpha = 0.25 \qquad (1)$$

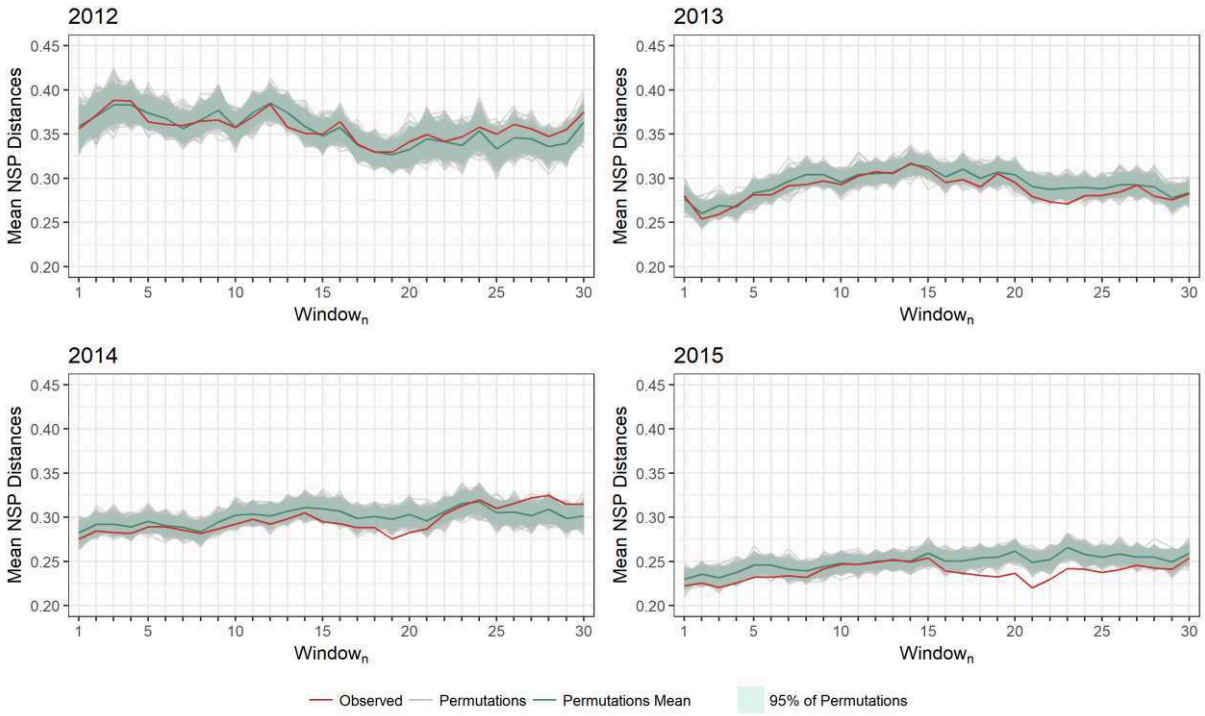
Table S2. Network-supported path distances of 2273 carbapenemase-producing *Enterobacteriaceae* episodes in France from 2012 to 2015 using an alternative measure of distance (1).

Year	2012	2013	2014	2015
Total episodes	242	403	672	956
Imported episodes	125	200	311	386
Non-imported episodes (%)	117 (48%)	203 (50%)	361 (54%)	570 (60%)
Non-imported episodes with potential infector (%)	97 (83%)	176 (87%)	307 (85%)	496 (87%)
NSP distance of observed data mean [95% CI]	0.35 [0.29-0.41]	0.28 [0.24-0.32]	0.29 [0.25-0.32]	0.22 [0.20-0.24]
NSP distance of permutations mean [95% CI]	0.34 [0.29-0.40]	0.29 [0.25-0.33]	0.30 [0.27-0.33]	0.25 [0.23-0.27]
P-value*	0.23	0.002	5.06×10^{-7}	5.07×10^{-23}

NSP: network-supported path

* Wilcoxon paired rank sum test p-value comparing of NSP distances between observed and permuted data

Figure S2. Sensitivity analysis using an alternative measure of distance (1): mean NSP distances between incident episodes and their closest potential infectors obtained for sliding 1-week time windows, 2012-2015. For each year, the mean NSP distance is plotted as a function of the first day of the 1-week time window $Window_n$, for observed data (in red) and permuted data (in black); for the observed data, 95% confidence bands are also provided.



There was a statistically significant difference at the baseline window $W_{[21,28]}$ for 2013, 2014, and 2015 using the alternative definition (Wilcoxon paired rank sum test, p-value = 0.002, 5.06×10^{-7}, 5.07×10^{-23} respectively) (Table S2). We observe very similar results in the original distance measure and the alternative definition at baseline and across the time windows (Figure S2). Applying the alternative method to measure distances in the healthcare network also showed the shortest mean NSP distances $W_{[20,27]}$ in 2014 and $W_{[21,28]}$ in 2015. Whether the distances are measured as the negative log of the annual transfer rate or the inverse of the total annual transfers to the power of an optimal scaling parameter, at the baseline window, observed NSP distances were shorter than what would be expected by chance in the random permutations; therefore, high rates of patient transfers may be linked to the CPE epidemic in France in recent years.

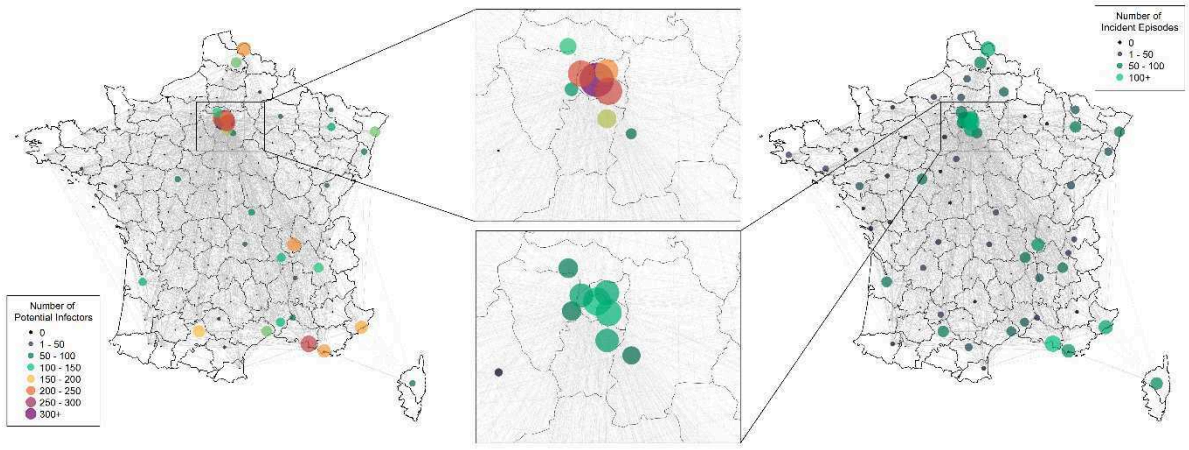
References

1. Donker T, Smieszek T, Henderson KL, Johnson AP, Walker AS, Robotham JV. Measuring distance through dense weighted networks: The case of hospital-associated pathogens. PLOS Computational Biology. 2017;13(8):e1005622. doi: 10.1371/journal.pcbi.1005622.
2. Opsahl T, Agneessens F, Skvoretz J. Node centrality in weighted networks: Generalizing degree and shortest paths. Social Networks. 2010;32(3):245-51. doi: https://doi.org/10.1016/j.socnet.2010.03.006.

Supplement 3: multiple spreading events

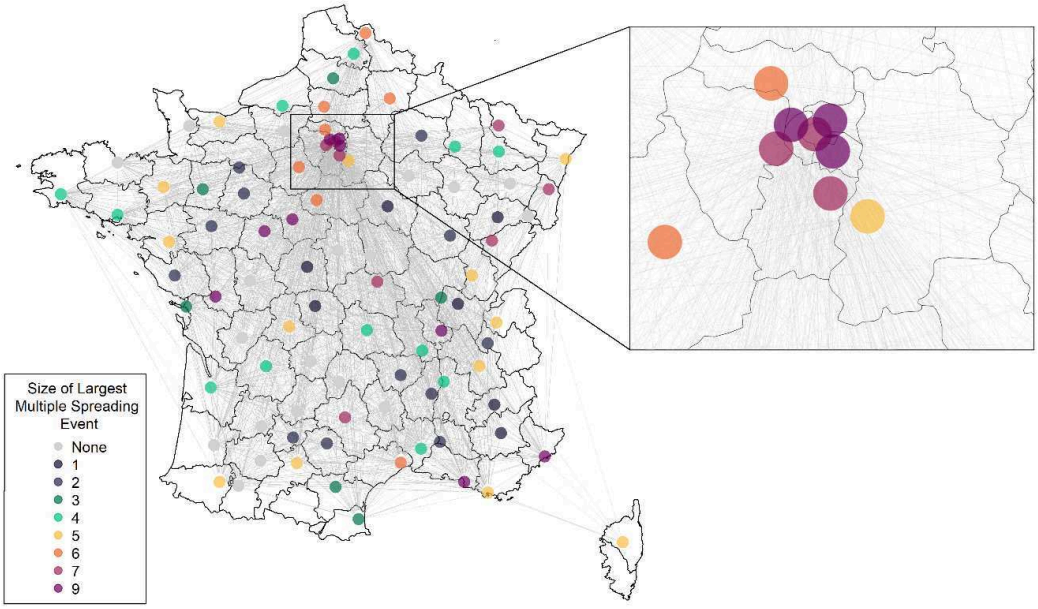
This supplementary material is hosted by *Eurosurveillance* as supporting information alongside the article "Assessing the role of inter-hospital patient transfer in the spread of carbapenemase-producing *Enterobacteriaceae*: the case of France between 2012 and 2015" by N Nekkab et al. on behalf of the authors who remain responsible for the accuracy and appropriateness of the content. The same standards for ethics, copyright, attributions and permissions as for the article apply. *Eurosurveillance* is not responsible for the maintenance of any links or email addresses provided therein.

Figure S3.1. Spatial distribution of linked episodes. Potential infector episodes are shown on the left and secondary episodes on the right per department from 2012 to 2015.



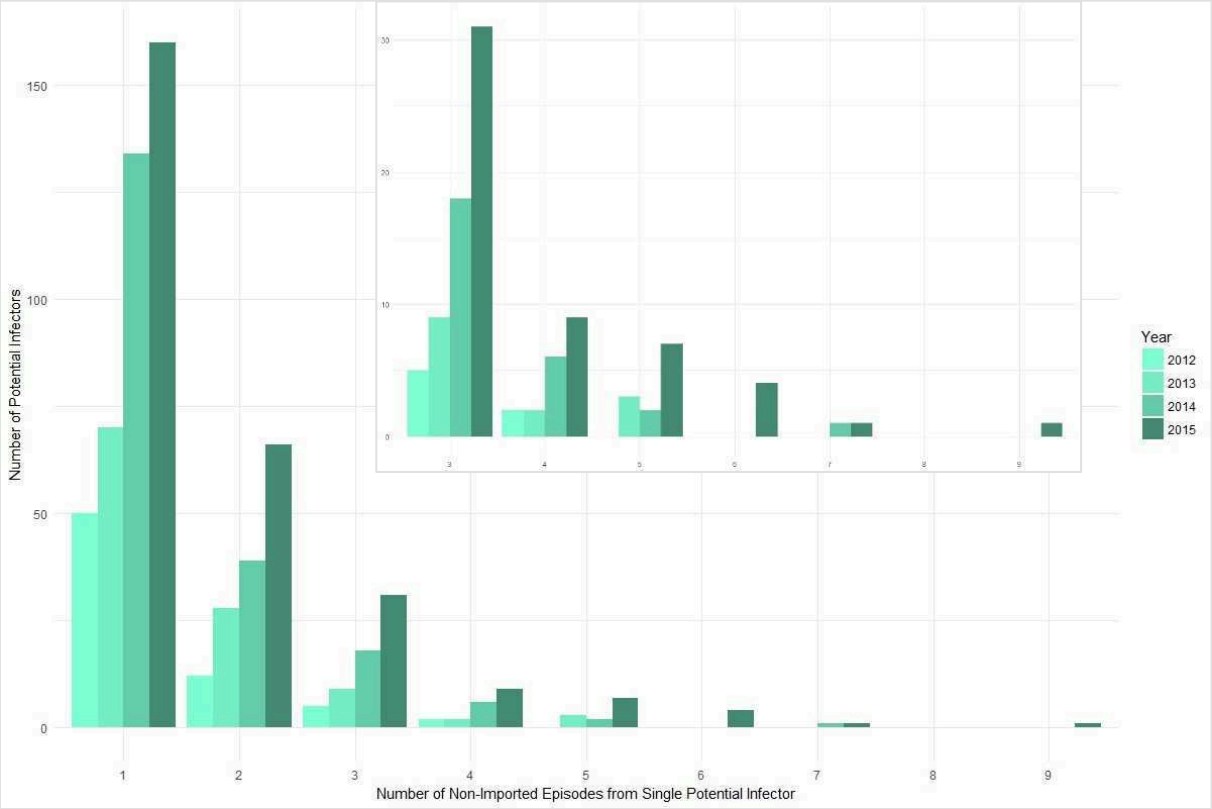
The Ile-de-France regions where Paris is located, is both the hub for potential infectors and secondary episodes (Figure S3.1).

Figure S3.2. Spatial distribution of departments with secondary episodes were associated with multiple spreading events. The size of the events are shown for the 2012 to 2015 period. Only the largest multiple spreading events are shown for each department.



Most of the largest multiple spreading events originated in the Paris department (for all events of size 9 and 7) and spread to neighboring Ile-de-France departments and to other hubs departments including large towns such as Lyon(69), Marseille(13), and Nice(06) and neighboring regions (Figure S3.2). Notably, the Var (83), the Rhône (69), the other Ile-de-France departments (91, 92, 93, and 94), the Nord (59), and the Loire (42) departments were also sources of multiple spreading events.

Figure S3.3. Distribution of the number of secondary episodes per infector. Number of potential infectors by the number of their secondary episodes (from 1 to 9 in the main graph and from 3 to 9 in the smaller graph) during the 2012 to 2015 period.

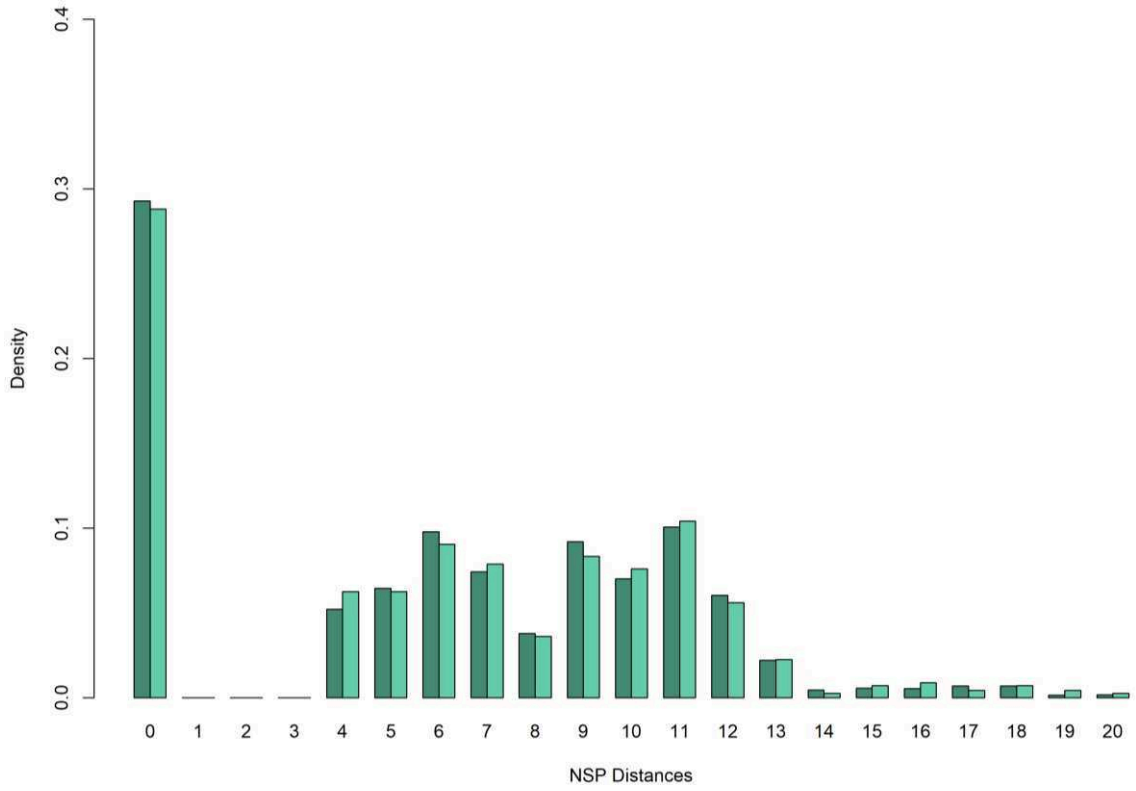


The percent of multiple spreading events significantly increased over time (X^2 test for trend in proportions, p-value = 0.02) from 27.5%, 37.5%, 33%, to 42.6% from 2012, 2013, 2014, and 2015 for potential infector episodes with two or more linked secondary episodes and also for potential infector episodes with three or more linked secondary episodes (10%, 12.5%, 13.5%, 19% respectively; X^2 test for trend in proportions, p-value = 0.03).

Supplement 4: results for the 2012 to 2015 period

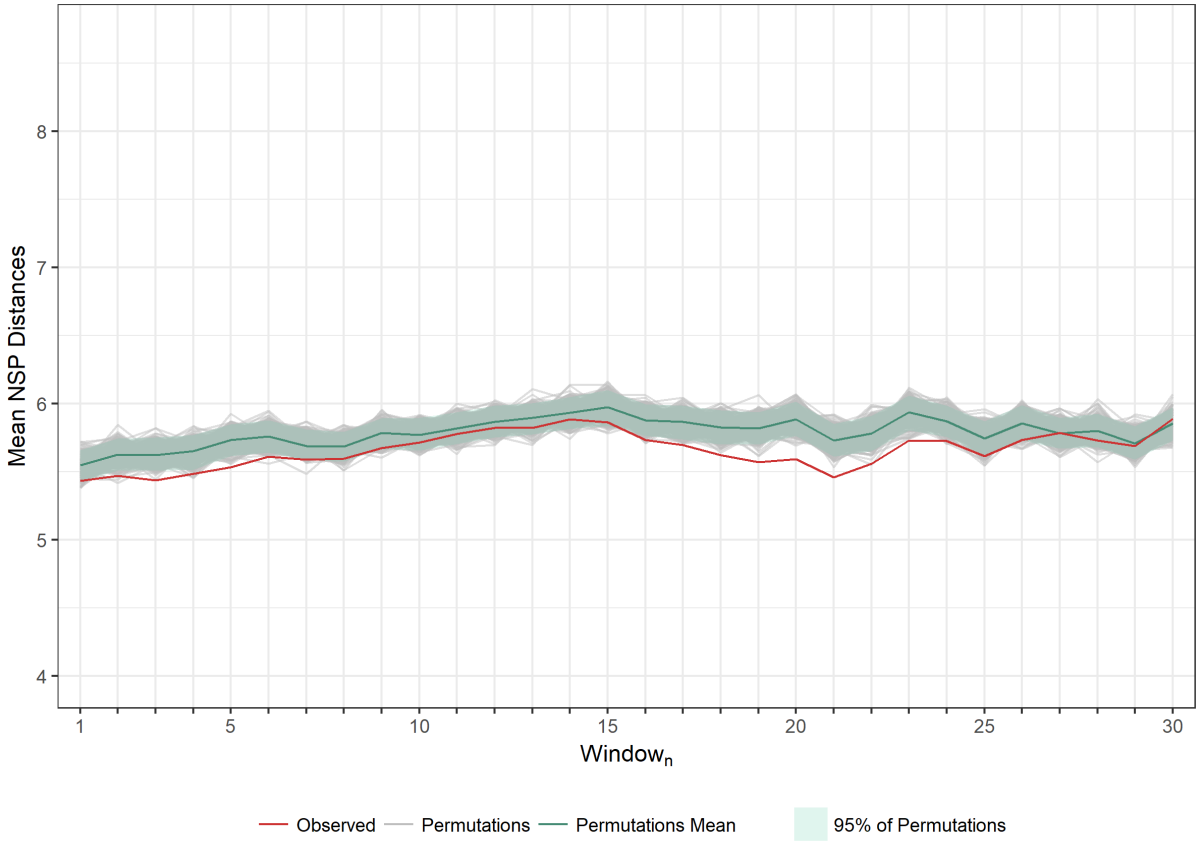
This supplementary material is hosted by *Eurosurveillance* as supporting information alongside the article "Assessing the role of inter-hospital patient transfer in the spread of carbapenemase-producing *Enterobacteriaceae*: the case of France between 2012 and 2015" by N Nekkab et al. on behalf of the authors who remain responsible for the accuracy and appropriateness of the content. The same standards for ethics, copyright, attributions and permissions as for the article apply. *Eurosurveillance* is not responsible for the maintenance of any links or email addresses provided therein.

Figure S4.1. Distribution of 1122 network-supported path distances for observed data and permutations for all episodes occurring between 2012 and 2015. The density distribution is shown for observed data NSP distances (light green) and the distances of permutations (dark green). Distances of zero correspond to paths that occur in the same department. Distances of one correspond to a range of distances between zero and one; the same applies to distances two through 20.



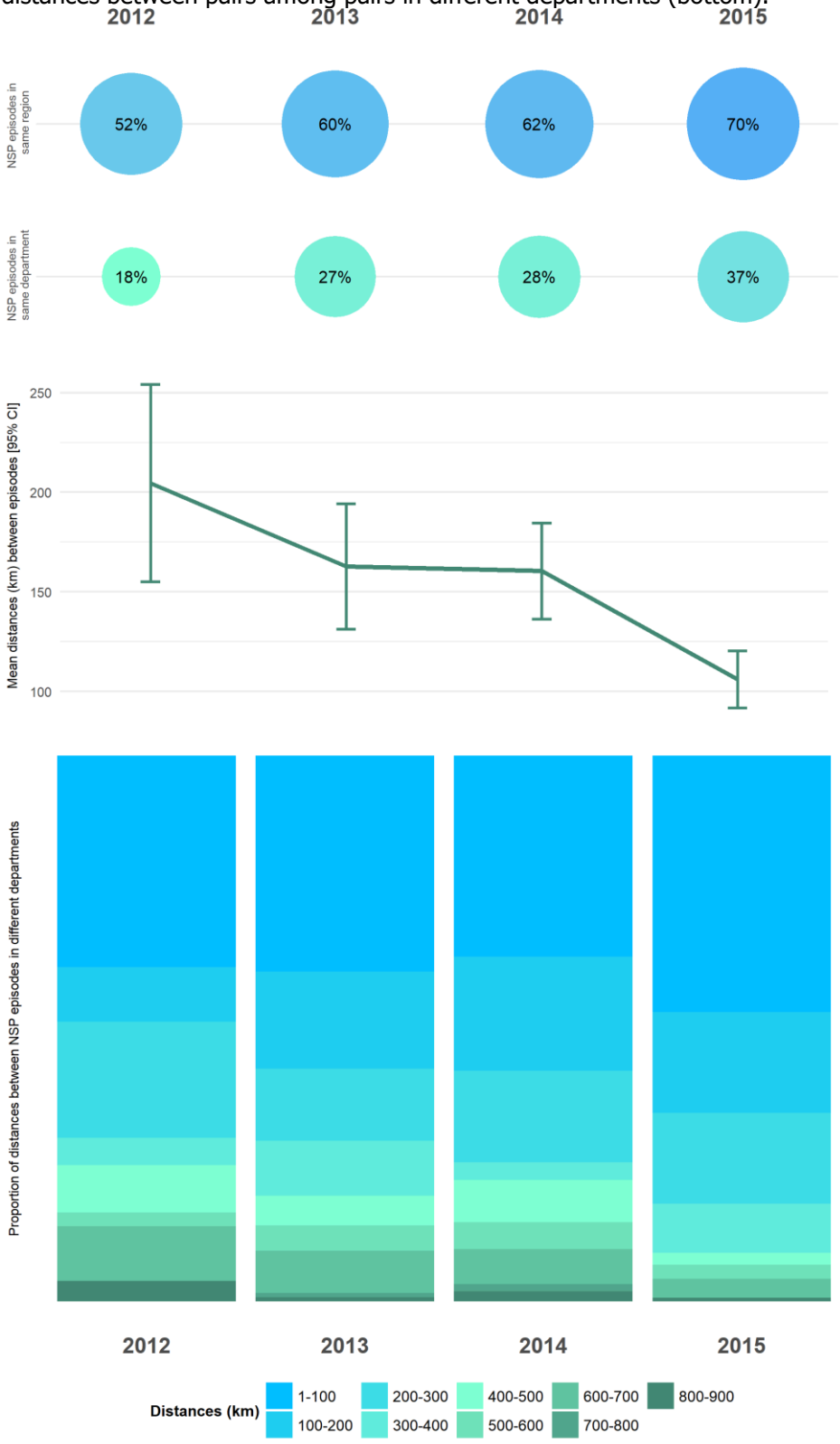
There was no statistically significant difference in the proportion of episodes occurring in the same department in the observed and permutated data (X^2 test for trend in proportions, p-value = 0.98).

Figure S4.2. Sensitivity analysis on the impact of the time window chosen to select candidate transmitters. The distribution of the mean NSP distances between incident episodes and their potential infectors obtained for sliding 1-week time windows for all episodes occurring between 2012 and 2015. The mean NSP distance is plotted as a function of the first day of the 1-week time window, for observed data (in red) and permuted data (in black). For the permutations, 95% confidence bands are also provided.



Supplement 5: geographic distances between potential infectors and their secondary cases

Figure S5. Distances to potential infectors. For observed data: the percent of potential infectors and NSP non-imported episodes (pairs) occurring the same department or region (top); the mean distance in kilometers (km) between pairs occurring in the same or different department (middle); the proportion of distances between pairs among pairs in different departments (bottom).



Supplement 6: p-values of sensitivity analysis of all time windows

Table S6. Sensitivity analysis. Mean NSP distances of observed data and 500 permutations: Wilcoxon paired rank sum test p-values for all time windows.

W_n	W_m	2012-2015			2012			2013			2014			2015		
		Mean NSP distances		P-value	Mean NSP distances		P-value	Mean NSP distances		P-value	Mean NSP distances		P-value	Mean NSP distances		P-values
		O	P		O	P		O	P		O	P		O	P	
1	8	5.43†	5.55	3.15E-18***	7.29	7.34	0.29	5.91	5.87	0.03*	5.83††	5.95	3.41E-07***	4.68	4.87	1.68E-12***
2	9	5.47	5.63	1.77E-20***	7.44	7.56	0.08	5.56†	5.59	3.41E-03**	5.99	6.13	5.92E-07***	4.74	4.96	3.24E-13***
3	10	5.44††	5.62	3.96E-27***	7.67	7.65	0.10	5.65††	5.77	1.07E-04***	5.92	6.12	2.00E-08***	4.64†	4.87	1.16E-16***
4	11	5.48	5.65	1.36E-25***	7.53	7.56	0.18	5.83	5.77	3.31E-03**	5.91	6.09	2.03E-07***	4.71	4.96	7.92E-18***
5	12	5.53	5.73	4.29E-27***	7.17††	7.30	0.05*	5.95	6.00	4.91E-04***	6.02	6.20	7.25E-07***	4.80	5.07	4.37E-19***
6	13	5.61	5.76	2.92E-23***	7.04†	7.12	0.05*	6.07	6.16	1.29E-04***	6.10	6.17	3.67E-05***	4.89	5.13	2.81E-15***
7	14	5.59	5.69	3.24E-21***	6.95	6.88	0.48	6.20	6.29	5.13E-05***	5.96	6.07	2.00E-06***	4.88	5.01	2.16E-13***
8	15	5.60	5.69	1.19E-23***	7.08	7.08	0.17	6.32	6.45	4.47E-05***	5.88	6.01	1.45E-07***	4.84	4.97	5.20E-15***
9	16	5.68	5.78	2.21E-24***	7.11	7.26	0.08	6.35	6.47	9.44E-05***	5.95	6.17	1.87E-09***	5.00	5.06	5.87E-13***
10	17	5.71	5.77	4.25E-18***	7.19	7.14	0.34	6.23	6.27	7.99E-04***	6.08	6.28	7.43E-08***	5.05	5.07	4.29E-09***
11	18	5.78	5.82	3.03E-17***	7.47	7.45	0.48	6.27	6.36	1.17E-04***	6.16	6.26	3.13E-07***	5.04	5.04	1.24E-08***
12	19	5.82	5.87	3.09E-15***	7.92	7.80	0.62	6.31	6.29	0.01*	6.11	6.22	7.97E-06***	5.05	5.12	1.09E-09***
13	20	5.82	5.90	4.56E-20***	7.58	7.82	0.01*	6.28	6.33	3.63E-04***	6.14	6.27	1.38E-06***	5.10	5.13	4.19E-11***
14	21	5.88	5.94	1.57E-18***	7.30	7.45	0.03*	6.59	6.61	2.16E-03**	6.32	6.37	6.35E-05***	5.12	5.17	2.52E-11***
15	22	5.86	5.98	3.93E-20***	7.37	7.31	0.30	6.56	6.58	8.99E-04***	6.13	6.32	9.82E-07***	5.21	5.32	1.30E-11***
16	23	5.73	5.88	6.72E-20***	7.58	7.39	0.47	6.32	6.41	2.56E-04***	6.10	6.25	2.56E-06***	4.95	5.18	1.22E-12***
17	24	5.70	5.87	1.30E-20***	7.23	7.08	0.54	6.39	6.52	1.40E-03**	6.01	6.12	1.71E-06***	4.92	5.20	3.71E-13***
18	25	5.62	5.83	4.03E-23***	6.95	6.88	0.55	6.31	6.38	9.79E-03**	6.01	6.16	7.95E-07***	4.87	5.24	1.56E-18***
19	26	5.57	5.82	1.65E-24***	6.86	6.76	0.61	6.50	6.45	0.15	5.73†	6.06	5.79E-08***	4.87	5.29	3.85E-20***
20	27	5.59	5.89	2.32E-30***	6.96	6.81	0.43	6.20	6.30	1.37E-03**	5.84	6.23	8.35E-09***	4.99	5.43	9.32E-23***
21	28	5.46	5.73	3.27E-28***	7.17	7.07	0.28	5.93	6.06	3.86E-03**	5.91	6.07	6.90E-06***	4.66††	5.17	1.88E-23***
22	29	5.56	5.78	5.64E-23***	6.96	6.93	0.10	5.83	5.99	1.72E-03**	6.11	6.19	1.92E-04***	4.81	5.21	1.07E-17***
23	30	5.73	5.94	6.56E-22***	7.00	6.93	0.33	5.75	6.04	1.46E-06***	6.33	6.36	8.40E-03**	5.09	5.45	4.05E-18***
24	31	5.72	5.87	8.71E-22***	7.08	7.11	0.07	5.87	6.04	2.05E-05***	6.39	6.38	1.82E-03**	5.05	5.31	8.65E-15***
25	32	5.61	5.74	2.92E-21***	6.98	6.79	0.33	5.83	5.95	4.48E-05***	6.21	6.18	2.23E-03**	4.95	5.22	1.70E-15***
26	33	5.73	5.86	2.17E-20***	7.29	7.08	0.33	5.95	6.11	5.82E-04***	6.33	6.25	8.32E-03**	5.01	5.29	1.03E-16***
27	34	5.79	5.78	4.34E-11***	7.42	7.26	0.24	6.00	6.04	8.45E-03**	6.42	6.13	0.53	5.06	5.24	1.17E-10***
28	35	5.73	5.80	1.71E-14***	7.28	7.04	0.75	5.85	6.05	1.39E-04***	6.43	6.25	0.11	5.05	5.28	3.08E-11***
29	36	5.69	5.71	1.23E-13***	7.38	7.12	0.80	5.72	5.84	2.56E-04***	6.31	6.11	0.03*	5.03	5.20	1.23E-11***
30	37	5.89	5.86	2.96E-09***	7.59	7.35	0.38	5.94	5.98	3.83E-03**	6.42	6.15	0.27	5.23	5.35	1.17E-09***

O: observed data; P: permutated data; * p-value < 0.05; ** p-value < 0.01; *** p-value < 0.001; † shortest significant observed NSP distance; †† second shortest significant observed NSP distance.

10.2 Additional analyses

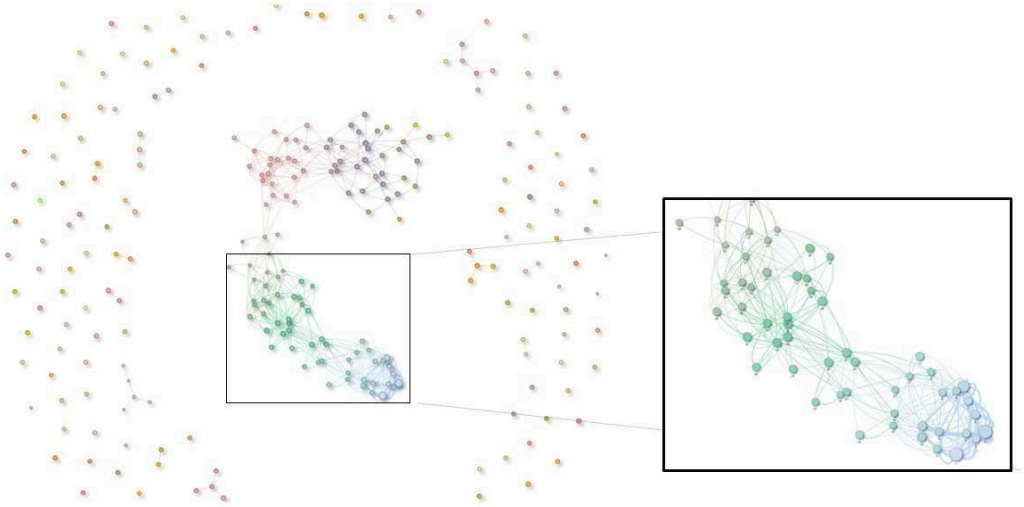
Based on our analysis of the contribution of patient transfers to CPE spread in France, it may be assumed that in 2015 at least, CPE do indeed spread over the French healthcare network to some extent. In a follow-up analysis, we hence attempted to reconstruct transmission chains of CPE episodes among French hospitals. This analysis was performed by adapting a Bayesian method initially developed to reconstruct outbreaks based on epidemiological and genomic data (Outbreaker2 R package, developed by T. Jombart, F. Campbell, R. Fitzjohn) (196, 197), without genomic data and for a large nationwide outbreak. The method was used to determine possible chains of transmission of CPE data based on knowledge of the episode dates, the weight of the links in the healthcare network, and the importation status of episodes. The study aimed to identify sources of CPE infection spread in the healthcare network and to determine what proportion of secondary infections could be explained from the network data.

The Outbreaker2 tool is an updated and more modular version of the earlier package Outbreaker, both of which are available for use on the R programming software.(198) The epidemiological likelihood of a given transmission tree is estimated using the distribution of a given generation time (the time interval between a primary and secondary infection) coupled with a model of sequence evolution defining the probability of the genetic changes along a chain of transmission.(198) The model allows the estimation of dates of infections, mutation rates, separate introductions of the pathogen, the presence of unobserved cases, and the transmission tree.

Five assumptions were made about transmission in the first analysis: 1) episodes were considered as “cases” (therefore the number of cases per episode was ignored) in the first analysis but not in the second 2) imported episodes were considered index cases and non-imported episodes were considered potential secondary cases, 3) an index case had to occur *n* number of days before a secondary case to be linked 4) an index case episode’s department and secondary case episode’s department had to be sufficiently linked in the healthcare network and 5) transmission could only occur between CPE episodes with the same mechanism of resistance since no genetic data was used. Since only the department of the CPE episodes was known and not the hospital, the department network of patient transfers was used for the analyses.

An example of reconstruction of transmission chains of 300 OXA-48 episodes occurring in 2015 is shown in Figure 27. Each link corresponds to two episodes having at least 5% probability of existing.

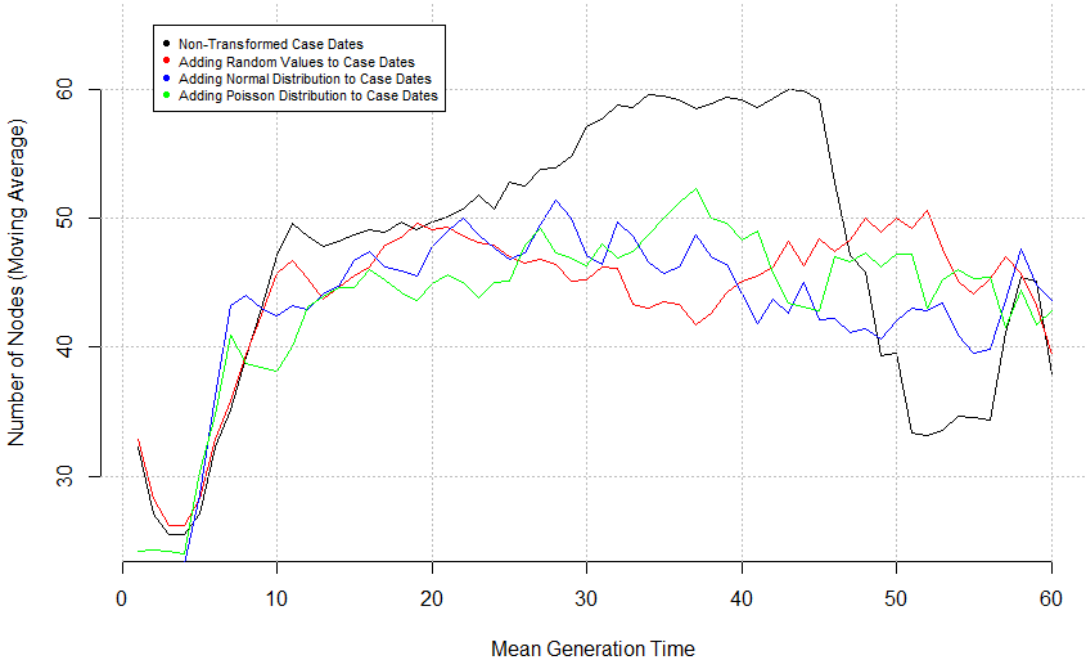
Figure 27. Reconstructed transmission chains of 300 OXA-48 episodes in 2015



A few observations may be made about some of these preliminary analyses: not all imported episodes were part of a chain of transmission, many were not linked to any episode, and many secondary cases had multiple potential sources of infection so uncertainty of the source was often high.

Since CPE episodes could also group many cases that could result in many other secondary cases, the second analysis assessed the entire dataset of each individual CPE case. Since multiple case episodes shared the same notification date but may not have occurred during that same date, several methods were developed to modify case dates: keeping all case dates as the original episode date, adding random values to case dates assuming the first reported date was the first reported case, adding a normal distribution of values to case dates, and adding a Poisson distribution to case dates. The total number of cases linked according to different mean generation time between cases is shown in Figure 28.

Figure 28. Transmission chain component size by mean generation time



The largest component size for transformed case dates corresponded to a generation time of 37 days using a Poisson distribution (Figure 28). This result was similar to findings in the previous study assessing the role of patient transfers on CPE spread where the sensitivity analysis supported a time window between potential infectors and incidence episodes between 20 to 30 days. However, depending on the transformation used, other time windows such as 10 to 20 days or 40 to 50 days could also have been just a pertinent. Component size or the number of connections may not be the most appropriate criteria for selection of the most appropriate generation time; therefore, other epidemiologically pertinent criteria should be considered.

This study showed the possibility to reconstruct large outbreaks without genomic data. Healthcare networks used to construct chains of transmission of CPE were able to explain a percentage of secondary CPE infections and to identify hospital hotspots of CPE spread; however, more work is needed to refine the understanding of the links between episodes.

Part Five: Discussion of thesis work and perspective

Chapter 11. Synthesis of results

After illustrating the integration of network data in mathematical modelling of pathogen spread in healthcare settings, this thesis has consisted of developing healthcare networks of patient transfers in France in order to better understand the transfer patterns of different patient populations, the role of the healthcare network structure in potential HAI spread, and the contribution of patient transfers, both international and local, on the CPE epidemic.

The first work of the thesis entailed conducting a systematic review of mathematical models in healthcare settings using real data on networks within and between institutions. This review presented an overview of the specific methods related to integration of network data in the various modelling studies identified and how they may improve our understanding and predictive capacity of HAI spread in healthcare settings. Models of HAI spread that incorporate either inter-individual contact or inter-institutional transfer network data were found to have become more frequent over time, and to have brought new insights into more effective HAI prevention and infection control strategies. However, to this date, they have also been limited to a few settings and a few pathogens. Further innovations in data collection and validation of parameter estimates of these models appear necessary for the improvement of our understanding of HAI spread dynamic.

The healthcare networks of France were reconstructed using the French medico-administrative database, the PMSI, after a literature review was conducted to assess its validity for detection of HAIs. At the heart of the thesis, the reconstruction and analysis of the French healthcare networks provided a first detailed description of the patient transfer patterns at the French national level, based on extensive social network analyses. The French healthcare networks were found to be characterised by heterogeneous patient flow, to demonstrate a scale-free and small-world nature, and to have a two-tier community clustering structure. A comparison showed that HAI-specific and suspected-HAI networks relied on the same underlying structure as that of the general patient network. The identified key hub healthcare centres, patient flow trajectories, and regional and local community clustering structure may help serve as a basis for novel infection control strategies.

Based on these reconstructed networks, a preliminary simulation study of pathogen spread was conducted in the early stages of the thesis, using a mathematical SIS model. As expected, epidemics starting in healthcare facilities with the highest centrality measures, especially high connectedness (degree) and a high number of patient transfers (strength), had the highest

probabilities of being sustained in time. Therefore, facility degree and strength should be considered as main targets for large-scale infection control strategies. More modelling work is required for a better and more comprehensive understanding for the potential impact of control measures on spread of specific pathogens in the healthcare networks.

Hospital-acquired infections cover a large spectrum of pathogens; in order to address the most pressing issue faced by healthcare systems – multi-drug resistance – the second part of the thesis focused on carbapenemase-producing *Enterobacteriaceae*. CPE episodes have been widespread in France since their first introduction in 2004 and have continued to rise in number every year. CPE episodes from September 2010 to December 2015 were described during the thesis and were used to forecast the number of expected episodes within the following four-year period using SARIMA models. Overall, the CPE epidemic in France was characterised by a majority of CPE strains carrying the OXA-48 resistance gene. The CPE epidemic was predicted to double in size in terms of the number of episodes by the end of 2019; however, the number of multi-case episodes and outbreaks were predicted to stabilize.

In order to help better understand and develop effective control strategies against the French CPR epidemic, the thesis then aimed at assessing the contribution of patient transfer patterns on CPE spread. This was done using the previously described general patient healthcare network in France in 2014 (191) and an adaptation of a previously published statistical method to empirically test the contribution of this network on CPE spread over the 2012 to 2015 period.(199) CPE episodes were found to be significantly supported by the patient transfer data. A transition was also observed in 2013, from an epidemic sustained by importation to an epidemic sustained by local transmission events. As a consequence, the study suggested that coordinated prevention and infection control strategies should now focus on at-risk patient transfers of carriers of CPE to reduce regional and inter-regional transmission. Preliminary work was performed to attempt to reconstruct CPE transmission chains among French hospitals; this should be pursued.

In summary, the work conducted during the thesis described the structure of French healthcare networks and assessed the role that patient transfers may have in the transmission of pathogens in the healthcare setting. The particular example of the role of patient transfers in CPE transmission serves to highlight the importance of considering the network structure in developing novel infection prevention and control strategies.

Chapter 12. Research limitations and perspectives

Although the research aims were met during the course of the thesis, several limitations may be identified. Even if the criteria of the systematic review was to identify mathematical and mechanistic models of HAI spread, the contributions of statistical models in the understanding of transmission dynamics of pathogen spread in healthcare settings could have also been useful. Statistical models are not only useful in describing observational studies but can also provide insights in spread dynamics of HAIs through multistate models such as survival models or competing risks models that can elucidate the transition probabilities of different states of infection.(200) The review did not include publications using social network analyses that can also provide valuable information on the impact of contact and transfer network structure on potential HAI spread dynamics.(109, 155, 170, 201, 202) Despite these exclusions, the review was able to identify over 200 studies and in which 76 were examined in detail, providing new insights into the implementation of contact and transfer network data to further develop mathematical models in healthcare settings. Future reviews either focused on the role of statistical methods in the improving the understanding of pathogen spread dynamics in these settings or the role of network structure in terms of network topology characteristics in order to detail complex social structures should also be conducted.

Three main limiting factors relating to the nature of the PMSI database were encountered in the development of the healthcare networks. The first limitation dealt with the coding bias of HAI infection in the PMSI database in order to appropriately identify patients with HAIs; however, the literature review and the inclusion of a second set of codes in order to identify other potential HAI patients addressed this limitation. However, it should be noted that even with the addition of a second set of more sensitive and specific codes, other data sources such as laboratory confirmed results of infections could have been used to identify patients. Nevertheless, this task would have been very difficult to achieve given the timeframe of the thesis due to the extensive size of the number of patients to evaluate for the entire country for the entire year of 2014 and potential issues regarding access to these data. Overall, for the purpose of the healthcare network analyses, it was deemed sufficient to have included the HAI-specific and suspected-HAI patient transfers in order to produce networks that reflected the overall structure of transfer patterns of HAI patients in order to sufficiently compare them to the general patient population.

The second limitation regarding the quality of the PMSI data was in regard to the healthcare facility identification numbers. Many of the university hospitals represented more than one

public hospital or healthcare facility due to the sharing of the same FINESS number. For example, the largest outlier hub in Paris (AP-HP) constituted 39 hospitals but was represented as one vertices in the network – a very significant outlier in terms of connectivity. Consequently, university hospital centres accommodated a larger patient population than hospital centres or local hospitals, influencing the network characteristics which may have led to overestimation of the specific patient movement patterns to and from these centres. However, the high concentration of other hospitals within proximity of these public hospital hubs demonstrates that despite this issue, major cities such as Paris play the most important role in connecting patients in the national network.

Finally, the third limitation we encountered concerning the use of the PMSI database was the fact that patients diagnosed with or patients suspected to have had an HAI could be treated and clear the infection before being transferred. In addition, pathogens due to asymptomatic carriage, the diagnoses performed within hospitals may not have allowed identifying all carriers of potential HAI pathogens so we may have missed these important patient transfers. However, despite these limitations in the sensitivity and specificity to be able to detect both asymptomatic carriers and carriage status during transfer, our results suggest that this information is not mandatory since all three networks rely on the same underlying structure. As a result, we are able to show that models of the general patient population can reliably inform us on the transfer patterns of different patient sub-populations.

Among the initial aims of the thesis was the development of a large meta-population-like model of pathogen spread in the French healthcare network. As discussed in the systematic review, mathematical models provide not only a theoretical framework to better understand pathogen spread dynamics in healthcare settings but can also lead to the testing and evaluation of novel infection control strategies that can inform healthcare systems on how to more effectively manage and reduce the burden of HAIs. A simple SIS-like model was developed in the early stages of the thesis and insights into how healthcare facility network characteristics may impact epidemic spread were given. Nonetheless, the model lacked the complexity to inform healthcare professionals on how to better manage at-risk HAI patients and the model did not entail the assessment of infection control measures such as the impact of screening based on different definitions of “at-risk” patients. In addition, the model should be further developed to include important hospital co-factors such as antibiotic exposure and the frequency of surgical procedures. One possible approach would be using heterogeneous probabilities of transmission

between hospitals, depending on hospital characteristics, to account for these important variations and better understand the dynamics of pathogen spread in the healthcare networks.

The thesis research, however, has been a precursor to a larger collaborative project that has been funded from 2018 to 2021 entitled "Spread of Pathogens on Healthcare Institutions Networks: a modelling study (SPHINx)" whose aim is to 1) develop a modelling platform to integrate different levels of HAI transmission from pathogen selection within wards to their dissemination between healthcare facilities and 2) use this platform to evaluate control strategies at the local, regional and national level. Specific strategies, such as the impact of antibiotic exposure in addition to the different antibiotic policies among different hospitals, can then be taken into account in the individual and hospital-level models. A national-level meta-population-like model of the healthcare network is also planned in the project to address the shortcomings of the thesis due to time constraints. The research that has been conducted so far during the thesis and a number of ideas of novel infection control strategies based on network structure that have been proposed will be of value for future development of the project.

A few limitations were also faced regarding data on CPE episodes in France. In the time series project for example, we had to rely on CPE data from the national surveillance network up to December 2015 even though the study began after June 2016. Not including valuable observations for the year 2016 in the model led to weaker estimates of the SARIMA model parameters, larger prediction intervals, and forecasts of 2016 CPE episodes with limited usefulness because they were in part outdated. Therefore, the advancement of the project relies on more recent CPE episodes becoming available in order to improve the predictive capacity of the models and to provide more timely and pertinent predictions of CPE episodes. This work is currently being prepared.

The second limitation was due to a lack of information regarding hospitals where the CPE episodes were reported. Indeed, the work regarding the assessment of the role of inter-facility transfers on CPE spread dynamics had to rely solely on the department of the episodes. As a result, transfer rates between hospitals of the same department were assumed to be homogenous that reduced the power of the study. However, statistically significant differences between the network-supported path distances in the observed CPE episode data and the permutations were still observed due to sufficient heterogeneity in the transfer rates between departments. Since hospital-level CPE data has been recently attained, future work involving modelling the

transmission of CPE on the healthcare network will lead to the use of the entire patient transfer network and thus, a better representation of the epidemic on the network.

Cautions regarding the interpretation of the forecasts of CPE incidence should be taken. Limitations concerning detection and declaration bias may have biased the results. CPE incidence data (and our forecasting results, as a consequence) could in fact reflect the increase in declaration rates over time by healthcare facilities and laboratories and not the true burden of disease. Some facilities may be more active in detecting CPE while others may not. In addition, some facilities may be less likely to declare and report single case CPE episodes while others may not. For example, university hospitals may not necessarily have more cases but rather declare CPE due to protocol or available resources that facilitate detection. On the other hand, some facilities may be less likely to report single case episodes or to investigate their links with other cases and as a result, we may under estimate the incidence of these episodes.

The SARIMA models gave simple conservative estimates concerning the predictions of CPE. These models did not take into account the mechanisms that may lead to spread and relied on the assumption that declaration behaviour, control strategies, and spreading dynamics will not change over time. In addition, these models consider a linear trend, while in fact pathogens do not spread linearly but rather may result in multiple cases. Therefore, the model estimates may actually underestimate the future incidence. However, despite the simplicity and potentially conservative estimates of future CPE episode incidence, the study aimed to underline the importance of improving prevention and infection control strategies aimed at CPE spread. It may be very likely that the burden of disease will grow over time and have consequences in terms of disease risk ad cost.

Finally, CPE infection from community settings should have also been considered. Potential CPE cases are likely to be occurring in the community which can lead to non-identification and poor control of any potential cross-infections. Future work should also assess the impact that CPE community transmission; and models should consider the impact of guidelines for screening and controlling CPE in the community.

In conclusion, many limitations were encountered regarding the quality of research of the thesis; however, some of these limitations were either addressed directly or will be addressed in future work as outlined in the SPHINx project.

Chapter 13. Public health implications

Improving the current knowledge of the dynamics and mechanisms that lead to the spread of pathogens in healthcare settings and the proposal of innovative prevention and infection control measures stemming from this knowledge are of major public health importance. HAIs strain healthcare systems all over the world due to several factors: increasing the burden of disease of patients who risk longer hospital stays and even death, posing a burden to healthcare staff, and a financial burden to the system in terms of providing an inadequate number of isolation rooms and hospital beds to high-risk and infected patients. Among all HAI pathogens, extensively drug-resistant strains pose a major threat to patient safety and in some cases, prevention measures may be the sole methods to combat their dissemination when treatment with antimicrobials is no longer an option. As a consequence of the rising need of novel prevention and infection control measures to combat the spread of pathogens in the healthcare setting, this thesis sought to elucidate new avenues of infection control research targeting common HAIs and multi-drug resistant *Enterobacteriaceae* through the modelling of pathogen spread dynamics. In order to achieve these objectives given the timeframe of the thesis, three publications in international peer-reviewed journals were published or submitted: a systematic review of mathematical modelling of pathogen spread in healthcare settings, an extensive analysis and comparison of the French healthcare networks, and an assessment of the role of patient transfers in the transmission of CPE in France. The following discussion aims to enlighten the public health implications of these findings and how they may provide a body of work to support novel ways of addressing HAI prevention and control at the national-scale, in particular for the case of France, but to other contexts as well.

One avenue in which innovative prevention and control measures have been proposed has been the field of mathematical modelling. Mathematical models have provided a theoretical framework for understanding complex transmission dynamics within healthcare settings for over 15 years (112-115). Furthermore, they have provided a quantitative approach to estimating the impact of various infection control strategies and their combined effects (113-115, 203). The number of publications on mathematical models of infections in healthcare settings have become more frequent over the years. Multiple factors may have led to this observed increase including perceived usefulness of models as tools for understanding the impact of infection prevention and control in the health field, for understanding drivers of recent major epidemics such as the 2002-2003 SARS outbreak (204-207) and the 2014-2015 Ebola epidemic (208-211) or growing awareness of factors contributing to the global impact of antibiotic resistance (212).

Due to growing use of digital data in the field of epidemiology over the years such as an increased availability of digitalized medical records and development of sensor technology to monitor inter-individual contacts have provided researchers with the means to build more realistic models. Further innovations in data collection on both network structure and infection, implementation of the data in modelling, and calibration and validation of the data in the models are required to further reinforce existing recommendations and to evaluate new control strategies in healthcare settings.

The analysis of the French healthcare networks provided a first description of patient transfer patterns at the French national level. The study showed that transfer patterns of patients who had an HAI and other patients with different diseases and comorbidities were subject to the same network dynamics. In comparison to other healthcare networks in England (14, 175), the Netherlands (13), Scotland (15), or the United States (16), the French healthcare networks were very centralized systems. Public university hospital centres and private hospitals in the main metropoles of France dominated patient flow. Studies in France have shown that highly connected hospitals may harbour more MRSA and MRSA bacteraemia cases (13, 167, 175, 213) and that HAIs were overall most prevalent in cancers centres, university hospitals, and armed forces hospitals (214). Therefore, these healthcare facilities may have the most potential to transmit HAIs in the entire network through carriage by infected or colonized patients. Pathogens in healthcare settings may spread at a higher rate than expected at random due to the centralization of patient movement and due to the small average number of transfers required for patients to move throughout the network.

These findings were in line with the preliminary SIS-like model results showing that the highest probability of a sustained epidemic occurred when hub hospitals were initially infected. A simulation model of the English healthcare network also identified university teaching hospitals as hubs for both MRSA incidence and recommended them as ideal targets for intervention measures such as screening of patients discharged from these hospitals as a more efficient means of control of pathogen spread compared to universal screening measures.(175) In parallel, the French healthcare network study also points to hubs as targets for sentinel surveillance in addition to priority targets of HAI control strategies to achieve the most effective reduction in transmission across the country.(161) Hence, in any context where patients are being transferred from hub healthcare centres, special attention should be paid to any potential HAI-related risks during admission.

In recent years, studies that have modelled the spread of infection in healthcare networks have argued for regional coordinated control strategies as one of the most effective ways to reduce pathogen dissemination in healthcare settings at regional and nation levels.(110, 175) The regional community clustering findings from the French healthcare network study were similar to that of the healthcare network of England in which the majority of transfers occurred through intra-regional community patient sharing and patient flows centred towards the regional university hospital within the community.(175) Furthermore, the French study demonstrated that a two-tier hospital community structure existed. Healthcare communities were identified at both the regional level, consistent with the French administrative regions, and at the sub-regional or department-level. Subtle differences between department-level communities of the suspected-HAI and the general network may be important in distinguishing hospitals with higher potential to harbour HAI pathogens, with possible consequences in terms of spread prediction; however, this requires further study. Coordinated local control measures such as screening at-risk patient transfers and increasing contact precautions based on the centrality of a neighbouring discharge hospital with known cases may be the first line of defence against pathogen spread within the regions before spread reaches the hub university hospitals through intra-community transfer.

Important intermediary trajectories may play a key role in the spread of pathogens between hub hospitals and between communities. A study has shown that modifying the number of patients moving between communities may reduce the spread of MRSA for example.(14) The same authors also showed that even though a healthcare facility's strong connections were important risk factors for a direct neighbour, weaker connections also offered ideal indirect routes for pathogens to travel further and faster in the network.(215) In addition, for the case of CPE, they found that in terms of absolute numbers of colonised patients admitted to a hospital by transfers from the same region compared to transfers coming from outside their region, transfers occurring within the same region posed more of threat.(216) Therefore, these studies parallel the observations of a two-tier-like network structure in which infection risks depends on the two levels. Slightly weaker links at the sub-regional level may play a more important role in the spread dynamics in terms of absolute number of transfers and thus potentially infection risk compared to the highly weighted and connected inter-community links between hubs. The simulation of pathogen spread in the networks showed that albeit the hubs introducing infection from the large metropole across to the regions, the highest number of infected hospitals resulted from regional university hospitals disseminating the infection to the local sub-regional clusters.

The two-tier structure may inform coordinated strategies at a more local level where healthcare facilities not only identify at-risk patients transferred from hub university hospital centres but also consider the risks coming from neighbouring local hospital centres. Infection control strategies – for short-term control – should rely more on the local department-level dynamics to minimize hospital-level outbreaks and transmission to neighbouring hospitals. In the long term, regional community dynamics may give clues regarding the gradual propagation of specific strains of pathogens over time (assuming carriage patterns follow that of patient flow patterns in the healthcare networks). Future studies are needed to validate these recommendations and to quantify the control measures. In addition, other studies are also required to assess the temporal dynamics of pathogen spread in networks in order to identify any potential seasonality patterns of flow and how to prevent emerging multi-drug resistant bacteria from becoming endemic.

Long-term prevention measures are also needed to prevent emerging pathogens from becoming endemic. Reducing hospital connectedness in order to reduce the risk of spread of pathogens in networks is at the core of many novel infection control proposals.(139, 175) Decentralization of the healthcare system and more specifically human resource and specialized health services towards the regional and department levels may help reduce the high connectedness of hubs in the metropole centres and redirect patient transfers; however, this remains to be tested. France previously moved towards regionalization strategies with the creation of regional hospital agencies, albeit not very effective and current trends gear more towards centralized healthcare.(217, 218) In addition, the number of university hospitals may be insufficient in France, below that of the UK, a country with a similar population size. Therefore, one structural solution to alleviate the burden of pathogen spread could be increasing the number facilities providing specialized services and distributing them at the local level to help redirect patient flow and potentially avoid large-scale pathogen dispersal.

One notable source of concern is the risk of CPE becoming endemic. CPE episodes have become widespread in France and the number of episodes continue to rise every year. The Ile-de-France region, which includes Paris and neighbouring French departments, has the highest incidence of CPE episodes, including the highest number of episodes linked to internationally imported cases. Paris serves as a healthcare hub and attracts a high number of patients seeking specialized care which may put them at higher risk for CPE infection. The infection risks are thus two-fold: patient seeking specialized care may increase individual risk factors to infection because specialized services may entail surgery and other invasive procedures; and also, patient

exposure to a higher number of potential contacts with CPE carriers due to high CPE incidence. The time series study predicted stabilization of episodes with multiple cases in France for the next few years. Therefore, one can assume that the control measures already in place during the 2010-2015 period to control outbreaks have been effective and should continue in order to control transmission to avoid recurring hospital outbreaks. It should be noted, however, that multiple case episodes represented a range between two to 200 CPE cases per episode. Therefore, a higher number of person-to-person transmission events may continue to occur; however, the number of outbreaks were predicted to occur at a stabilized frequency. The study also predicted an increase in the number of single case episodes; but, this may be due the surveillance system failing to link transmission events or incomplete surveillance data.

The previously mentioned prevention and control strategies for pathogen spread in healthcare settings are especially pertinent for the CPE epidemic. The dynamics of CPE transmission in France have changed over time. Mounting evidence for local spread through transfers emerged in 2014 followed by the strongest evidence for transfer network-supported CPE transmission in 2015. These results suggest that between 2013 and 2014 there was a growing contribution of regional and inter-regional transfers in the spread of CPE in France which is in concordance with reports by the ECDC.(89) In addition, the estimated delay in notification of CPE episodes between hospitals due to patient transfer was observed between 20 to 30 days after initial notification in an index hospital. Therefore, this recent study not only linked CPE transmission events to patient transfers in recent years, but also estimated delay in notification of these events. These two findings reinforce recommendations highlighting the importance of considering patient transfer as critical risk factors for CPE introduction that could help the surveillance system to estimate high risk periods for outbreaks linked to patient transfers from hospitals with known CPE cases, and that could lead to public health authorities to take action during critical periods to control spread.

Network structure can also explain the observations of these local transmission events. Both the number of linked episodes occurring in the same department increased over time and the proportion of linked episodes occurring in different departments occurred within shorter geographic proximity over time as well. As previously mentioned, local department-level community patient sharing was expected to play an important role in infection spread dynamics. The increasing number of these possible CPE transmission events occurring in the same department may be explained by the fact that most patient transfers in the healthcare network occurred at the local level. Therefore, the largest proportion of CPE transmission may have

occurred between neighbouring hospitals and public health authorities should take in to consideration the importance of monitoring local patient transfers for potential at-risk CPE carriers in order to have the most effective impact on the CPE epidemic.

While patient transfers are certainly not the sole explanation for the augmentation in observed CPE episodes between 2013 and 2015 and other factors such as antibiotic exposure should have been considered, these studies suggest that patient transfer has played an increasingly significant role over time. Episodes from international importation could also have contributed to almost half of the spread of CPE in France for example. These results are consistent with the outbreak descriptions observe in the literature in which both imported and non-imported cases have led to secondary cases of CPE in different hospitals. In addition, the heterogeneity in infection control policies across different types of healthcare facilities in France and limited implementation of specific strategies to control CPE may have led to poor control of CPE and in consequence, dissemination over time.(219)

Our results may also help elucidate how patient transfers can serve as a mechanism of spread of CPE in France. Since asymptomatic carriage of CPE could occur and there may be heterogeneous detection of colonised patients across different healthcare facilities, understanding patient transfer patterns may also help elucidate the potential risks for new introductions in the healthcare network.

There was no observed association between the number of cases per potential infector episode and the number of secondary episodes. On one hand this may suggest that control measures have prevented large hospital outbreaks from causing multi-department outbreaks during the 2014-2015 period; on the other hand, as previously mentioned in the time series study, most reports were single-case episodes suggesting a potential failure of surveillance authorities in identifying single-cases as part of the same chain of transmission of other reported episodes.

Network dynamics may also help explain CPE outbreaks. The number of CPE spreading events involving multiple episodes has increased over time. For example, an imported OXA-48 episode in Paris was linked to nine other episodes in nine different departments in France in 2015. Paris has been identified as the largest hub for not only CPE episodes linked to importation but also for patient transfer. Other examples of multiple spreading events involving hubs further highlight the important role of links with a large number of patient transfers in connecting geographically distant hospitals in terms of HAI transmission. These observations underline the importance for health authorities to improve control efforts in large and highly

connected metropoles and healthcare facilities. These efforts also go hand-in-hand with coordination between the regional surveillance systems, local expert laboratories and regional health authorities in order to rapidly identify CPE cases. Healthcare facilities should also be urged to quickly notify any cases or potential contacts. In addition, screening measures, contact precautions, and strict cohorting of patients (which have been shown especially effective in one particular CPE outbreak (9, 95)) should be implemented once cases are identified.

Conclusion

Healthcare networks have been important in elucidating the role of patient transfer patterns in pathogen spread in healthcare settings. The major limiting factor of this work has been the lack of a detailed modelling study quantifying the impact of pathogen spread in the healthcare networks; however, this body of work provides a foundation for future work on modelling spread. In addition, the various findings of the studies conducted during the course of the thesis may help enlighten public health implications of patient transfers on spread dynamics in order to provide a body of work to support novel ways of addressing prevention and control in healthcare settings at the national-scale. These studies support regional coordinated efforts at the local, regional, and national level between healthcare facilities that requires the aid of the surveillance system in order to coordinate these efforts and mobilize facilities to implement new measures accordingly. In addition, these efforts require the cooperation of university hospitals who play an important role in the healthcare network structure. At-risk patient transfer identification using network topology measures may prove useful for specific pathogens such as multi-drug resistance bacteria. CPE episodes have been linked to both international importation and local spread; as a result, it is of upmost importance for healthcare network dynamics to be considered in the prevention and infection control process. This thesis serves to highlight the importance of healthcare network structure in the development of effective prevention and infection control measures.

References

1. Report on the burden of endemic health care-associated infection worldwide. Geneva: WHO; 2011.

2. ECDC. Point prevalence survey of healthcare associated infections and antimicrobial use in European acute care hospitals. 2013.

3. ECDC. Annual epidemiological report on communicable diseases in Europe 2008: report on the state of communicable diseases in the EU and EEA/EFTA countries. 2008.

4. Carlet J. Les infections liées aux soins médicaux. Haut Conseil de la Santé Publique; 2002.

5. Les coûts induits par les infections nosocomiales: Sénat; [updated 19 April 2018. Available from: https://www.senat.fr/rap/r05-421/r05-42113.html.

6. Surveillance des bactéries multirésistantes dans les établissements de santé. Réseau BMR-Raisin, France. Résultats 2015. . Saint-Maurice: Santé publique France; 2017.

7. Risk Assessment on the Spread of Carbapenemase-Producing Enterobacteriaceae (CPE) through Patient Transfer between Healthcare Facilities, with Special Emphasis on Cross-Border Transfer. European Centre for Disease Prevention and Control; 2011.

8. Nordmann P, Naas T, Poirel L. Global Spread of Carbapenemase-producing Enterobacteriaceae. Emerging Infectious Diseases. 2011;17(10):1791-8.

9. Vaux S, Carbonne A, Thiolet JM, Jarlier V, Coignard B, RAISIN, et al. Emergence of carbapenemase-producing Enterobacteriaceae in France, 2004 to 2011. Eurosurveillance. 2011;16(22):19880.

10. Naas T. Quand, comment, et pourquoi dépister des entérobactéries productrices de carbapénèmases. Apport de la PCR et des tests rapides.: Cpias Auvergne Rhone Alpes; 2017 [Available from: http://www.cpias-auvergnerhonealpes.fr/Journees/JASE/2017/9_naas_resistance_enterobacteries.pdf.

11. Eveillard M, Quenon J-L, Rufat P, Mangeol A, Fauvelle F. Association Between Hospital-Acquired Infections and Patients' Transfers. Infection Control & Hospital Epidemiology. 2001;22(11):693-6.

12. Robotham JV, Scarff CA, Jenkins DR, Medley GF. Meticillin-resistant *Staphylococcus aureus* (MRSA) in hospitals and the community: model predictions based on the UK situation. Journal of Hospital Infection. 2007;65:93-9.

13. Donker T, Wallinga J, Grundmann H. Patient referral patterns and the spread of hospital-acquired infections through national health care networks. Plos Computational Biology. 2010;6(3).

14. Donker T, Wallinga J, Grundmann H. Dispersal of antibiotic-resistant high-risk clones by hospital networks: changing the patient direction can make all the difference. Journal of Hospital Infection. 2013;86(1):34-41.

15. van Bunnik BAD, Ciccolini M, Gibbons CL, Edwards G, Fitzgerald R, McAdam PR, et al. Efficient national surveillance for health-care-associated infections. BMC Public Health. 2015;15.

16. Fernández Gracia J, Onnela J-P, Barnett ML, Eguíluz VM, Christakis NA. Spread of pathogens in the patient transfer network of US hospitals. Physics and Society. 2015.

17. Assab* R, Nekkab* N, Crepey P, Astagneau P, Guillemot D, Opatowski L, et al. Mathematical models of infection transmission in healthcare settings: recent advances from the use of network structured data. Current Opinion in Infectious Diseases. 2017;30(4):410-8.

18. Gerbier S, Bouzbid S, Pradat E, Baulieux J, Lepape A, Berland M, et al. Use of the French medico-administrative database (PMSI) to detect nosocomial infections in the University Hospital of Lyon. Revue d'Epidémiologie et de Santé Publique. 2011;59(1):3-14.

19. Clauset A, Shalizi CR, Newman MEJ. Power-law distributions in empirical data. 2007.

20. Rosvall M, Bergstrom CT. Maps of random walks on complex networks reveal community structure. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(4):1118-23.

21. Prevention of hospital-acquired infections: a practical guide. Geneva: World Health Organization (WHO); 2002. Report No.: WHO/CDS/CSR/EPH/2002.12.

22. Horan TC, Andrus M, Dudeck, Margaret A. CDC/NHSN surveillance definition of health care–associated infection and criteria for specific types of infections in the acute care setting. American Journal of Infection Control. 2008;36(5):309-32.

23. von Wintersdorff CJH, Penders J, van Niekerk JM, Mills ND, Majumder S, van Alphen LB, et al. Dissemination of Antimicrobial Resistance in Microbial Ecosystems through Horizontal Gene Transfer. Front Microbiol. 2016;7.

24. Tacconelli E, Carrara E, Savoldi A, Harbarth S, Mendelson M, Monnet DL, et al. Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. The Lancet Infectious Diseases.

25. Hirshon JM, Risko N, Calvello EJB, Stewart de Ramirez S, Narayan M, Theodosis C, et al. Health systems and services: the role of acute care. Bulletin of the World Health Organization. 2013;91(5):386-8.

26. Magill SS, Edwards JR, Bamberg W, Beldavs ZG, Dumyati G, Kainer MA, et al. Multistate Point-Prevalence Survey of Health Care–Associated Infections. New England Journal of Medicine. 2014;370(13):1198-208.

27. Klevens RM, Edwards JR, Richards CL, Horan TC, Gaynes RP, Pollock DA, et al. Estimating Health Care-Associated Infections and Deaths in U.S. Hospitals, 2002. Public Health Reports. 2007;122(2):160-6.

28. ECDC. Healthcare-associated infections acquired in intensive care units. Stockholm; 2017.

29. Cassini A, Plachouras D, Eckmanns T, Abu Sin M, Blank H-P, Ducomble T, et al. Burden of Six Healthcare-Associated Infections on European Population Health: Estimating Incidence-Based Disability-Adjusted Life Years through a Population Prevalence-Based Modelling Study. PLOS Medicine. 2016;13(10):e1002150.

30. Scott RD. The direct medical costs of healthcare-associated infections in U.S. hospitals and the benefits of prevention. Atlanta: Division of Healthcare Quality Promotion National Center for Preparedness, Detection, and Control of Infectious Diseases Coordinating Center for Infectious Diseases; 2009.

31. Tacconelli E, Smith G, Hieke K, Lafuma A, Bastide P. Epidemiology, medical outcomes and costs of catheter-related bloodstream infections in intensive care units of four European countries: literature- and registry-based estimates. Journal of Hospital Infection. 2009;72(2):97-103.

32. Surveillance des infections du site opératoire dans les établissements de santé français. Résultats 2015. Saint-Maurice: Santé publique France; 2017.

33. Surveillance des infections nosocomiales en réanimation adulte, Réseau REA-Raisin, France. Résultats 2015. Saint-Maurice: Santé publique France; 2017.

34. Recent trends in antimicrobial resistance among Streptococcus pneumoniae and Staphylococcus aureus isolates: the French experience. Eurosurveillance. 2008;13(46):19035.

35. Mayhall GC. Hospital epidemiology and infection control. 4th ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2012.

36. Hidron AI, Edwards JR, Patel J, Horan TC, Sievert DM, Pollock DA, et al. Antimicrobial-Resistant Pathogens Associated With Healthcare-Associated Infections: Annual Summary of Data Reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2006–2007. Infection Control & Hospital Epidemiology. 2015;29(11):996-1011.

37. Watkins RR, Bonomo RA. 140 - β-Lactam Antibiotics. In: Cohen J, Powderly WG, Opal SM, editors. Infectious Diseases (Fourth Edition): Elsevier; 2017. p. 1203-16.e2.

38. Rolinson GN. Evolution of beta-lactamase inhibitors. Reviews of infectious diseases. 1991;13 Suppl 9:S727-32.

39. Ambler RP. The structure of β-lactamases. Philosophical Transactions of the Royal Society of London B, Biological Sciences. 1980;289(1036):321-31.

40. Papp-Wallace KM, Endimiani A, Taracila MA, Bonomo RA. Carbapenems: Past, Present, and Future. Antimicrobial Agents and Chemotherapy. 2011;55(11):4943-60.

41. Birnbaum J, Kahan FM, Kropp H, Macdonald JS. Carbapenems, a new class of beta-lactam antibiotics. The American Journal of Medicine. 1985;78(6):3-21.

42. Kahan J, Kahan F, Goegelman R, Currie S, Jackson M, Stapley E, et al. Abstracts XVI, Interscience Conference on Antimicrobial Agents and Chemotherapy, Chicago, Ill. 1976;227.

43. Kahan JS, Kahan FM, Goegelman R, Currie SA, Jackson M, Stapley EO, et al. Thienamycin, a new beta-lactam antibiotic. I. Discovery, taxonomy, isolation and physical properties. The Journal of antibiotics. 1979;32(1):1-12.

44. Meletis G. Carbapenem resistance: overview of the problem and future perspectives. Therapeutic Advances in Infectious Disease. 2016;3(1):15-21.

45. Falagas ME, Karageorgopoulos DE, Nordmann P. Therapeutic options for infections with Enterobacteriaceae producing carbapenem-hydrolyzing enzymes. Future microbiology. 2011;6(6):653-66.

46. Falagas ME, Lourida P, Poulikakos P, Rafailidis PI, Tansarli GS. Antibiotic Treatment of Infections Due to Carbapenem-Resistant Enterobacteriaceae: Systematic Evaluation of the Available Evidence. Antimicrobial Agents and Chemotherapy. 2014;58(2):654-63.

47. Rafailidis PI, Falagas ME. Options for treating carbapenem-resistant Enterobacteriaceae. Curr Opin Infect Dis. 2014;27(6):479-83.

48. Lee J, Lee S. Carbapenem Resistance in Gram-negative Pathogens: Emerging Non-metallo-carbapenemases. Research Journal of Microbiology. 2006;1:1-22.

49. Queenan AM, Bush K. Carbapenemases: the Versatile β-Lactamases. Clinical Microbiology Reviews. 2007;20(3):440-58.

50. Yamamoto T, Tanaka M, Nohara C, Fukunaga Y, Yamagishi S. Transposition of the oxacillin-hydrolyzing penicillinase gene. Journal of Bacteriology. 1981;145(2):808-13.

51. Yang YJ, Wu PJ, Livermore DM. Biochemical characterization of a beta-lactamase that hydrolyzes penems and carbapenems from two Serratia marcescens isolates. Antimicrobial Agents and Chemotherapy. 1990;34(5):755-8.

52. Naas T, Vandel L, Sougakoff W, Livermore DM, Nordmann P. Cloning and sequence analysis of the gene for a carbapenem-hydrolyzing class A beta-lactamase, Sme-1, from Serratia marcescens S6. Antimicrobial Agents and Chemotherapy. 1994;38(6):1262-70.

53. Osano E, Arakawa Y, Wacharotayankun R, Ohta M, Horii T, Ito H, et al. Molecular characterization of an enterobacterial metallo beta-lactamase found in a clinical isolate of Serratia marcescens that shows imipenem resistance. Antimicrobial Agents and Chemotherapy. 1994;38(1):71-8.

54. Yotsuji A, Minami S, Inoue M, Mitsuhashi S. Properties of novel beta-lactamase produced by Bacteroides fragilis. Antimicrobial Agents and Chemotherapy. 1983;24(6):925-9.

55. Medeiros A, Hare R. Beta-lactamase mediated resistance to penems and carbapenems among Enterobacteriaceae [abstract 116]. 26th Interscience Conference on Antimicrobial Agents and Chemotherapy; New Orleans. Washington DC: American Society for Microbiology; 1986.

56. Paton R, Miles RS, Hood J, Amyes SGB, Miles RS, Amyes SGB. ARI 1: β-lactamase-mediated imipenem resistance in Acinetobacter baumannii. International Journal of Antimicrobial Agents. 1993;2(2):81-7.

57. Donald HM, Scaife W, Amyes SGB, Young H-K. Sequence Analysis of ARI-1, a Novel OXA β-Lactamase, Responsible for Imipenem Resistance in Acinetobacter baumannii 6B92. Antimicrobial Agents and Chemotherapy. 2000;44(1):196-9.

58. Nordmann P, Mariotte S, Naas T, Labia R, Nicolas MH. Biochemical properties of a carbapenem-hydrolyzing beta-lactamase from Enterobacter cloacae and cloning of the gene into Escherichia coli. Antimicrobial Agents and Chemotherapy. 1993;37(5):939-46.

59. Queenan AM, Torres-Viera C, Gold HS, Carmeli Y, Eliopoulos GM, Moellering RC, et al. SME-Type Carbapenem-Hydrolyzing Class A β-Lactamases from Geographically Diverse Serratia marcescens Strains. Antimicrobial Agents and Chemotherapy. 2000;44(11):3035-9.

60. Poirel L, Naas T, Nicolas D, Collet L, Bellais S, Cavallo J-D, et al. Characterization of VIM-2, a Carbapenem-Hydrolyzing Metallo-β-Lactamase and Its Plasmid- and Integron-Borne Gene from a Pseudomonas aeruginosa Clinical Isolate in France. Antimicrobial Agents and Chemotherapy. 2000;44(4):891-7.

61. Lauretti L, Riccio ML, Mazzariol A, Cornaglia G, Amicosante G, Fontana R, et al. Cloning and Characterization of bla(VIM), a New Integron-Borne Metallo-β-Lactamase Gene from a Pseudomonas aeruginosa Clinical Isolate. Antimicrobial Agents and Chemotherapy. 1999;43(7):1584-90.

62. Yigit H, Queenan AM, Anderson GJ, Domenech-Sanchez A, Biddle JW, Steward CD, et al. Novel Carbapenem-Hydrolyzing β-Lactamase, KPC-1, from a Carbapenem-Resistant Strain of Klebsiella pneumoniae. Antimicrobial Agents and Chemotherapy. 2001;45(4):1151-61.

63. Riccio ML, Franceschini N, Boschi L, Caravelli B, Cornaglia G, Fontana R, et al. Characterization of the Metallo-β-Lactamase Determinant of Acinetobacter baumannii AC-54/97 Reveals the Existence of bla(IMP) Allelic Variants Carried by Gene Cassettes of Different Phylogeny. Antimicrobial Agents and Chemotherapy. 2000;44(5):1229-35.

64. Poirel L, Le Thomas I, Naas T, Karim A, Nordmann P. Biochemical Sequence Analyses of GES-1, a Novel Class A Extended-Spectrum β-Lactamase, and the Class 1 Integron In52 from Klebsiella pneumoniae. Antimicrobial Agents and Chemotherapy. 2000;44(3):622-32.

65. Yigit H, Queenan AM, Rasheed JK, Biddle JW, Domenech-Sanchez A, Alberti S, et al. Carbapenem-Resistant Strain of Klebsiella oxytoca Harboring Carbapenem-Hydrolyzing β-Lactamase KPC-2. Antimicrobial Agents and Chemotherapy. 2003;47(12):3881-9.

66. Toleman MA, Simm AM, Murphy TA, Gales AC, Biedenbach DJ, Jones RN, et al. Molecular characterization of SPM-1, a novel metallo-beta-lactamase isolated in Latin America: report from the SENTRY antimicrobial surveillance programme. The Journal of antimicrobial chemotherapy. 2002;50(5):673-9.

67. Giakkoupi P, Tzouvelekis LS, Tsakris A, Loukova V, Sofianou D, Tzelepi E. IBC-1, a Novel Integron-Associated Class A β-Lactamase with Extended-Spectrum Properties Produced by an Enterobacter cloacae Clinical Strain. Antimicrobial Agents and Chemotherapy. 2000;44(9):2247-53.

68. Yan J-J, Hsueh P-R, Ko W-C, Luh K-T, Tsai S-H, Wu H-M, et al. Metallo-β-Lactamases in Clinical Pseudomonas Isolates in Taiwan and Identification of VIM-3, a Novel Variant of the VIM-2 Enzyme. Antimicrobial Agents and Chemotherapy. 2001;45(8):2224-8.

69. Poirel L, Weldhagen GF, Naas T, De Champs C, Dove MG, Nordmann P. GES-2, a Class A β-Lactamase from Pseudomonas aeruginosa with Increased Hydrolysis of Imipenem. Antimicrobial Agents and Chemotherapy. 2001;45(9):2598-603.

70. Woodford N, Tierno PM, Young K, Tysall L, Palepou M-FI, Ward E, et al. Outbreak of Klebsiella pneumoniae Producing a New Carbapenem-Hydrolyzing Class A β-Lactamase, KPC-3, in a New York Medical Center. Antimicrobial Agents and Chemotherapy. 2004;48(12):4793-9.

71. Poirel L, Héritier C, Tolün V, Nordmann P. Emergence of Oxacillinase-Mediated Resistance to Imipenem in Klebsiella pneumoniae. Antimicrobial Agents and Chemotherapy. 2004;48(1):15-22.

72. Castanheira M, Toleman MA, Jones RN, Schmidt FJ, Walsh TR. Molecular Characterization of a β-Lactamase Gene, bla(GIM-1), Encoding a New Subclass of Metallo-β-Lactamase. Antimicrobial Agents and Chemotherapy. 2004;48(12):4654-61.

73. Lee K, Yum JH, Yong D, Lee HM, Kim HD, Docquier J-D, et al. Novel Acquired Metallo-β-Lactamase Gene, bla(SIM-1), in a Class 1 Integron from Acinetobacter baumannii Clinical Isolates from Korea. Antimicrobial Agents and Chemotherapy. 2005;49(11):4485-91.

74. Palepou M-FI, Woodford N, Hope R, Colman M, Glover J, Kaufmann ME, et al. Novel class A carbapenemase, KPC-4, in an Enterobacter isolate from Scotland. 15th European Congress of Clinical Microbiology and Infectious Diseases (ECCMID); Copenhagen, Denmark: Clinical Microbiology and Infection; 2005. p. S106.

75. Queenan AM, Shang W, Schreckenberger P, Lolans K, Bush K, Quinn J. SME-3, a Novel Member of the Serratia marcescens SME Family of Carbapenem-Hydrolyzing β-Lactamases. Antimicrobial Agents and Chemotherapy. 2006;50(10):3485-7.

76. Yong D, Toleman MA, Giske CG, Cho HS, Sundman K, Lee K, et al. Characterization of a New Metallo-β-Lactamase Gene, bla(NDM-1), and a Novel Erythromycin Esterase Gene Carried on a Unique Genetic Structure in Klebsiella pneumoniae Sequence Type 14 from India. Antimicrobial Agents and Chemotherapy. 2009;53(12):5046-54.

77. Poirel L, Ros A, Carricajo A, Berthelot P, Pozzetto B, Bernabeu S, et al. Extremely Drug-Resistant Citrobacter freundii Isolate Producing NDM-1 and Other Carbapenemases Identified in a Patient Returning from India. Antimicrobial Agents and Chemotherapy. 2011;55(1):447-8.

78. Egowa R, Sawai T, Mitsuhashi S. Drug resistance to enteric bacteria. XII. Unique substrate specificity of penicillinase produced by R factor. Jpn J Microbiol. 1967;11:173-8.

79. Lee SH, Jeong SH. Nomenclature of GES-Type Extended-Spectrum β-Lactamases. Antimicrobial Agents and Chemotherapy. 2005;49(5):2148-50.

80. Doi Y, Paterson DL. Carbapenemase-Producing Enterobacteriaceae. Seminars in respiratory and critical care medicine. 2015;36(1):74-84.

81. Ludden C, Reuter S, Judge K, Gouliouris T, Blane B, Coll F, et al. Sharing of carbapenemase-encoding plasmids between Enterobacteriaceae in UK sewage uncovered by MinION sequencing. Microbial Genomics. 2017;3(7):e000114.

82. ECDC. Risk assessment on the spread of carbapenemase-producing Enterobacteriaceae (CPE) through patient transfer between healthcare facilities, with special emphasis on cross-border transfer. Stockholm; 2011.

83. van Duin D, Doi Y. The global epidemiology of carbapenemase-producing Enterobacteriaceae. Virulence. 2017;8(4):460-9.

84. Munoz-Price LS, Poirel L, Bonomo RA, Schwaber MJ, Daikos GL, Cormican M, et al. Clinical epidemiology of the global expansion of Klebsiella pneumoniae carbapenemases. The Lancet Infectious diseases. 2013;13(9):785-96.

85. Dortet L, Poirel L, Nordmann P. Worldwide Dissemination of the NDM-Type Carbapenemases in Gram-Negative Bacteria. BioMed Research International. 2014;2014:249856.

86. Benouda A, Touzani O, Khairallah MT, Araj GF, Matar GM. First detection of oxacillinase-mediated resistance to carbapenems in Klebsiella pneumoniae from Morocco. Annals of tropical medicine and parasitology. 2010;104(4):327-30.

87. Levast M, Poirel L, Carrër A, Deiber M, Decroisette E, Mallaval F-O, et al. Transfer of OXA-48-positive carbapenem-resistant Klebsiella pneumoniae from Turkey to France. Journal of Antimicrobial Chemotherapy. 2011;66(4):944-5.

88. Mairi A, Pantel A, Sotto A, Lavigne J-P, Touati A. OXA-48-like carbapenemases producing Enterobacteriaceae in different niches. European Journal of Clinical Microbiology & Infectious Diseases. 2018;37(4):587-604.

89. ECDC. Rapid risk assessment: Carbapenem-resistant Enterobacteriaceae. Stockholm; 2016 8 April 2016.

90. ECDC. Surveillance of antimicrobial resistance in Europe 2016. Annual Report of the European Antimicrobial Resistance Surveillance Network (EARS-Net). Stockholm; 2017.

91. Grundmann H, Glasner C, Albiger B, Aanensen DM, Tomlinson CT, Andrasević AT, et al. Occurrence of carbapenemase-producing Klebsiella pneumoniae and Escherichia coli in the European survey of carbapenemase-producing Enterobacteriaceae (EuSCAPE): a prospective, multinational study. The Lancet Infectious Diseases. 2016;17(2):153-63.

92. Robert J, Pantel A, Mérens A, Lavigne J-P, Nicolas-Chanoine M-H, on behalf of OsCRSG. Incidence rates of carbapenemase-producing Enterobacteriaceae clinical isolates in France: a prospective nationwide study in 2011–12*. Journal of Antimicrobial Chemotherapy. 2014;69(10):2706-12.

93. Entérobactéries productrices de carbapénèmases (EPC) Saint-Maurice: Santé Publique France; [updated 28 March 2017. Available from: http://invs.santepubliquefrance.fr/Dossiers-thematiques/Maladies-infectieuses/Infections-associees-aux-soins/Surveillance-des-infections-associees-aux-soins-IAS/Enterobacteries-productrices-de-carbapenemases-EPC.

94. Kassis-Chikhani N, Decré D, Gautier V, Burghoffer B, Saliba F, Mathieu D, et al. First outbreak of multidrug-resistant Klebsiella pneumoniae carrying blaVIM-1 and blaSHV-5 in a French university hospital. Journal of Antimicrobial Chemotherapy. 2006;57(1):142-5.

95. Kassis-Chikhani N, Saliba F, Carbonne A, Neuville S, Decre D, Sengelin C, et al. Extended measures for controlling an outbreak of VIM-1 producing imipenem-resistant Klebsiella pneumoniae in a liver transplant centre in France, 2003–2004. Eurosurveillance. 2010;15(46):19713.

96. Épisodes impliquant des EPC en France. Bilan épidémiologique national au 31 décembre 2015, France. Saint-Maurice: Santé publique France; 2017.

97. Episodes impliquant des entérobactéries productrices de carbapénèmases. Situation épidémiologique du 4 octobre 2011. Saint-Maurice: Santé Publique France; 2010 [updated 19 October 2010. Available from: http://invs.santepubliquefrance.fr/Dossiers-thematiques/Maladies-infectieuses/Infections-associees-aux-soins/Surveillance-des-infections-associees-aux-soins-IAS/Enterobacteries-productrices-de-carbapenemases-EPC/Episodes-impliquant-des-enterobacteries-productrices-de-carbapenemases.-Situation-epidemiologique-du-4-octobre-2011.

98. Naas T, Cuzon G, Babics A, Fortineau N, Boytchev I, Gayral F, et al. Endoscopy-associated transmission of carbapenem-resistant Klebsiella pneumoniae producing KPC-2 β-lactamase. Journal of Antimicrobial Chemotherapy. 2010;65(6):1305-6.

99. Cuzon G, Ouanich J, Gondret R, Naas T, Nordmann P. Outbreak of OXA-48-Positive Carbapenem-Resistant Klebsiella pneumoniae Isolates in France. Antimicrob Agents Chemother. 2011;55(5):2420-3.

100. Decre D, Birgand G, Geneste D, Maury E, Petit JC, Barbut F, et al. Possible importation and subsequent cross-transmission of OXA-48-producing Klebsiella pneumoniae, France, 2010. Eurosurveillance. 2010;15(46):19718.

101. Naas T, Nordmann P, Vedel G, Poyart C. Plasmid-Mediated Carbapenem-Hydrolyzing β-Lactamase KPC in a Klebsiella pneumoniae Isolate from France. Antimicrob Agents Chemother. 492005. p. 4423-4.

102. Kassis-Chikhani N, Decré D, Ichai P, Sengelin C, Geneste D, Mihaila L, et al. Outbreak of Klebsiella pneumoniae producing KPC-2 and SHV-12 in a French hospital. Journal of Antimicrobial Chemotherapy. 2010;65(7):1539-40.

103. Lepelletier D, Berthelot P, Lucet JC, Fournier S, Jarlier V, Grandbastien B. French recommendations for the prevention of 'emerging extensively drug-resistant bacteria' (eXDR) cross-transmission. The Journal of hospital infection. 2015;90(3):186-95.

104. Dortet L, Cuzon G, Ponties V, Nordmann P. Trends in carbapenemase-producing Enterobacteriaceae, France, 2012 to 2014. Eurosurveillance. 2017;22(6):30461.

105. Holman A-M. Étude épidémiologique des entérobactéries productrices de carbapénèmase à La Réunion de 2010 à 2015. Médecine humaine et pathologie.

106. Otter JA, Yezli S, French GL. The Role Played by Contaminated Surfaces in the Transmission of Nosocomial Pathogens. Infection Control & Hospital Epidemiology. 2011;32(7):687-99.

107. Cattuto C, Van den Broeck W, Barrat A, Colizza V, Pinton J, Vespignani A. Dynamics of person-to-person interactions from distributed RFID sensor networks. Plos One. 2010;5(7).

108. Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, et al. Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases. PLOS Medicine. 2008;5(3):e74.

109. Duval A, Obadia T, Martinet L, Boëlle P-Y, Fleury E, Guillemot D, et al. Measuring dynamic social contacts in a rehabilitation hospital: effect of wards, patient and staff characteristics. Scientific Reports. 2018;8(1):1686.

110. Ciccolini M, Donker T, Kock R, Mielke M, Hendrix R, Jurke A, et al. Infection prevention in a connected world: the case for a regional approach. International Journal of Medical Microbiology. 2013;303(6-7):380-7.

111. Ke W, Huang SS, Hudson LO, Elkins KR, Nguyen CC, Spratt BG, et al. Patient sharing and population genetic structure of methicillin-resistant Staphylococcus aureus. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(17):6763-8.

112. Opatowski L, Guillemot D, Boelle P, Temime L. Contribution of mathematical modeling to the fight against bacterial antibiotic resistance. Current Opinion in Infectious Diseases. 2011;24(3):279-87.

113. Bonten M, Austin D, Lipsitch M. Understanding the spread of antibiotic resistant pathogens in hospitals: mathematical models as tools for control. Clinical Infectious Diseases. 2001;33(10):1739-46.

114. Grundmann H, Hellriegel B. Mathematical modelling: a tool for hospital infection control. Lancet Infectious Diseases. 2006;6(1):39-45.

115. van Kleef E, Robotham J, Jit M, Deeny S, Edmunds W. Modelling the transmission of healthcare associated infections: a systematic review. BMC Infectious Diseases. 2013;13.

116. Austin DJ, Anderson RM. Studies of antibiotic resistance within the patient, hospitals and the community using simple mathematical models. Philosophical Transactions of the Royal Society B: Biological Sciences. 1999;354(1384):721-38.

117. Salathé M, Bengtsson L, Bodnar T, Brewer D, Brownstein J, Buckee C, et al. Digital Epidemiology. Plos Computational Biology. 2012;8(7).

118. Salathé M. Digital epidemiology: what is it, and where is it going? Life Sciences, Society and Policy. 2018;14(1):1.

119. Cooper B, Medley G, Scott G. Preliminary analysis of the transmission dynamics of nosocomial infections: stochastic and management effects. Journal of Hospital Infection. 1999;43(2):131-47.

120. Austin D, Bonten M, Weinstein R, Slaughter S, Anderson R. Vancomycin-resistant enterococci in intensive-care hospital settings: Transmission dynamics, persistence, and the impact of infection control programs. Proceedings of the National Academy of Sciences of the United States of America. 1999;96(12):6908-13.

121. de Celles M, Salomon J, Marinier A, Lawrence C, Gaillard J, Herrmann J, et al. Identifying More Epidemic Clones during a Hospital Outbreak of Multidrug-Resistant Acinetobacter baumannii. Plos One. 2012;7(9).

122. Hetem D, Bootsma M, Troelstra A, Bonten M. Transmissibility of Livestock-associated Methicillin-Resistant Staphylococcus aureus. Emerging Infectious Diseases. 2013;19(11):1797-802.

123. Salathé M, Kazandjieva M, Lee J, Levis P, Feldman M, Jones J. A high-resolution human contact network for infectious disease transmission. Proceedings of the National Academy of Sciences of the United States of America. 2010;107(51):22020-5.

124. Liljeros F, Edling C, Amaral L, Stanley H, Aberg Y. The web of human sexual contacts. Nature. 2001;411(6840):907-8.

125. Jones J, Handcock M. Sexual contacts and epidemic thresholds. Nature. 2003;423(6940):605-6.

126. Meyers LA, Newman MEJ, Martin M, Schrag S. Applying network theory to epidemics: control measures for *Mycoplasma pneumoniae* outbreaks. Emerging Infectious Diseases. 2003;9(2):204-10.

127. Meyers L, Pourbohloul B, Newman M, Skowronski D, Brunham R. Network theory and SARS: predicting outbreak diversity. Journal of Theoretical Biology. 2005;232(1):71-81.

128. Isella L, Romano M, Barrat A, Cattuto C, Colizza V, Van den Broeck W, et al. Close encounters in a pediatric ward: measuring face-to-face proximity and mixing patterns with wearable sensors. Plos One. 2011;6(2).

129. Huang S, Avery T, Song Y, Elkins K, Nguyen C, Nutter S, et al. Quantifying interhospital patient sharing as a mechanism for infectious disease spread. Infection Control and Hospital Epidemiology. 2010;31(11):1160-9.

130. Avitzour M, Aharonson-Daniel L, Peleg K. Secondary transfer of trauma patients: rationale and characteristics. The Israel Medical Association journal : IMAJ. 2006;8(8):539-42.

131. Kulshrestha A, Singh J. Inter-hospital and intra-hospital patient transfer: Recent concepts. Indian Journal of Anaesthesia. 2016;60(7):451-7.

132. Iwashyna TJ, Christie JD, Moody J, Kahn JM, Asch DA. The structure of critical care transfer networks. Medical care. 2009;47(7):787-93.

133. Iwashyna TJ. The incomplete infrastructure for interhospital patient transfer. Critical Care Medicine. 2012;40(8):2470-8.

134. Nuemi G, Afonso F, Roussot A, Billard L, Cottenet J, Combier E, et al. Classification of hospital pathways in the management of cancer: application to lung cancer in the region of burgundy. Cancer epidemiology. 2013;37(5):688-96.

135. Li A, Cornelius SP, Liu YY, Wang L, Barabási AL. The fundamental advantages of temporal networks. Science. 2017;358(6366):1042.

136. Barthélemy M. Spatial networks. Physics Reports. 2011;499(1–3):1-101.

137. Newman M. The Structure and Function of Complex Networks. SIAM Review. 2003;45(2):167-256.

138. Dijkstra EW. A note on two problems in connexion with graphs. Numer Math. 1959;1(1):269-71.

139. Newman MEJ. The spread of epidemic disease on networks. 2002.

140. Keeling MJ. The effects of local spatial structure on epidemiological invasions. Proceedings of the Royal Society B: Biological Sciences. 1999;266(1421):859-67.

141. Keeling MJ, Eames KT. Networks and epidemic models. J R Soc Interface. 22005. p. 295-307.

142. Sattenspiel L, Simon CP. The spread and persistence of infectious diseases in structured populations. Mathematical Biosciences. 1988;90(1):341-66.

143. Longini IM. A mathematical model for predicting the geographic spread of new infectious agents. Mathematical Biosciences. 1988;90(1):367-83.

144. Kretzschmar M, Morris M. Measures of concurrency in networks and the spread of infectious disease. Mathematical Biosciences. 1996;133(2):165-95.

145. Ball F, Mollison D, Scalia-Tomba G. Epidemics with two levels of mixing. Ann Appl Probab. 1997;7(1):46-89.

146. Riley S, Fraser C, Donnelly CA, Ghani AC, Abu-Raddad LJ, Hedley AJ, et al. Transmission Dynamics of the Etiological Agent of SARS in Hong Kong: Impact of Public Health Interventions. Science. 2003;300(5627):1961-6.

147. Schaefler S, Jones D, Perry W, Baradet T, Mayr E, Rampersad C. Methicillin-resistant Staphylococcus aureus strains in New York City hospitals: inter-hospital spread of resistant strains of type 88. Journal of clinical microbiology. 1984;20(3):536-8.

148. Nicolle LE, Dyck B, Thompson G, Roman S, Kabani A, Plourde P, et al. Regional Dissemination and Control of Epidemic Methicillin-Resistant Staphylococcus aureus. Infection Control & Hospital Epidemiology. 2015;20(3):202-5.

149. Trick WE, Kuehnert MJ, Quirk SB, Arduino MJ, Aguero SM, Carson LA, et al. Regional Dissemination of Vancomycin-Resistant Enterococci Resulting from Interfacility Transfer of Colonized Patients. The Journal of Infectious Diseases. 1999;180(2):391-6.

150. Ostrowsky BE, Trick WE, Sohn AH, Quirk SB, Holt S, Carson LA, et al. Control of Vancomycin-Resistant Enterococcus in Health Care Facilities in a Region. New England Journal of Medicine. 2001;344(19):1427-33.

151. Manikal VM, Landman D, Saurina G, Oydna E, Lal H, Quale J. Endemic Carbapenem-Resistant Acinetobacter Species in Brooklyn, New York: Citywide Prevalence, Interinstitutional Spread, and Relation to Antibiotic Usage. Clinical Infectious Diseases. 2000;31(1):101-6.

152. Smith DL, Dushoff J, Perencevich EN, Harris AD, Levin SA. Persistent colonization and the spread of antibiotic resistance in nosocomial pathogens: resistance is a regional problem. Proceedings of the National Academy of Sciences of the United States of America. 2004;101(10):3709-14.

153. Sohn AH, Ostrowsky BE, Sinkowitz-Cochran RL, Quirk SB, Jarvis WR. Evaluation of a successful vancomycin-resistant Enterococcus prevention intervention in a community of health care facilities. American Journal of Infection Control. 2001;29(1):53-7.

154. Vriens M, Blok H, Fluit A, Troelstra A, van der Werken C, Verhoef J. Costs Associated with a Strict Policy to Eradicate Methicillin-Resistant Staphylococcus aureus in a Dutch University Medical Center: A 10-Year Survey. European Journal of Clinical Microbiology and Infectious Diseases. 2002;21(11):782-6.

155. Lee BY, McGlone SM, Song Y, Avery TR, Eubank S, Chang C-C, et al. Social Network Analysis of Patient Sharing Among Hospitals in Orange County, California. American Journal of Public Health. 2011;101(4):707-13.

156. Lee BY, Wong KF, Bartsch SM, Yilmaz SL, Avery TR, Brown ST, et al. The Regional Healthcare Ecosystem Analyst (RHEA): a simulation modeling tool to assist infectious disease control in a health system. Journal of the American Medical Informatics Association. 2013;20(E1):E139-E46.

157. Lee BY, Singh A, Bartsch SM, Wong KF, Kim DS, Avery TR, et al. The potential regional impact of contact precaution use in nursing homes to control methicillin-resistant *Staphylococcus aureus*. Infection Control and Hospital Epidemiology. 2013;34(2):151-60.

158. Lee BY, Bartsch SM, Wong KF, McKinnell JA, Slayton RB, Miller LG, et al. The potential trajectory of carbapenem-resistant *Enterobacteriaceae*, an emerging threat to health-care facilities, and the impact of the centers for disease control and prevention toolkit. American Journal of Epidemiology. 2016;183(5):471-9.

159. Bartsch SM, Huang SS, Wong KF, Avery TR, Lee BY. The spread and control of norovirus outbreaks among hospitals in a region: a simulation model. Open Forum Infectious Diseases. 2014;1(2).

160. Lee BY, Yilmaz SL, Wong KF, Bartsch SM, Eubank S, Song Y, et al. Modeling the regional spread and control of vancomycin-resistant enterococci. American Journal of Infection Control. 2013;41(8):668-73.

161. Ciccolini M, Donker T, Grundmann H, Bonten MJM, Woolhouse MEJ. Efficient surveillance for healthcare-associated infections spreading between hospitals. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(6):2271-6.

162. Slayton RB, Toth D, Lee BY, Tanner W, Bartsch SM, Khader K, et al. Vital Signs: estimated effects of a coordinated approach for action to reduce antibiotic-resistant infections in health care facilities - United States. Morbidity and Mortality Weekly Report. 2015;64(30):826-31.

163. Robotham JV, Deeny SR, Fuller C, Hopkins S, Cookson B, Stone S. Cost-effectiveness of national mandatory screening of all admissions to English National Health Service hospitals for methicillin-resistant *Staphylococcus aureus*: a mathematical modelling study. Lancet Infectious Diseases. 2016;16(3):348-56.

164. Lesosky M, McGeer A, Simor A, Green K, Low DE, Raboud J. Effect of patterns of transferring patients among healthcare institutions on rates of nosocomial methicillin-resistant *Staphylococcus aureus* transmission: a Monte Carlo simulation. Infection Control and Hospital Epidemiology. 2011;32(2):136-47.

165. Lee BY, Bartsch SM, Wong KF, Singh A, Avery TR, Kim DS, et al. The importance of nursing homes in the spread of methicillin-resistant *Staphylococcus aureus* (MRSA) among hospitals. Medical Care. 2013;51(3):205-15.

166. Lee BY, Bartsch SM, Wong KF, Yilmaz SL, Avery TR, Singh A, et al. Simulation shows hospitals that cooperate on infection control obtain better results than hospitals acting alone. Health Affairs. 2012;31(10):2295-303.

167. Lee BY, McGlone SM, Wong KF, Yilmaz SL, Avery TR, Song Y, et al. Modeling the spread of methicillin-resistant *Staphylococcus aureus* (MRSA) outbreaks throughout the hospitals in Orange County, California. Infection Control and Hospital Epidemiology. 2011;32(6):562-72.

168. Karkada UH, Adamic LA, Kahn JM, Iwashyna TJ. Limiting the spread of highly resistant hospital-acquired microorganisms via critical care transfers: a simulation study. Intensive Care Medicine. 2011;37(10):1633-40.

169. Polgreen PM, Tassier TL, Pemmaraju SV, Segre AM. Prioritizing healthcare worker vaccinations on the basis of social network analysis. Infection Control and Hospital Epidemiology. 2010;31(9):893-900.

170. Cusumano-Towner M, Li DY, Tuo SS, Krishnan G, Maslove DM. A social network of hospital acquired infection built from electronic medical record data. Journal of the American Medical Informatics Association. 2013;20(3):427-34.

171. Hornbeck T, Naylor D, Segre AM, Thomas G, Herman T, Polgreen PM. Using sensor networks to study the effect of peripatetic healthcare workers on the spread of hospital-associated infections. Journal of Infectious Diseases. 2012;206(10):1549-57.

172. Ueno T, Masuda N. Controlling nosocomial infection based on structure of hospital social networks. Journal of Theoretical Biology. 2008;254(3):655-66.

173. Barnes S, Golden B, Wasil E, editors. A dynamic patient network model of hospital-acquired infections. 2010 Winter Simulation Conference; 2010 Dec 05-08; Baltimore, MD. New York: IEEE; 2010.

174. Ferrer J, Salmon M, Temime L. Nosolink: an agent-based approach to link patient flows and staff organization with the circulation of nosocomial pathogens in an intensive care unit. 2013 International Conference on Computational Science. 2013;18:1485-94.

175. Donker T, Wallinga J, Slack R, Grundmann H. Hospital networks and the dispersal of hospital-acquired pathogens by patient transfer. Plos One. 2012;7(4):8.

176. Simmering JE, Polgreen LA, Campbell DR, Cavanaugh JE, Polgreen PM. Hospital Transfer Network Structure as a Risk Factor for Clostridium difficile Infection. Infection Control and Hospital Epidemiology. 2015;36(9):1031-7.

177. Salathé M, Jones J. Dynamics and Control of Diseases in Networks with Community Structure. Plos Computational Biology. 2010;6(4).

178. Lamarsalle L, Hunt B, Schauf M, Szwarcensztein K, Valentine WJ. Evaluating the clinical and economic burden of healthcare-associated infections during hospitalization for surgery in France. Epidemiology and Infection. 2013;141(12):2473-82.

179. Fourquet F, Demont F, Lecuyer AI, Rogers MA, Bloc DH. French medical hospital information system and cross infection surveillance: theory and feasibility. Médecine Mal Infect. 2003;33(2):110-3.

180. Moulis G, Lapeyre-Mestre M, Palmaro A, Pugnet G, Montastruc JL, Sailler L. French health insurance databases: What interest for medical research? La Revue de Médecine Interne. 2015;36(6):411-7.

181. Banaei A, Bensadon M, Vuagnat A, Chodorge M. Management of Modernizing French DRG based Information System (PMSI): How Subsidiarity Principle Made It Possible. Public Health Informatics. 2005:1315:20.

182. Bouzbid S, Gicquel Q, Gerbier S, Chomarat M, Pradat E, Fabry J, et al. Automated detection of nosocomial infections: evaluation of different strategies in an intensive care unit 2000-2006. Journal of Hospital Infection. 2011;79(1):38-43.

183. Legras B, Feldmann L, Burdin J, Weber M, Hartemann P. Evaluation des infections nosocomiales à partir des données du laboratoire et des résumés d'hospitalisation. Médecine Mal Infect. 1993;23(7):307-15.

184. Gerbier-Colomban S, Bourjault M, Cetre JC, Baulieux J, Metzger MH. Evaluation study of different strategies for detecting surgical site infections using the hospital information system at Lyon University Hospital, France. Annals of surgery. 2012;255(5):896-900.

185. Nuemi G, Astruc K, Aho S, Quantin C. Comparing results of methicillin-resistant Staphylococcus aureus (MRSA) surveillance using the French DRG-based information system (PMSI). Rev Epidemiol Sante Publique. 2013;61(5):455-61.

186. Grammatico-Guillon L, Baron S, Gaborit C, Rusch E, Astagneau P. Quality Assessment of Hospital Discharge Database for Routine Surveillance of Hip and Knee Arthroplasty–Related Infections. Infection Control & Hospital Epidemiology. 2016;35(6):646-51.

187. Sahli L, Lapeyre-Mestre M, Derumeaux H, Moulis G. Positive predictive values of selected hospital discharge diagnoses to identify infections responsible for hospitalization in the French national hospital database. Pharmacoepidemiology and Drug Safety. 2016;25(7):785-9.

188. Girard D, Antoine D, Che D. Epidemiology of pulmonary tuberculosis in France. Can the hospital discharge database be a reliable source of information? Médecine et Maladies Infectieuses. 2014;44(11):509-14.

189. Bernard S, Mailles A, Stahl JP. Epidemiology of infectious encephalitis, differences between a prospective study and hospital discharge data. Epidemiology and Infection. 2012;141(11):2256-68.

190. Grammatico-Guillon L, Baron S, Rosset P, Gaborit C, Bernard L, Rusch E, et al. Surgical Site Infection After Primary Hip and Knee Arthroplasty: A Cohort Study Using a Hospital Database. Infection Control & Hospital Epidemiology. 2015;36(10):1198-207.

191. Nekkab N, Astagneau P, Temime L, Crepey P. Spread of hospital-acquired infections: A comparison of healthcare networks. PLoS Comput Biol. 2017;13(8):e1005666.

192. Clauset A. Finding local community structure in networks. Physical Review E. 2005;72(2).

193. FINESS Extraction du Fichier des établissements: Le ministère des solidarités et de la santé; [updated 4 April 2018. Available from: https://www.data.gouv.fr/fr/datasets/finess-extraction-du-fichier-des-etablissements/#_.

194. Diekmann O, Heesterbeek JAP. Mathematical epidemiology of infectious diseases: model building, analysis and interpretation. Chichester: John Wiley and Sons; 2000.

195. Hyndman RJ, Athanasopoulos G. Forecasting: principles and practice. Section 8/9 [Internet]. 2013; 2018(7 April). Available from: https://www.otexts.org/fpp/8/9.

196. Jombart T, Aanensen DM, Baguelin M, Birrell P, Cauchemez S, Camacho A, et al. OutbreakTools: a new platform for disease outbreak analysis using the R software. Epidemics. 2014;7:28-34.

197. Jombart T, Eggo RM, Dodd PJ, Balloux F. Reconstructing disease outbreaks from genetic data: a graph approach. Heredity. 2010;106(2):383-90.

198. Jombart T, Cori A, Didelot X, Cauchemez S, Fraser C, Ferguson N. Bayesian Reconstruction of Disease Outbreaks by Combining Epidemiologic and Genomic Data. PLOS Computational Biology. 2014;10(1):e1003457.

199. Obadia T, Silhol R, Opatowski L, Temime L, Legrand J, Thiebaut ACM, et al. Detailed contact data and the dissemination of *Staphylococcus aureus* in hospitals. Plos Computational Biology. 2015;11(3):16.

200. Schumacher M, Allignol A, Beyersmann J, Binder N, Wolkewitz M. Hospital-acquired infections—appropriate statistical treatment is urgently needed! International Journal of Epidemiology. 2013;42(5):1502-8.

201. Vanhems P, Barrat A, Cattuto C, Pinton J-F, Khanafer N, Régis C, et al. Estimating Potential Infection Transmission Routes in Hospital Wards Using Wearable Proximity Sensors. PLOS ONE. 2013;8(9):e73970.

202. Chambers D, Wilson P, Thompson C, Harden M. Social Network Analysis in Healthcare Settings: A Systematic Scoping Review. PLoS ONE. 2012;7(8):e41911.

203. Gingras G, Guertin MH, Laprise JF, Drolet M, Brisson M. Mathematical modeling of the transmission dynamics of *Clostridium difficile* infection and colonization in healthcare settings: a systematic review. PLoS One. 2016;11(9).

204. Nishiura H, Kuratsuji T, Quy T, Phi NC, Van Ban V, Ha LD, et al. Rapid awareness and transmission of severe acute respiratory syndrome in Hanoi French Hospital, Vietnam. American Journal of Tropical Medicine and Hygiene. 2005;73(1):17-25.

205. Masuda N, Konno N, Aihara K. Transmission of severe acute respiratory syndrome in dynamical small-world networks. Physical Review E. 2004;69(3).

206. Webb GF, Blaser MJ, Zhu HP, Ardal S, Wu JH. Critical role of nosocomial transmission in the Toronto SARS outbreak. Mathematical Biosciences and Engineering. 2004;1(1):1-13.

207. Lloyd-Smith JO, Galvani AP, Getz WM. Curtailing transmission of severe acute respiratory syndrome within a community and its hospital. Proceedings of the Royal Society B-Biological Sciences. 2003;270(1528):1979-89.

208. Camacho A, Kucharski AJ, Funk S, Breman J, Piot P, Edmunds WJ. Potential for large outbreaks of Ebola virus disease. Epidemics. 2014;9:70-8.

209. Barbarossa MV, Denes A, Kiss G, Nakata Y, Rost G, Vizi Z. Transmission dynamics and final epidemic size of Ebola virus disease outbreaks with varying interventions. Plos One. 2015;10(7):21.

210. Dong FL, Xu DL, Wang Z, Dong MW. Evaluation of Ebola spreading in West Africa and decision of optimal medicine delivery strategies based on mathematical models. Infection Genetics and Evolution. 2015;36:35-40.

211. Vanhems P, Von Raesfeldt R, Ecochard R, Voirin N. Emergence of Ebola virus disease in a french acute care setting: a simulation study based on documented inter-individual contacts. Scientific Reports. 2016;6.

212. O'Neill J. Antimicrobial resistance: tackling a crisis for the health and wealth of nations. London: Review on Antimicrobial Resistance 2014.

213. Chase-Topping CLG, Bram ADvB, Oliver B, Chris R, Thibaud P, Laura I, et al. Not just a matter of size: a hospital-level risk factor analysis of MRSA bacteraemia in Scotland. BMC Infectious Diseases. 2016;16(1):222.

214. Réseau d'alerte d'investigation et de surveillance des infections nosocomiales (Raisin). Enquête nationale de prévalence des infections nosocomiales et des traitements anti-infectieux en établissements de santé, France, mai-juin 2012. Résultats. Saint-Maurice: Institut de veille sanitaire; 2013.

215. Donker T, Smieszek T, Henderson KL, Johnson AP, Walker AS, Robotham JV. Measuring distance through dense weighted networks: The case of hospital-associated pathogens. PLOS Computational Biology. 2017;13(8):e1005622.

216. Donker T, Henderson KL, Hopkins KL, Dodgson AR, Thomas S, Crook DW, et al. The relative importance of large problems far away versus small problems closer to home: insights into limiting the spread of antimicrobial resistance in England. BMC Medicine. 2017;15(1):86.

217. Saltman RB, Bankauskaite V, Vrangbaek K. Decentralization in health care. Strategies and outcomes. New York: Open University Press; 2007. 149 p.

218. Polton D. Décentralisation des systèmes de santé: Quelques réflexions à partir d'expériences étrangères. Questions d'économie de la santé. 2003.

219. Lepelletier D, Berthelot P, Lucet JC, Fournier S, Jarlier V, Grandbastien B, et al. French recommendations for the prevention of 'emerging extensively drug-resistant bacteria' (eXDR) cross-transmission. Journal of Hospital Infection. 2015;90(3):186-95.

Narimane NEKKAB

SPREAD OF PATHOGENS IN HEALTHCARE NETWORKS: ASSESSMENT OF THE ROLE OF INTER-FACILITY PATIENT TRANSFERS ON INFECTION RISKS AND CONTROL MEASURES

Résumé

La propagation des pathogènes, notamment des bactéries multi-résistantes, au sein du réseau des hôpitaux, est un grand enjeu de santé publique. L'évaluation du rôle joué par les transferts inter-établissements des patients sur cette propagation pourrait permettre l'élaboration de nouvelles mesures de contrôle. L'utilisation des données de réseaux de contact inter-individus et de transferts inter-établissement dans la modélisation mathématique ont rendu ces modèles plus proches de la réalité. Toutefois, ces derniers restent limités à quelques milieux hospitaliers et quelques pathogènes. La thèse a eu pour objectifs de 1) mieux comprendre la structure des réseaux hospitaliers français et leur impact sur la propagation des pathogènes dans le milieu hospitalier ; et 2) évaluer le rôle des transferts sur la propagation des entérobactéries productrices de carbapenemase (EPC). Les réseaux hospitaliers français sont caractérisés par des flux de patients vers des hubs et par deux niveaux de communautés des hôpitaux. La structure du réseau de transfert des patients présentant une infection nosocomiale (IN) n'est pas différente de celle du réseau général de transfert des patients. Ce travail a également montré que, depuis 2012, les transferts de patients jouent avec les années un rôle de plus en plus important sur la diffusion des EPC en France. En conséquence, la structure du réseau des hôpitaux pourrait servir de base pour la proposition des nouvelles stratégies de contrôles des IN en général, et des EPC en particulier.

Infections nosocomiales ; réseaux hospitaliers ; entérobactéries

Résumé en anglais

The spread of pathogens and multi-drug resistance in healthcare networks is a major public health issue. Evaluating the role of inter-facility patient transfers may provide insights on novel infection control measures. The increasing use of inter-individual contact and inter-facility transfer network data in mathematical modelling of pathogen spread in healthcare settings has helped these models become more realistic; however, they remain limited to a few settings and pathogens. The main objectives of this thesis were two-fold: 1) to better understand the structure of the healthcare networks of France and their impact on pathogen spread dynamics; and 2) to assess the role of transfers on the spread of Carbapenemase-producing *Enterobacteriaceae* (CPE). The French healthcare networks are characterized by centralized patient flows towards hubs hospitals and a two-tier community clustering structure. We also found that networks of patients with healthcare-associated infections (HAIs) form the same underlying structure as that of the general patient population. The general patient network was used to show that, since 2012, patient transfers have played an increasingly important role over time in the spread of CPE in France. Therefore, the structure of healthcare networks may help serve as a basis for novel infection control strategies to tackle HAIs in general, and CPE in particular.

Hospital-acquired infections; healthcare networks; *Enterobacteriaceae*