UNIVERSITE PARIS-EST CRETEIL

École doctorale Sciences de la vie et de la santé (ED402 SVS)

THÈSE

Présentée en vue d'obtenir le

DOCTORAT EN PATHOLOGIE ET RECHERCHE CLINIQUE
SPECIALITE SANTE PUBLIQUE, ENVIRONNEMENT ET SOCIETE

Soutenue le 16/12/2025

par

Antoine GAILLET

Utilisation de l'Entrepôt des Données de Santé de l'Assistance Publique – Hôpitaux de Paris pour la conduite d'études épidémiologiques dans le domaine des soins critiques

Thèse dirigée par le Pr Nicolas de Prost

Co-dirigée par le Pr Armand Mekontso Dessap

Membres du jury :

Pr Nicolas de Prost	Université Paris-Est Créteil	Directeur
Pr Armand Mekontso Dessap	Université Paris-Est Créteil	Co-directeur
Pr Marie-France Mamzer	Université Paris Cité	Présidente du jury
Dr Nadia Oubaya	Université Paris-Est Créteil	Examinatrice
Pr Tai Pham	Université Paris-Saclay	Examinateur
Pr Lara Zafrani	Université Paris Cité	Rapporteuse
Pr Guillaume Dumas	Université Grenoble Alpes	Rapporteur

Remerciements

Je remercie tout particulièrement le Pr Nicolas de Prost pour son accompagnement depuis le début et de tous les instants, rigoureux et sympathique. Merci au Pr Armand Mekontso Dessap pour sa co-direction de la thèse. Merci aux membres des services de médecine intensive des hôpitaux Henri Mondor et Bichat-Claude Bernard pour leur aide, de prêt ou de loin, à la réalisation de ce projet. Merci au Dr Richard Layese pour ses lumières plus que précieuses dans le domaine parfois opaque des biostatistiques et du codage. Merci au Pr Romain Sonneville pour le suivi pertinent qu'il a apporté à ce travail au cours des quatre dernières années. Mes remerciements vont également à l'ensemble des membres du jury, qui me font l'honneur de juger ce travail. Veuillez trouver ici l'expression de mon profond respect.

Enfin, un profond merci à ma famille, tolérante et aimante.

Résumé

Introduction : L'épidémiologie actuelle est le fruit d'une lente évolution scientifique au cours du temps, qui a vu sa mutation s'accélérer ces dernières décennies à l'occasion de l'apparition des données massives et de l'interconnexion numérique. L'entrepôt des données de santé (EDS) de l'assistance publique des hôpitaux de Paris (AP-HP) en est l'illustration la plus contemporaine. De par la multiplicité de ses données, sa taille, et son accessibilité théorique, elle offre une source d'information précieuse et quasi inépuisable. Le but de ce travail de thèse était d'évaluer les capacités et les limites de l'EDS de l'AP-HP à soutenir la recherche clinique dans des contextes pathologiques contrastés en contexte de soins critiques.

Méthode : Deux projets de recherche clinique ayant des méthodologies distinctes ont été menés à l'aide de l'EDS AP-HP, de 2017 à 2023 : le premier, ciblant une pathologie fréquente et bien codée ; le deuxième, une affection rare et non codée.

Résultats : Le premier travail a notablement permis d'établir trois phénotypes distincts de patients infectés par le virus respiratoire syncytial (VRS) et hospitalisés en soins critiques avec l'établissement d'une cohorte de 474 patients. Le second a permis de mettre en évidence, au sein d'une cohorte de 235 patients, des facteurs de vulnérabilité de patients hospitalisés pour un *Drug Reaction with Eosinophilia and Systemic Symptoms* (DRESS). Dans les deux cas, on a eu recours aux codages structurés de l'EDS (incluant celui de la Classification Statistique Internationale des Maladies (CIM-10) et la classification commune des actes médicaux (CCAM)). De la recherche textuelle a également été employée pour les deux études, avec cependant des moyens et une finalité différents (extraction de données descriptives pour l'étude VRS (co-infection, thérapeutiques), identification des patients pour l'étude DRESS).

Discussion : L'EDS a permis de traiter des questions épidémiologiques très différentes en mettant à profit différents types de données (structurées et non structurées). Néanmoins, le faible volume d'études publiées (n=108, correspondant à 21% des projets déposés) contraste avec la richesse et le potentiel de cet entrepôt, le plus volumineux d'Europe. Des difficultés d'exploitation de l'EDS sont probablement le fait de difficultés techniques engendrées par la centralisation (via JupyterHub, pourtant gage de sécurité) et le caractère non structuré des données (pourtant gage de sa richesse).

Conclusion : L'EDS de l'AP-HP permet de traiter des situations cliniques diverses. Sa pleine exploitation est potentiellement encore bridée du fait de contraintes techniques. La simplification de ces dernières constituera le principal défi d'avenir permettant un plus large usage de l'EDS de l'AP-HP.

Abstract

Introduction: Modern epidemiology is the result of a gradual scientific evolution that has accelerated over the past decades with the emergence of big data and the digital interconnection of healthcare systems. The Entrepôt des Données de Santé (EDS, Health Data Warehouse) of the Assistance Publique–Hôpitaux de Paris (AP-HP) represents the most contemporary illustration of this transformation within French medical research. Owing to the multiplicity, scale, and theoretical accessibility of its data, the EDS constitutes a valuable and nearly inexhaustible source of information. The aim of this doctoral work was to assess the capabilities and limitations of the AP-HP EDS in supporting clinical research across contrasting pathological contexts within critical care settings.

Methods: To this end, two distinct clinical research projects were conducted across the 39 hospitals of the AP-HP network between 2017 and 2023: (1) the first focused on a frequent and well-coded condition, and (2) the second on a rare and non-coded one.

Results: The first study established three distinct phenotypes of patients infected with respiratory syncytial virus (RSV) in the intensive care context, based on a cohort of 474 patients. The second identified vulnerability factors among 235 hospitalized patients diagnosed with Drug Reaction with Eosinophilia and Systemic Symptoms (DRESS).

Both studies relied on structured data from the EDS — including the International Classification of Diseases, 10th Revision (ICD-10) and the Common Classification of Medical Procedures (CCAM) — as well as unstructured textual data. Text mining was used in both studies but with differing goals: descriptive data extraction (e.g., co-infections, treatments) in the RSV study, and patient identification in the DRESS study.

Discussion: The EDS enabled the exploration of very different epidemiological questions by leveraging both its structured and unstructured data. However, the relatively low volume (n=108) and publication rate (21%) of EDS-based projects contrasts with the richness and potential of the database. This discrepancy likely reflects technical challenges inherent to its centralized architecture (via JupyterHub, which ensures data security) and the complexity of handling unstructured data, despite their analytical value.

Conclusion: The AP-HP Health Data Warehouse holds tremendous potential for investigating a wide range of clinical scenarios. Yet, its full exploitation remains limited by current technical constraints. Simplifying data access and analytical workflows will represent a key future challenge to unlock the full capacity of the EDS for conducting large-scale clinical and epidemiological research.

Liste des abréviations

AP-HP : Assistance publique – Hôpitaux de Paris

CDW : Clinical Data Warehouse

CIM-10 : International Statistical Classification of Diseases and Related Health Problems 10th Revision.

CNIL : Commission nationale de l'informatique et des libertés

CSE : Comité Scientifique et Éthique

DME : Dossiers médicaux électroniques

DPI : Dossier Patient Informatisé

DRESS : Drug Reaction with Eosinophilia and Systemic Symptoms

EDC : Entrepôts de données cliniques

EDS : entrepôts des données de santé

ETL : Extract – Transform – Load

FHIR : Fast Healthcare Interoperability Resources

I2B2 : Informatics for Integrating Biology and the Bedside

IA : intelligence artificielle

IRM : imagerie par résonance magnétique

MIMIC : Medical Information Mart for Intensive Care

NLP : Natural language processing

OHDSI : Observational Health Data Sciences and Informatics

OMOP : Observational Medical Outcomes Partnership

PACS : Picture archiving and communication system

PMSI : Programme de Médicalisation des Systèmes d'Information

RGPD : Règlement général sur la protection des données

SNDS : Système National des Données de Santé

SQL : Structured Query Language

VRS : virus respiratoire syncytial

Table des matières

I. Introduction 10
A. Genèse de l'épidémiologie 10
B. Epidémiologie moderne 12
1. Révolution statistique 12
2. Révolution méthodologique 13
C. Epidémiologie post-moderne : les données massives 14
1. Définition 15
2. Applications 15
3. Entrepôt des données des santé 18
II. Problématique 21
III. Objectif 23
IV. Résultats 24
A. Article 1 24
B. Article 2 47
V. Discussion 87
A. Revue de la littérature 87
B. Accessibilité des données 99
C. Sécurité des données 100
D. Recherche textuelle 101
E. Qualité des données 103
VI. Conclusion 105
VII. Bibliographie 107

I. Introduction

Littéralement, « l'étude du peuple », l'épidémiologie est une discipline qui étudie la fréquence, la distribution et les déterminants des problèmes de santé dans les populations. Elle est aujourd'hui un pilier de la santé publique et de la médecine préventive. Relativement récente, elle nait en Angleterre au XVIIe siècle. Initialement centrée sur les épidémies et les maladies transmissibles, elle s'est élargie à l'étude des facteurs de risque des maladies chroniques et complexes, grâce à une méthodologie et des analyses statistiques de plus en plus élaborées. Aujourd'hui, l'émergence des données massives ouvre de nouvelles perspectives pour approfondir la compréhension des déterminants de santé mais expose à de nouveaux enjeux de qualité et de sécurité de la donnée inhérents à la gestion de base de données aussi grandes.

A. Genèse de l'épidémiologie

Dans la médecine occidentale, Hippocrate compte parmi les premiers à avoir établi un lien entre l'apparition des maladies, les facteurs environnementaux et les modes de vie, en particulier dans son traité *Des airs, des eaux et des lieux* (vers 400 av. J.-C.) [1]. Bien plus tard, à la Renaissance, l'Italien Girolamo Fracastoro franchit une étape décisive avec son ouvrage *De contagione et contagiosis morbis* (1546), où il propose une explication scientifique des maladies contagieuses fondée sur l'étude de la transmission de la syphilis, alors largement répandue en Europe et au-delà.

Au XVIIe siècle, l'Angleterre voit apparaître les premiers jalons de l'épidémiologie contemporaine. À partir de 1538, les *Bills of mortality* recensent chaque décès, et dès 1603, l'enregistrement systématique des causes de mortalité permet la diffusion hebdomadaire de statistiques en temps réel. En 1662, John Graunt exploite ces données pour étudier les

variations de mortalité liées aux épidémies de peste, conférant à ces registres administratifs une fonction de véritable outil de surveillance sanitaire. Il conçoit également la première table de mortalité, destinée à estimer la probabilité de survie selon l'âge, ouvrant ainsi la voie aux calculs d'espérance de vie réalisés par les mathématiciens contemporains. Ces avancées constituent les fondements communs de l'épidémiologie et de la démographie.

Au XVIIIe siècle, Daniel Bernoulli, médecin et mathématicien suisse, inaugure une approche innovante en posant les bases de l'épidémiologie mathématique. Dans un essai publié en 1760 sur la mortalité liée à la variole et les bénéfices potentiels de l'inoculation, il élabore un modèle destiné à évaluer le rapport bénéfice/risque de cette pratique alors controversée. Il démontre ainsi l'intérêt de la modélisation mathématique pour orienter les choix de santé publique. Ses travaux suscitent néanmoins des objections, concernant notamment la qualité des données utilisées pour valider les modèles, soulignant déjà l'intérêt porté à la robustesse des systèmes d'information.

Au XIXe siècle, débute véritablement la santé publique fondée sur les preuves. Pierre-Charles-Alexandre Louis introduit la *Méthode numérique*, fondée sur la comparaison systématique des résultats thérapeutiques afin d'établir des preuves objectives. Par ses travaux sur les infections respiratoires, il démontre la faible efficacité des saignées, imposant l'autorité des faits face aux dogmes médicaux [2]. Son approche, prolongée par Jules Gavarret, jette les bases de la statistique médicale en introduisant les notions d'échantillon, de comparaison intergroupe et d'incertitude [3]. Parallèlement, William Farr systématise l'exploitation des registres de mortalité pour identifier les tendances sanitaires et les facteurs de risque, fondant l'épidémiologie descriptive et analytique. Enfin, John Snow illustre magistralement cette évolution : en cartographiant les cas de choléra à Londres en 1854, il

identifie une pompe à eau contaminée comme source de l'épidémie, démontrant la puissance de la collecte et de l'analyse spatiale des données dans la compréhension des maladies infectieuses.

À la fin du XIXe siècle, l'essor de la microbiologie (Henle, Koch, Pasteur), puis les progrès de l'hygiène, de la nutrition, de la vaccination et de l'antibiothérapie au début du XXe siècle, entraînent une diminution spectaculaire de la mortalité par maladies infectieuses dans les pays occidentaux (de 797 à 36 décès pour 100 000 habitants entre 1900 et 1980) [4]. Cette transition marque le passage de « l'âge de la pestilence et de la famine » à celui des « maladies dégénératives et provoquées par l'homme » [5]. Ainsi, si l'épidémiologie reposait déjà sur des bases solides, avec ses volets descriptif, expérimental et théorique adaptés aux maladies aiguës et infectieuses, le défi du XXe siècle fut de développer une épidémiologie des maladies chroniques, intégrant l'identification des facteurs de risque génétiques et environnementaux. C'est dans ce contexte qu'émerge véritablement l'épidémiologie moderne.

B. Epidémiologie moderne

1. Révolution statistique

Dans le sillage de l'épidémiologie quantitative du XIXe siècle, appuyée sur les progrès de la statistique et de la théorie des probabilités amorcés par Laplace, s'est développée au début du XXe siècle une épidémiologie mathématique marquée par une grande inventivité méthodologique permettant d'évaluer rigoureusement l'efficacité des traitements. Des figures telles que Ronald A. Fisher, qui introduit les tests d'hypothèse, l'analyse de variance et les plans expérimentaux randomisés, ou Jerzy Neyman et Egon Pearson, à l'origine de la théorie des tests d'hypothèse formalisée, ont permis d'évaluer rigoureusement l'efficacité des traitements

[6]. À cela s'ajoutent Karl Pearson, pionnier de la statistique biométrique et des méthodes de corrélation et de régression, et William Gosset, auteur du test t (Student) pour la comparaison de moyennes sur petits échantillons.

2. Révolution méthodologique

Au XXe siècle, l'essor des méthodes statistiques a permis la mise en place de plans d'étude plus élaborés, fondés sur la comparaison de groupes de patients et l'identification des facteurs de risque. Les compagnies d'assurance-vie nord-américaines ont joué un rôle moteur en constituant de vastes cohortes dès les années 1920-1930, en croisant données cliniques, démographiques et d'état civil pour quantifier les risques de handicap et de mortalité prématurée. Ce cadre a favorisé le développement des études prospectives de cohorte, particulièrement adaptées à l'analyse étiologique et au calcul de risques relatifs. La cohorte la plus emblématique, financée et coordonnée à l'époque par le *National Heart Institute*, est la *Framingham Heart Study* (1947), dont les résultats ont profondément marqué la compréhension des maladies cardiovasculaires et, plus largement, des pathologies chroniques, tout en illustrant les apports des analyses multivariées dans l'étude de la multifactorialité. Dans son sillage, d'autres cohortes majeures – *Nurses' Health Study*, EPIC – ont élargi le champ à la nutrition, au cancer et aux maladies métaboliques. Parallèlement, les études cas-témoins ont apporté des preuves décisives, notamment grâce aux travaux pionniers de Bradford Hill et Richard Doll établissant le lien entre tabagisme et cancer du poumon [7]. À la fin des années 1950, David L. Sackett a introduit une définition structurée du biais comme déviation systématique des résultats ou des inférences par rapport à la réalité. En identifiant 35 formes de biais dans les enquêtes observationnelles, il a mis en évidence leur impact sur la validité des conclusions épidémiologiques [8]. Cette conceptualisation a conduit

à l'établissement d'une hiérarchie des plans d'étude et des niveaux de preuve en pratique clinique [9], tout en soulignant la nécessité de prévenir les biais majeurs, qu'ils soient liés à la sélection, à l'information ou à la confusion. A partir de 1954, avec le premier essai d'efficacité du vaccin de la poliomyélite [10], les essais randomisés contrôlés se sont développés comme la référence méthodologique pour évaluer les interventions.

L'épidémiologie ne s'est constituée en discipline académique qu'au cours de la deuxième partie du XXe siècle, sous l'impulsion notamment d'Alvan Feinstein (Yale University), qui a formalisé l'application stricte des méthodes épidémiologiques à la pratique médicale. La création d'unités universitaires dédiées a favorisé l'intégration de la recherche, des soins et de l'enseignement. À partir des années 1990, le concept de médecine fondée sur les preuves, popularisé par David Sackett s'est largement diffusé. Parallèlement, le développement des registres de maladies, des grandes cohortes et des bases de données administratives concomitamment au développement numérique a abouti à la création des bases de données massives, faisant entrer l'épidémiologie dans une nouvelle ère.

C. Epidémiologie post-moderne : les données massives

L'essor des technologies numériques a entraîné une explosion des données de santé, issues des dossiers médicaux électroniques, de l'imagerie, de la génomique ou encore des objets connectés. L'analyse de ces ensembles massifs repose désormais sur les méthodes de *data science*, d'intelligence artificielle (IA) et d'apprentissage automatique, ouvrant la voie à la médecine de précision, qui intègre données cliniques, biologiques et environnementales pour adapter les soins à chaque individu. Toutefois, cette évolution s'accompagne de défis majeurs : garantir la qualité et la validité des informations collectées, protéger la confidentialité et les

droits des patients, limiter les biais algorithmiques et, enfin, traduire ces analyses complexes en décisions cliniques pertinentes.

1. Définition

Le concept de « données massives », ou « *Big Data* », désigne des ensembles d'informations caractérisés par leur volume, leur vélocité et leur variété, dont la gestion et l'analyse dépassent les capacités des outils informatiques traditionnels [11]. Si l'expression « *Big Data* » apparaît pour la première fois en 1997 dans un article de l'*Association for Computing Machinery*, son essor s'inscrit dans une évolution progressive : émergence des premières bases de données dans les années 1960-1970 (IBM), généralisation des bases relationnelles et du langage SQL (*Structured Query Language*) dans les années 1980, puis explosion de la production d'informations avec l'essor d'Internet dans les années 1990. Les années 2000 marquent une étape décisive avec la numérisation massive des échanges, l'avènement des réseaux sociaux, des smartphones et des objets connectés. A titre d'illustration, entre 1986 et 2007, la quantité de données stockées au niveau mondial a été multipliée par 100, avec une part issue du numérique passé de 0.7% à 94% [12].

2. Applications

Initialement, les deux domaines médicaux d'application des données massives étaient la génétique et l'imagerie [13]. La génomique, considérée comme la « discipline reine » des données massives, illustre en effet l'explosion des capacités de collecte et d'analyse de données, notamment, via le séquençage complet du génome humain (terminé en 2003). Là où l'analyse se limitait autrefois à quelques gènes chez un petit nombre de patients, elle porte désormais sur l'ensemble du génome de milliers d'individus, enrichi de données cliniques et démographiques. L'exploitation de ces informations massives, stockées dans des bases

numériques accessibles via Internet, constitue le principal défi actuel de la recherche biomédicale. Des ressources comme la *UK Biobank* ou le *1000 Genomes Project* illustrent cette capacité à intégrer données génétiques, cliniques et environnementales à grande échelle. L'imagerie constitue une autre discipline majeure de la donnée massive, qu'il s'agisse de l'imagerie médicale ou de la microscopie, générant tous deux des flux impossibles à traiter manuellement. L'exploitation de ces données massives repose donc sur des outils informatiques spécifiques, avec un recours croissant à l'IA, qui permet d'analyser et d'interpréter des images à grande échelle grâce à l'augmentation simultanée de la puissance de calcul et des volumes de données disponibles. Parmi les ressources de référence figurent des bases d'imagerie médicale telles que *UK Biobank Imaging* ou le *NIH Chest X-ray dataset*.

En parallèle, la recherche clinique s'est également développée via l'exploration de diverses données massives. Les registres spécialisés (cancer, transplantation d'organe) offrent des données standardisées et longitudinales sur des maladies ou conditions particulières, permettant d'étudier l'évolution des pathologies et l'efficacité des traitements, notamment pour les maladies rares, mais leur couverture géographique est souvent limitée et leur maintenance coûteuse. Les bases médico-administratives nationales (Système National des Données de Santé (SNDS) pour la France, *Medicare & Medicaid Claims* pour les Etats-Unis) sont exhaustives et couvrent de larges populations, fournissant des informations sur les hospitalisations, les prescriptions et les diagnostics, pour un faible coût de collecte mais elles manquent souvent de détails cliniques (absence d'accès aux comptes-rendus de séjour) et peuvent comporter des biais de codage ou des imprécisions temporelles. Plus récemment, la mise à disposition, parfois publique, des dossiers médicaux électroniques (EPIC, Cerner), constitués de données multidimensionnelles (cliniques, biologiques, imagerie, prescriptions), laisse entrevoir la possibilité de bases hybrides, associant l'accès des données de bases

médico-administratives (codages, actes) et des données plus cliniques de registres (notes médicales, biologie, imagerie) [14].

Dans le domaine particulier des soins critiques, plusieurs bases de données ouvertes ont été développées et sont à l'origine de nombreuses publications [15, 16]. La principale, MIMIC (*Medical Information Mart for Intensive Care*), est américaine. Elle a vu le jour en 2001 à la suite d'un partenariat entre le *Massachusetts Institute of Technology* et le *Beth Israel Deaconess Medical Center* et compte plus de 70 000 séjours en réanimation adulte. Les données y sont riches (démographie, scores de sévérité signes vitaux, biologie, traitements, procédures, imagerie, notes cliniques et évolution), très granulaires (notamment en termes de relevé des paramètres vitaux), longitudinales, et en libre accès après formation éthique via la plateforme *PhysioNet*. De 2004 à 2024, cette base a été à l'origine d'environ 3000 articles scientifiques, essentiellement des études de cohorte observationnelles notamment à l'origine du développement de modèle de *machine learning* à partir des données de monitoring physiologique. A titre de comparaison, la base *eICU* (également américaine et accessible sur la même plateforme) offre, elle, une meilleure représentativité du fait de son caractère multicentrique, avec plus de 200 000 séjours au sein de 200 hôpitaux, mais ne permet pas d'étude longitudinales car restreinte à la période 2014–2015, et fournit des données moins granulaires que MIMIC. L'équivalent européen de MIMIC est la base *Amsterdam University Medical Center database* (AmsterdamUMCdb), qui, bien que plus petite (23 000 admissions), offre des données de qualité similaire. En comparaison aux études traditionnelles en soins intensifs utilisant des données hospitalières locales, les productions scientifiques issue des bases de données ouvertes permettent de générer des cohortes de plus grande taille et de développer une large variété de modèles de *machine learning* appliqués. Les bases de données reposant sur des cohortes classiques, incluant un volume de patients plus faible, collectent

cependant généralement des données plus détaillées [17]. D'un point de vue purement numérique, les études sur bases de données ouvertes ont un impact scientifiquement plus élevé avec en moyenne 40% de citations supplémentaire par article. Néanmoins, ces études souffrent d'une grande hétérogénéité, d'un manque de standardisation et d'un contrôle qualité imparfait.

3. Entrepôt des données des santé

C'est dans ce contexte que les entrepôts des données de santé (EDS) se sont développés au sein des centres hospitalo-universitaires français dans le courant des années 2010 [18, 19]. Chaque EDS constitue ainsi une infrastructure centralisée destinée à agréger, structurer et harmoniser des informations médicales issues de multiples systèmes hospitaliers ou cliniques. La finalité d'un EDS est double : contribuer à l'amélioration continue de la qualité des soins et soutenir la recherche clinique. Contrairement aux systèmes opérationnels tels que les dossiers patients, les prescriptions ou les bases de laboratoire, les EDS sont conçus comme des bases multidimensionnelles (Table 1) optimisées pour l'exploration et l'analyse de données complexes.

Table 1 : détails des informations réunies au sein des entrepôts des données de santé

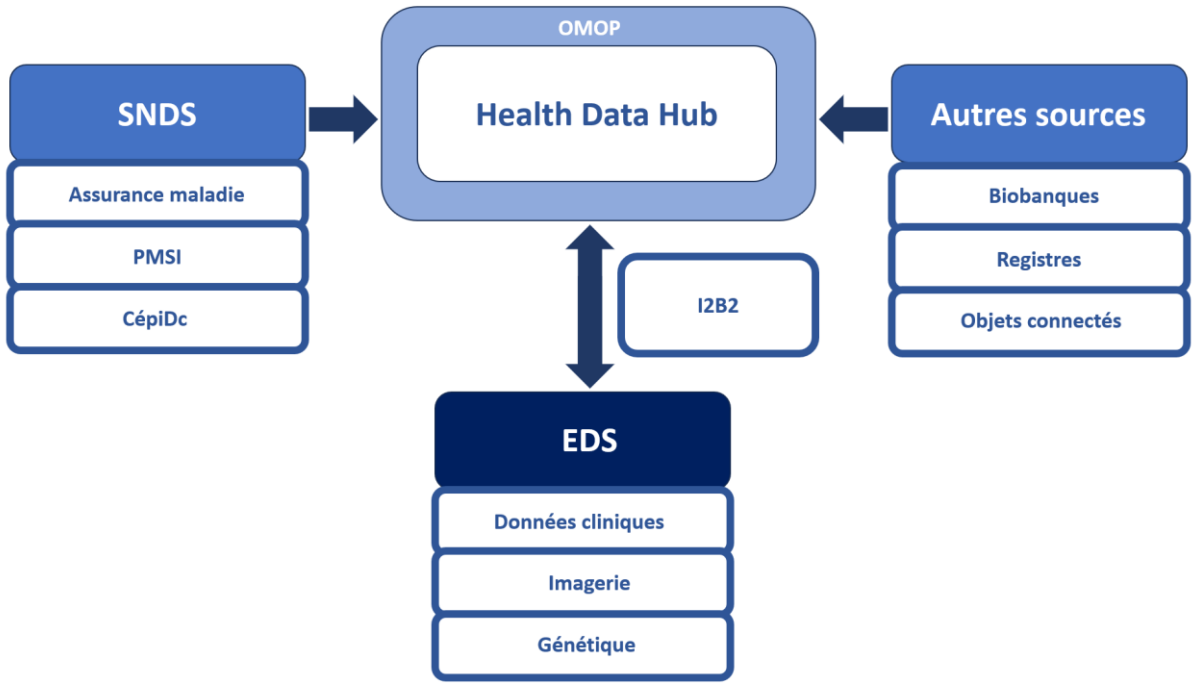
Source	**Type de données**
Dossier Patient Informatisé (DPI)	Données cliniques : diagnostics, antécédents, notes médicales
Systèmes de laboratoire	Résultats biologiques, hématologie, microbiologie, biochimie
Pharmacie hospitalière	Prescriptions, dispensation, administration de médicaments
Imagerie médicale	Radiologie, scanner, IRM, échographie, PACS
Monitoring / dispositifs médicaux	Signes vitaux, ECG, ventilation, perfusions
Données administratives	Séjours, actes, remboursements, codage PMSI
Données génomiques / biomarqueurs	Séquençage, biobanques associées
Capteurs / objets connectés	Glucomètres, tensiomètres, trackers, applications mobiles santé

L'architecture d'un EDS repose sur un modèle de type ETL (*Extract – Transform – Load*). La première étape consiste à extraire les données brutes depuis les systèmes sources. Celles-ci sont ensuite transformées par un processus de nettoyage incluant la standardisation des formats, l'harmonisation des codages (tels que le CIM-10), la pseudonymisation ou dé-identification à des fins de recherche, ainsi que la vérification de cohérence et la gestion des valeurs manquantes. Enfin, les données traitées sont chargées dans l'entrepôt central, organisé de manière à optimiser les requêtes et les analyses.

La gestion de ces données massives, qui sont souvent fragmentées et hétérogènes de par leur nature (génomique, physiologique, biologique, clinique, sociale), leur format (texte, valeurs numériques, signaux, images 2D et 3D, séquences génomiques) ou leur dispersion au sein de plusieurs systèmes d'information (établissements de soin, laboratoires de recherche, bases publiques), nécessite de s'appuyer sur différents cadres techniques et organisationnels (Figure 1) : la plateforme i2b2 (*Informatics for Integrating Biology and the Bedside*) constitue une solution logicielle largement utilisée pour construire et interroger des EDS hospitaliers locaux grâce à une interface graphique intuitive permettant la constitution de cohortes et l'exploration des données sans expertise avancée en codage et avec la possibilité d'ajouter des modules complémentaires, comme de la recherche textuelle pour l'analyse de textes médicaux [20, 21]; le modèle OMOP (*Observational Medical Outcomes Partnership*), issu du consortium OHDSI (*Observational Health Data Sciences and Informatics*), fournit une structure normalisée et des vocabulaires communs permettant d'harmoniser ces entrepôts et de rendre comparables leurs données au niveau international [22] ; enfin, le *Health Data Hub* représente une déclinaison nationale centralisée, visant à fédérer les EDS hospitaliers, les bases médico-administratives (SNDS) et les registres français dans une infrastructure sécurisée respectant les exigences légales et éthiques françaises et européennes (RGPD, CNIL), mettant à

disposition des jeux de données pseudonymisées et standardisées, afin de faciliter la recherche en santé publique [23]. Ainsi, i2b2, OMOP et le *Health Data Hub* se situent à des niveaux complémentaires : outil local, standardisation internationale et plateforme nationale intégrative.

Figure 1 : encadrement technique et organisationnel de l'EDS



Il existe ainsi des EDS dans la plupart des CHU français mais celui de l'AP-HP est le plus ancien (2017) et concerne le plus grand nombre de patients (19 millions depuis son autorisation). Avec 258 projets de recherche approuvés par le comité scientifique et éthique de l'AP-HP depuis sa création (et jusqu'à 2025 inclus), ayant abouti à 108 publications scientifiques et 32 tableaux de bord développés et mis à disposition de la communauté hospitalière, c'est également l'EDS le plus développé à ce jour. Il fera ainsi l'objet du travail de thèse ci-présent.

II. Problématique

Les bases de données habituelles présentent chacune des qualités mais également des défauts qui en limitent leur exploitation. Classiquement :

- Les recueils rétrospectifs sont à l'origine d'études cas-témoin (ex : études sur l'exposition au tabac et le risque de cancer du poumon [7]) ou de cohorte. Ils sont faciles à mettre en place, peu chers et se prêtent respectivement à l'étude des expositions et des maladies rares. En revanche, ces études sont sujettes aux biais de sélection et de mémoire, rendant difficile l'établissement d'un lien de causalité entre la maladie étudiée et les facteurs associés.
- Les cohortes prospectives sont moins biaisées et permettent d'établir des séquences temporelles approchant la causalité, mais leur coût et le temps nécessaire à leur constitution et à leur suivi limitent leur développement. Elles sont par ailleurs peu adaptées à l'étude des maladies rares.
- Les cohortes prospectives adossées à la constitution d'une biobanque réunissent des données multimodales (*e.g.,* génétique, mode de vie, biologie) associées à un suivi longitudinal. Elles sont par conséquent également chères et chronophages, mais également soumises à un biais de sélection car reposant sur le volontariat.
- Les bases médico-administratives ont, elles, une forte puissance du fait de leur très grande taille et permettent l'analyse de populations entières. Néanmoins, elles souffrent d'un manque de précision, sont limitées aux maladies ayant un codage bien défini, et sont soumises aux erreurs de codage.

- Les bases de données massives ouvertes ont un caractère hybride (entre données structurées et non structurées) mais souffrent individuellement de plusieurs limites (monocentriques, transversales, manque de granularité, absence de contrôle qualité).

III. Objectif

L'objectif de ce travail de thèse est d'étudier la faisabilité de la conduite de travaux épidémiologiques dans le domaine des soins critiques grâce à l'EDS AP-HP. Pour ce faire, nous nous sommes intéressés à des contextes cliniques contrastés, une affection fréquente et faisant l'objet d'un codage CIM-10 (l'infection à VRS [24]) et une affection rare sans codage (le *Drug Reaction with Eosinophilia and Systemic Symptoms*, DRESS). Les approches méthodologiques mises en œuvre ainsi que leurs forces et faiblesses sont présentées et discutées dans la partie résultats. La discussion de ce travail de thèse repose en grande partie sur un travail de synthèse des travaux réalisés grâce à l'EDS AP-HP depuis 2017 et de l'analyse comparative de leur méthodologie.

IV. Résultats

A. Article 1

Clinical Phenotypes and Outcomes Associated With Respiratory Syncytial Virus Infection in Critically Ill Patients: A Retrospective Multicenter Cohort Study in Greater Paris Area Hospitals, 2017–2023

Antoine Gaillet, Richard Layese, Slim Fourati, Héloise Celante, Tài Pham, Sarah Benghanem, Armand Mekontso Dessap, Etienne de Montmollin, Jeoffrey Pirault, Antoine Vieillard-Baron, Eric Maury, Yves Cohen, Maxens Decavèle, Muriel Fartoukh, Nicholas Heming, Charles Damoisel, Nadia Oubaya, Nicolas de Prost; Prono-RSV study group of the clinical data warehouse of Greater Paris University Hospitals

J Infect Dis. 2025 Sep 15;232(3):679-690.

Rationnel:

Le virus respiratoire syncytial (VRS), appartenant à la famille des Pneumoviridae et divisé en deux sous-types (A et B), est responsable d'infections respiratoires aiguës fréquentes chez l'adulte, comparables en incidence à la grippe (3–10 % par an). Bien que principalement associé à la bronchiolite du nourrisson, le VRS entraîne chez l'adulte des formes allant d'une atteinte bénigne des voies respiratoires supérieures à des tableaux graves, notamment chez les sujets ≥65 ans, porteurs de comorbidités cardio-respiratoires (insuffisance cardiaque, BPCO) ou en situation d'immunodépression. Chez les patients hospitalisés, la mortalité liée au

VRS est similaire à celle liée à la grippe, avec environ 20 % de recours à la réanimation et une mortalité hospitalière de 13 à 26 % dans ce sous-groupe. Les patients admis en réanimation présentent généralement plus de comorbidités que ceux infectés par la grippe. Le rôle pronostique des différents types d'immunodépression reste cependant mal caractérisé. L'émergence récente de stratégies préventives efficaces, incluant plusieurs vaccins dirigés contre la protéine F préfusion du VRS chez les sujets âgés et l'anticorps monoclonal nirsevimab chez le nourrisson, renforce la nécessité d'identifier les profils de patients adultes les plus à risque de formes sévères. L'étude présentée visait ainsi, dans une cohorte rétrospective de patients en réanimation, à analyser les facteurs de risque de mortalité à 30 jours avec un focus sur l'immunodépression. Nous avons également identifié, à l'aide d'une méthode de clustering non supervisé, des profils cliniques distincts et exploré leur lien avec le pronostic.

Matériels et méthodes:

Cette étude multicentrique, rétrospective, a inclus tous les patients ≥ 18 ans admis en réanimation dans l'un des 39 hôpitaux de l'AP-HP, dont 17 avec un service de réanimation, entre août 2017 et mai 2023 avec une infection documentée à VRS, identifiée par codes CIM-10. Le diagnostic reposait en pratique sur une RT-PCR positive, bien que l'accès à la donnée virologique ne soit pas rendu possible par l'EDS. Les données recueillies à l'admission en réanimation et durant le séjour incluaient l'âge, le sexe, les comorbidités (identifiées par codes CIM-10), le score de gravité SAPS II (lorsqu'il était disponible), la durée d'hospitalisation, ainsi que les variables liées à la prise en charge (recours à une ventilation invasive ou non invasive, à une épuration extrarénale ou à des vasopresseurs, selon la nomenclature CCAM). L'utilisation de corticostéroïdes, d'antibiotiques et de ribavirine, ainsi que la présence de co-infections respiratoires, a été identifiée à l'aide d'une méthode de traitement automatique du langage naturel (*Natural Language Processing*) basée sur des expressions régulières (regex), en

recourant à des mots-clés spécifiques et à leurs variantes. Lorsqu'une négation apparaissait dans une fenêtre de ±50 caractères autour du mot-clé ou de sa variante, le terme était exclu par l'algorithme. Les différentes sections des comptes rendus cliniques (telles que les antécédents médicaux, l'histoire de la maladie, l'évolution et la conclusion) ont été segmentées à l'aide de regex, les sections « antécédents médicaux » et « histoire de la maladie actuelle » étant retirées. L'algorithme a été validé par une relecture des comptes rendus cliniques et affiné de manière itérative par l'ajout de nouveaux mots-clés. Concernant la ribavirine, un échantillon aléatoire de 100 dossiers a été sélectionné, la relecture manuelle par un clinicien constituant la référence (« *gold standard* »). Les performances obtenues étaient les suivantes : exactitude 0,99 ; précision 0,93 ; rappel 0,99 ; F1-score 0,96. En ce qui concerne les mots-clés, pour les co-infections, nous avons utilisé une liste de bactéries et de termes définissant la présence d'une co-infection. Pour l'antibiothérapie, nous avons employé une liste de traitements antibiotiques ainsi que le terme « antibiotique » afin de détecter la prise d'antibiotiques durant l'hospitalisation. Pour la ribavirine, le mot « ribavirine » a été utilisé pour identifier son administration au cours du séjour. Les performances des algorithmes de recherche textuelle développés ont été vérifiées et améliorées par comparaison avec les informations pertinentes extraites lors de la lecture manuelle de comptes rendus hospitaliers par un médecin réanimateur. L'immunodépression était définie par un antécédent de cancer solide, transplantation, infection VIH, hémopathie, neutropénie ou autre cause. Le critère de jugement principal était la mortalité toutes causes à J30. Les critères secondaires comprenaient la mortalité en réanimation et à l'hôpital, la durée de séjour, et le recours aux supports vitaux. Les analyses statistiques comprenaient des comparaisons uni- et multivariées avec régressions logistiques ajustées sur des facteurs de confusion a priori (immunodépression, âge, sexe, comorbidités cardio-respiratoires, insuffisance respiratoire

aiguë, co-infections, créatininémie). Des modèles mixtes intégrant effet centre et période ont été utilisés, avec imputations multiples pour la créatinine manquante. L'immunodépression a été testée sous trois formes (binaire, type de pathologie, nombre de conditions). Enfin, une analyse exploratoire non supervisée par cartes auto-organisées de Kohonen (*self-organized maps*, SOM) a permis d'identifier des profils cliniques distincts et d'étudier leur association au pronostic.

Cette étude de cohorte a été approuvée par le comité d'éthique de l'Entrepôt de Données de Santé (EDS) de l'AP-HP (avis n° CSE-20-83), et aucun des patients inclus n'a exprimé d'opposition à y participer. L'Entrepôt de Données de Santé de l'AP-HP a, par ailleurs, reçu l'autorisation de la « Commission Nationale Informatique et Libertés » (CNIL).

Résultats:

Au total, 474 patients atteints d'infection à VRS ont été admis en réanimation sur la période étudiée. L'âge moyen était de 65±17 ans, 56% étaient des hommes, et 34% étaient immunodéprimés. Les comorbidités cardiovasculaires et respiratoires étaient fréquentes. La mortalité intra-hospitalière atteignait 13% (62/474). En analyse multivariée, l'âge >65 ans, l'insuffisance respiratoire aiguë et l'immunodépression étaient indépendamment associés à la mortalité (ORa=2,23 ; IC95% [1,25–4,00] ; p=0,006). L'analyse de clusters non supervisée distinguait trois profils : (i) un cluster exclusivement constitué de patients immunodéprimés (98/474 ; 20% de mortalité), (ii) un cluster de patients âgés avec comorbidités cardio-respiratoires (202/474 ; 13% de mortalité) et (iii) un cluster de patients plus jeunes (174/474 ; 9% de mortalité).

Conclusion du travail n°1 :

Concernant les infections à VRS, cette large cohorte rétrospective multicentrique ayant inclus des patients en réanimation/unités de soins intensifs, montre que l'immunodépression était systématiquement associée à la mortalité toutes causes à 30 jours. Trois clusters homogènes de patients, définis selon leurs profils de comorbidités, ont été identifiés et associés à des pronostics différents, ce qui pourrait guider l'implémentation des nouvelles thérapeutiques anti-VRS visant à prévenir l'admission en réanimation.

Concernant la conduite de la recherche dans l'EDS, la pathologie traitée était accessible via le codage CIM-10, cependant l'accès aux dossiers médicaux fournis par la plateforme a permis d'obtenir, via une recherche textuelle par traitement automatique du langage naturel, des informations essentielles concernant les co-infections et la prise en charge de l'infection à VRS, des données qui n'auraient pas été accessibles sans étude individuelle des comptes-rendus (réalisée pour une centaine de patients afin de valider l'algorithme de NLP). Le gain de temps, sans perte de sens, apporté par le recours au NLP, a permis de constituer une cohorte de grande taille tout en apportant des informations précises.

The Journal of Infectious Diseases

MAJOR ARTICLE

Infectious Diseases Society of America

HIV Medicine Association

OXFORD

Clinical Phenotypes and Outcomes Associated With Respiratory Syncytial Virus Infection in Critically Ill Patients: A Retrospective Multicenter Cohort Study in Greater Paris Area Hospitals, 2017–2023

Antoine Gaillet,[1,2,3] Richard Layese,[3,4] Slim Fourati,[3,5,6] Héloise Celante,[1,2] Tài Pham,[7,8] Sarah Benghanem,[9] Armand Mekontso Dessap,[1,2,3] Etienne de Montmollin,[10] Jeoffrey Pirault,[11] Antoine Vieillard-Baron,[12] Eric Maury,[13] Yves Cohen,[14] Maxens Decavèle,[15,16] Muriel Fartoukh,[17] Nicholas Heming,[18] Charles Damoisel,[19] Nadia Oubaya,[3,5] and Nicolas de Prost[1,2,3,5]; on behalf of the Prono-RSV Study Group of the Clinical Data Warehouse of Greater Paris University Hospitals

[1]Service de Médecine Intensive Réanimation, Hôpitaux Universitaires Henri Mondor–Albert Chenevier, Assistance Publique–Hôpitaux de Paris, Créteil, France; [2]Groupe de Recherche Clinique Cardiovascular and Respiratory Manifestations of Acute lung injury and Sepsis (CARMAS), Université Paris Est–Créteil, Créteil, France; [3]Inserm, Institut Mondor de Recherche Biomédicale, Université Paris Est–Créteil, Créteil, France; [4]Clinical Research Unit, Assistance Publique–Hôpitaux de Paris, Henri Mondor Hospital, Créteil, France; [5]Inserm U955, Team "Viruses, Hepatology, Cancer," Henri Mondor Hospital, Créteil, France; [6]Department of Virology, Hôpitaux Universitaires Henri Mondor, Assistance Publique–Hôpitaux de Paris, Créteil, France; [7]Service de Médecine Intensive Réanimation, Hôpital de Bicêtre, département médico-universitaire "Maladie du coeur et des vaisseaux" (CORREVE), Fédérations Hospitalo-universitaires SEPSIS, Groupe de Recherche Clinique Cardiovascular and Respiratory Manifestations of Acute lung injury and Sepsis (CARMAS), Université Paris-Saclay, Assistance Publique–Hôpitaux de Paris, Le Kremlin-Bicêtre, France; [8]Inserm U1018, Equipe d'Epidémiologie Respiratoire Intégrative, Centre de Recherche en Epidémiologie et Santé des Populations, Université Paris-Saclay, Université de Versailles Saint-Quentin-en-Yvelines, Université Paris Sud, Villejuif, France; [9]Medical Intensive Care Unit, Cochin University Hospital, Paris, France; [10]Medical and Infectious Diseases Intensive Care Unit, Bichat Claude Bernard University Hospital, Université Paris Cité, Assistance Publique–Hôpitaux de Paris, Paris, France; [11]Service de Médecine Intensive Réanimation, Université Paris Cité, Assistance Publique–Hôpitaux de Paris, Hôpital Louis Mourier, département médico-universitaire Enseignements et Soins de Proximité, Recherche, Innovation et Territoires (ESPRIT), Colombes, France; [12]Service de Médecine Intensive Réanimation, Assistance Publique–Hôpitaux de Paris, Hôpital Ambroise Paré, Boulogne-Billancourt, France; [13]Medical Intensive Care Unit, Hôpital Saint-Antoine, Assistance Publique–Hôpitaux de Paris, Saint-Antoine University Hospital, Sorbonne University, Paris, France; [14]Intensive Care Unit, Avicennes Hospital, Assistance Publique–Hôpitaux de Paris, University Sorbonne Paris Nord, Bobigny, France; [15]Service de Pneumologie, Médecine Intensive et Réanimation (Département R3S), Groupe Hospitalier Universitaire Assistance Publique–Hôpitaux de Paris, Sorbonne Université, site Pitié-Salpêtrière, Paris, France; [16]Inserm, UMR S1158 Neurophysiologie Respiratoire Expérimentale et Clinique, Sorbonne Université, Paris, France; [17]Médecine Intensive Réanimation, Assistance Publique–Hôpitaux de Paris, Sorbonne Université, Hôpital Tenon, Paris, France; [18]Médecine Intensive Réanimation, Hôpital Raymond Poincaré, Assistance Publique–Hôpitaux de Paris, Garches, France; and [19]Service de Réanimation Polyvalente et Surveillance Continue, Assistance Publique–Hôpitaux de Paris, Université Paris Sud, Hôpital Antoine Béclère, Clamart, France

Background. Respiratory syncytial virus (RSV) is one of the deadliest respiratory viruses. There is a need to better identify prognostic factors in RSV-infected patients, particularly those requiring intensive care unit (ICU) admission, with a focus on immunosuppressed patients.

Methods. This was a multicenter, retrospective cohort study of RSV-infected adults hospitalized in 17 ICUs in the Greater Paris area between 1 August 2017 and 1 May 2023. The primary endpoint was all-cause day 30 mortality. Supervised and unsupervised analyses were performed.

Results. During the study, 474 RSV-infected patients were admitted (56% male, mean age 65±17 years, 34% immunosuppressed). Day 30 mortality was 14%. Immunosuppression was linked to mortality (aOR=2.10, 95% CI [1.14;3.80], p=0.035). Cluster analysis identified three groups: (1) immunosuppressed (17%, highest mortality 21%), (2) older with comorbidities (43%, mortality 14%), and (3) younger (37%, lowest mortality 9%).

Conclusions. One-third of ICU patients with RSV infection were immunosuppressed, and both supervised and unsupervised methods linked immunosuppression to day 30 mortality. Anti-RSV therapies preventing ICU admission should be evaluated in this subgroup.

Keywords. acute respiratory failure; respiratory syncytial virus; RSV; immunosuppression; intensive care unit.

Respiratory syncytial virus (RSV) is an RNA virus of the Pneumoviridae family, divided into 2 antigenically distinct subgroups: A (RSV-A) and B (RSV-B) [1]. While RSV is best known for causing potentially fatal acute bronchiolitis in young children [2], it is also one of the most common causes of acute respiratory viral infections in adults, infecting 3%–10% of adults annually, comparable to influenza infections [3, 4]. Infected adults present with symptoms ranging from limited upper respiratory tract symptoms to potentially fatal disease, particularly in those aged ≥65 years [5], with cardiorespiratory

Received 04 January 2025; editorial decision 10 March 2025; accepted 10 March 2025

Correspondence: Antoine Gaillet, MD, Service de Médecine Intensive Réanimation, Hôpitaux Universitaires Henri Mondor, AP-HP, 1 rue Gustave Eiffel, Créteil Cedex 94010, France (antoine.gaillet@aphp.fr).

The Journal of Infectious Diseases®

© The Author(s) 2025. Published by Oxford University Press on behalf of Infectious Diseases Society of America. All rights reserved. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

https://doi.org/10.1093/infdis/jiaf129

comorbidities, including congestive heart failure and chronic obstructive pulmonary disease (COPD) and/or severe immunosuppression [6–8].

RSV infection in hospitalized patients has recently been shown to be associated with hospital mortality rates similar to those of adult patients with influenza infection [9, 10]. Previous studies have reported a greater need for intensive care unit (ICU) admission in patients infected with RSV infection than influenza [11–13], but this has not been replicated in all studies and the role of influenza vaccination may be a confounding factor. In addition, adult patients admitted to the ICU with RSV infection tend to have a higher prevalence of comorbidities, including cardiovascular and respiratory diseases and immunosuppressive conditions, compared to those with influenza infection [10].

In previously published cohorts, approximately 20% of patients hospitalized for RSV infection required ICU admission and, among these, in-hospital mortality ranged from 13% to 26% [9, 10, 14]. Immunosuppressed patients have previously been shown to be at risk of severe RSV infection, but there are limited data on their clinical profile and, in particular, the prognostic implications of different subtypes of immunosuppression.

Recently, new effective RSV immunization strategies have emerged: Several randomized controlled trials have demonstrated the efficacy of 3 RSV prefusion F protein vaccines in preventing RSV-associated acute respiratory infection and lower respiratory tract disease in elderly patients [15]. In addition, a new long-acting monoclonal antibody, nirsevimab, has been approved by the US Food and Drug Administration (FDA) and the European Medicines Agency and is now available in several countries to prevent RSV infection in newborns and infants <1 year of age [16]. With the advent of this new range of preventive therapies (and the possible future development of curative therapies), it seems necessary to better identify patients at highest risk of death in order to optimize the personalization of care and improve prognosis.

Therefore, the aim of the following study was to analyze in a retrospective cohort of critically ill RSV patients: (1) in a supervised approach, the risk factors for day 30 mortality, with a focus on immunosuppressed patients; (2) in an unsupervised approach, using clustering techniques based on machine learning algorithms, to identify and characterize clinical profiles and, secondarily, to study their associations with outcomes.

METHODS

Study Design and Population

We conducted a multicenter, retrospective cohort study including all patients aged ≥18 years, admitted to the ICUs of Assistance Publique–Hôpitaux de Paris (AP-HP) hospitals between 1 August 2017 and 1 May 2023 with a diagnosis of RSV infection. The diagnosis of RSV infection was defined by 1 of the following *International Classification of Diseases, 10th Revision* (*ICD-10*) codes: J121, J205, J210, or B974, in any hospital stay during the study period using the French national hospital database (Program for Medicalization of Information Systems) (Supplementary Table 1). In practice, all patients coded with an RSV diagnosis had a positive RSV reverse-transcription polymerase chain reaction test, but unfortunately we did not have access to the type of specimen (ie, nasopharyngeal swab, tracheal aspirate, bronchoalveolar lavage fluid) from which the diagnosis was made (as well as for coinfections).

We used the AP-HP Health Data Warehouse, comprising 39 hospitals in the Greater Paris area, 17 of which have an ICU. The AP-HP Data Warehouse aggregates all data from inpatients and outpatients from any of the AP-HP hospitals, including sociodemographics, clinical summaries, biological data, medication prescriptions, and medical records with diagnoses and interventions related to the hospital stays and vital status at hospital discharge.

Ethical Considerations

This cohort study was approved by the Institutional Review Board of the AP-HP Data Warehouse (approval number CSE-20-83), and patients enrolled in this study did not refuse to participate in the study. The AP-HP Data Warehouse was approved by the Commission Nationale Informatique et Libertés.

Data Recorded

The following data were collected on admission to the ICU and during ICU stay: age; sex; comorbidities (identified by *ICD-10* codes, listed in Supplementary Table 2); Simplified Acute Physiology Score II (SAPS II) if available; duration of hospital stay; and variables related to patient management, including need for invasive or noninvasive ventilatory support, renal replacement therapy, or vasopressors, using the medical procedure nomenclature (Classification Commune des Actes Médicaux; Supplementary Table 3), and use of corticosteroids, antibiotics, and ribavirin and presence of coinfection, using a natural language processing method. Regex-based methods were used to develop the algorithm, using specific keywords and their variants. If a negation appeared within a window of ±50 characters from the position of the keyword or variant, it was excluded by the algorithm. The different sections of the clinical reports (such as medical history, disease history, progression, and conclusion) were segmented using regex, with the sections "medical history" and "history of current disease" removed. The algorithm was validated through a review of clinical reports and iteratively refined with the addition of new keywords. For ribavirin, a random sample of 100 reports was selected, with the gold standard being the clinician's manual review of each report. The following performance metrics were observed: accuracy 0.99, precision 0.93, recall 0.99, F1 score

0.96. Biological data were collected on the first 2 days of the stay (D1 and D2). The variable "immunosuppression" was defined by the presence of 1 of the following variables: "solid cancer" and/or "solid organ transplant" and/or "human immunodeficiency virus infection" (HIV) and/or "allogeneic hematopoietic stem cell transplantation" (HSCT) and/or "neutropenia" and/or "hematological disease" and/or "other immunosuppression," using *ICD-10* codes (Supplementary Table 2). The variable coinfection was defined by the documentation of pathogen in a respiratory sample (ie, bacteria, virus, or fungus) other than RSV within 48 hours of ICU admission together with a consistent clinical context, documented in the patient's medical record.

Endpoints

The primary endpoint was all-cause mortality at day 30. Secondary endpoints included the need for noninvasive and invasive mechanical ventilation support, hemodialysis, or vasopressor support during ICU stay; duration of ICU and hospital stay; and in-hospital and ICU mortality.

Statistical Analyses

Categorical variables were presented as numbers (percentages) and quantitative variables as mean ± standard deviation or median (interquartile range [IQR]), as appropriate. Characteristics of patients who died in hospital and those who did not were compared using χ^2 tests or Fisher exact tests for categorical variables and Student *t* tests or Mann-Whitney tests for quantitative variables, as appropriate.

Risk factors for all-cause day 30 mortality were assessed using univariable and multivariable logistic regression models, adjusting for a priori–selected confounders previously shown to be associated with mortality (ie, immunosuppression [10], age [10], sex [17], chronic respiratory and heart failure [18], acute respiratory failure [19], bacterial and viral coinfection [20–22], and serum creatinine [23]). Adjusted odds ratios (aORs) along with their 95% confidence intervals (CIs) were computed. We used a ratio of "5 events/adjustment variable ratio" as recommended [24, 25]. Center and period mixed effects were also included in the logistic regression models. Because the relationship between immunosuppression and mortality has not been widely reported in the literature, we decided to express the immunosuppression variable in 3 different ways within 3 otherwise similar regression models: (1) immunosuppression (yes/no); (2) underlying immunosuppressive condition (no [reference], solid cancer, solid organ transplantation [SOT], hematological disease and/or allogeneic HSCT, other immunosuppressive condition); and (3) immunosuppressive conditions (number). Only the variable "serum creatinine" (30% missing data) had to be imputed (multiple imputation) for the multivariable regression analysis. Likelihood ratio tests were used to search for clinically relevant interactions between variables using. Log-linearity (not respected only for age, dichotomized accordingly at the threshold of age >65 years), absence of multicollinearity, and outliers were checked. Discrimination of the models was measured by their areas under the curve and 95% CIs. All significance tests were 2-tailed and the statistical significance level was set at 5%.

To illustrate differences in the phenotypes of critically ill RSV-infected patients, we performed an exploratory unsupervised clustering analysis using Kohonen's self-organized map (SOM) method [26], which allows us to build 2-dimensional maps from multidimensional datasets. In a nutshell, each map is divided into districts in which patients are located by the SOM algorithm on the basis of their characteristics: Patients with similar features are closely located on the maps, while patients with distinct profiles are farther from each other, hence allowing to identify key differences or similarities among them. The number of patients per district is not uniformly distributed. Quality control of the SOMs was achieved by checking the distribution of data points on each map using histograms and by directly visualizing potential subgroups of problematic data points. The SOMs were obtained with the Numero package framework for the R statistical platform after principal component analysis adapted for mixtures of qualitative and quantitative variables was applied [27]. Once the clusters were identified, we looked at their prognostic value for outcomes.

All analyses were performed with Python software (version 3.7) and RStudio software, version 4.0.5 (R project for Statistical Computing, Vienna, Austria).

RESULTS

Patient Population

From August 2017 to May 2023, 2555 patients were admitted to the 17 participating centers with RSV infection, of whom 474 (19%) were admitted to the ICU (Figure 1). The temporal distribution of ICU admissions of RSV-infected patients over the study period followed a winter seasonal pattern, as shown in Supplementary Figure 1.

The majority of patients were male (n = 265/474 [56%]), with a mean age of 65 ± 17 years, and had a high prevalence of underlying comorbidities, including metabolic conditions (ie, obesity, n = 154/474 [32%]; diabetes, n = 135/474 [28%]), cardiovascular disease (ie, hypertension, n = 243/474 [51%]), chronic heart failure (n = 174/274 [37%]), and respiratory disease (ie, COPD, n = 189/474 [40%]) (Table 1). In addition, approximately one-third of patients (n = 162/474 [34%]) had at least 1 immunosuppressive condition, including hematological malignancies (n = 61/474 [13%]), solid cancers (n = 40/474 [8%]), SOT (n = 34/474 [7%]), allogeneic HSCT (n = 20/474 [4%]), neutropenia (n = 14/474 [3%]), and HIV infection (n = 14/474 [3%]). The majority of patients (n = 355/474 [75%]) presented with acute respiratory failure, as expected in critically ill RSV-infected patients. Coinfections were documented in 23% of cases (n = 107/474), mainly bacterial (n = 95/474 [20%]),

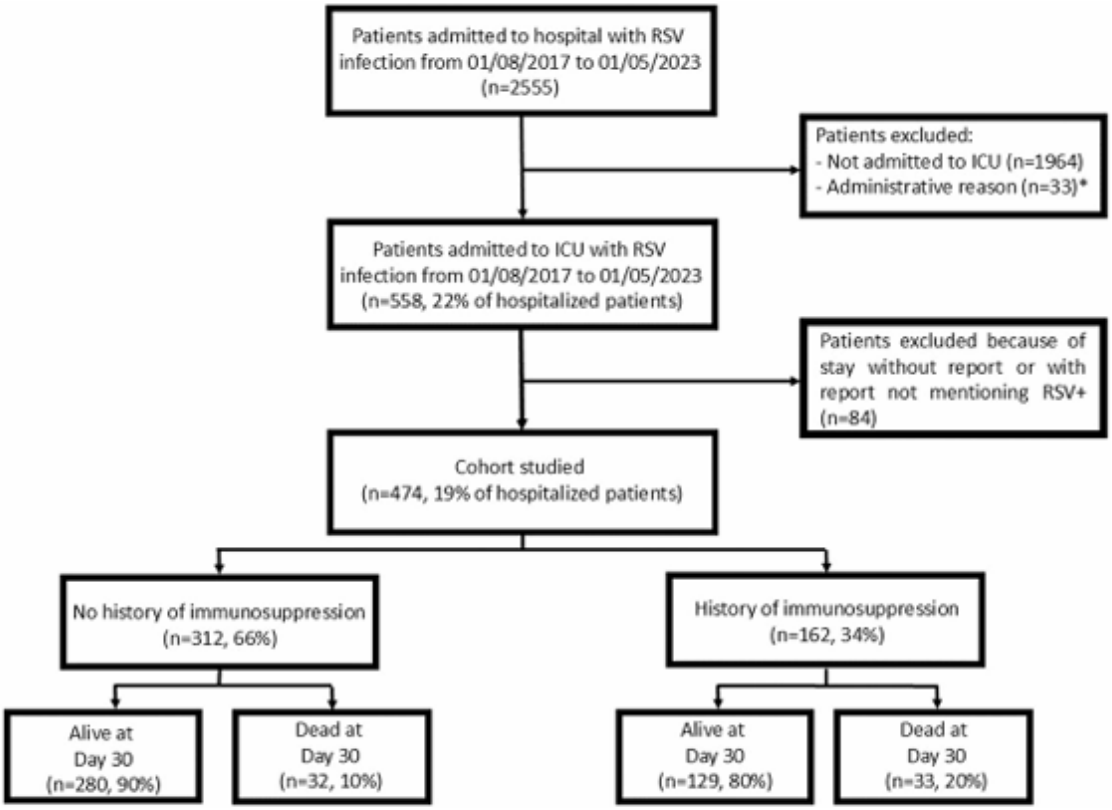


Figure 1. Flowchart of the cohort. *A total of 278 patients, who were initially part of the previous study [19] and were hospitalized between 1 January 2015 and 31 July 2017, were ineligible for inclusion in the current study due to changes in regulatory requirements. More precisely, patients admitted to Assistance Publique–Hôpitaux de Paris before 1 August 2017 were no longer eligible for analysis. Among these 278 patients, 33 were hospitalized in clinical wards between 1 January 2015 and 31 July 2017 and were subsequently admitted to the intensive care unit (ICU) for another respiratory syncytial virus (RSV) infection between 1 August 2017 and 1 May 2023. Thus, these 33 patients were not included in the current study.

involving *Streptococcus pneumoniae* (n = 35), *Pseudomonas aeruginosa* (n = 16), *Haemophilus influenzae* (n = 15), and *Klebsiella pneumoniae* (n = 7).

Throughout their ICU stay, most patients received antibiotics (n = 392/474 [83%]) and half received corticosteroids (n = 245/474 [52%]), whereas ribavirin was rarely used (n = 15/474 [3%]) (Table 2). Half of the patients required noninvasive ventilation support (n = 263/474 [55%]) and more than a quarter required invasive mechanical ventilation support (n = 131/474 [28%]). The median duration of hospital and ICU stay was 13 (IQR, 8–23) and 5 (IQR, 3–11) days, respectively, with a very short pre-ICU hospital stay (median, 0 days).

Risk Factors for All-Cause Day 30 Mortality

All-cause day 30 mortality was 14% (n = 65/474), with no significant difference in the annual mortality rate observed over the study time period (Table 2, Supplementary Figure 2). Patients who died were significantly older and more likely to have immunosuppressive conditions, including hematological malignancies, allogenic HSCT, or solid cancers (Table 1).

Regardless of how the immunosuppression variable was expressed in the 3 logistic regression models tested, it was consistently associated with day 30 all-cause mortality in univariable and multivariable analyses (Table 3). Whether immunosuppression was considered as a binary variable (aOR, 2.10 [95% CI, 1.14–3.80]; $P = .035$), qualitatively distinguishing submodalities of immunosuppression ($P = .038$, with excess risk for all submodalities, but statistically significant only for the hematological malignancy/allograft category; aOR, 2.68 [95% CI, 1.47–6.43]), or quantitatively (aOR, 1.64 [95% CI, 1.28–2.66]; $P = .002$, for each additional immunosuppressive condition), it remained significantly associated with day 30 mortality. Other covariables included in the models and significantly associated with day 30 all-cause mortality were age >65 years and acute respiratory failure. The variance explained by the random effect on centers and time period was <1%.

Unsupervised Clustering Analysis

Figure 2 shows the clustering analysis obtained with the SOM method, displaying polarized distributions of patient characteristics across the maps. From these results, 3 clusters of patients with homogeneous phenotypes could be individualized. The cluster boundaries are shown in Figure 2 and the detailed patient characteristics are shown in Table 4.

Table 1. Characteristics of Patients With Respiratory Syncytial Virus Infection at Intensive Care Unit Admission, According to Vital Status at Day 30

Characteristic	All Patients (n = 474)	Day 30 Survivors (n = 409)	Day 30 Nonsurvivors (n = 65)	*P* Value
Age, y	65 ± 17	62 ± 15	67 ± 14	**<.001**
Female sex	209 (44)	182 (44)	27 (44)	.901
BMI ≥30 kg/m²	154 (32)	131 (32)	23 (35)	.467
Diabetes	135 (28)	113 (28)	22 (34)	.489
Hypertension	243 (51)	209 (51)	34 (52)	.785
Peripheral arterial obstructive disease	30 (6)	27 (7)	3 (5)	.245
Chronic heart failure	174 (37)	147 (36)	27 (41)	.691
COPD	189 (40)	166 (41)	23 (35)	.389
Chronic respiratory failure	109 (23)	93 (23)	16 (25)	.835
Cirrhosis	16 (3)	12 (3)	4 (6)	.144
Chronic alcoholism	32 (7)	27 (7)	5 (8)	.920
Chronic kidney disease	34 (7)	29 (7)	5 (8)	.771
Chronic hemodialysis	23 (5)	18 (4)	5 (8)	.390
Immunosuppression	162 (34)	130 (32)	32 (49)	**.004**
Solid cancer	40 (8)	30 (7)	10 (15)	**.021**
Solid organ transplant[a]	34 (7)	27 (7)	7 (11)	.744
Hematologic malignancy	61 (13)	46 (11)	15 (23)	**.003**
Allogeneic HSCT	20 (4)	12 (3)	8 (12)	**.007**
Neutropenia	14 (3)	10 (2)	4 (6)	.091
HIV infection	14 (3)	14 (3)	2 (3)	1
Other[b]	13 (3)	10 (2)	3 (5)	.390
No. of immunosuppressive conditions	0 (0–1)	0 (0–0)	0 (0–1)	**<.001**
Volume of patients (per center and per year)				
<2	46 (10)	40 (10)	6 (9)	.111
2–6	163 (34)	133 (32)	30 (46)	
>6	265 (46)	237 (59)	28 (43)	
Year of admission				
2017	55 (12)	49 (12)	6 (9)	.231
2018	85 (18)	76 (18)	9 (14)	
2019	89 (19)	80 (19)	9 (14)	
2020	33 (7)	26 (6)	7 (11)	
2021	56 (12)	47 (11)	9 (14)	
2022	131 (28)	115 (28)	16 (25)	
2023	25 (5)	16 (4)	9 (14)	
Acute respiratory failure	355 (75)	298 (73)	57 (88)	**.007**
SAPS II (n = 105)	38 ± 17	38 ± 15	53 ± 29	**<.001**
Coinfections	107 (23)	92 (22)	15 (23)	.701
Bacteria	95 (20)	87 (21)	8 (12)	.087
Gram-positive cocci	57 (12)	17 (4)	2 (3)	
Gram-negative bacilli	48 (10)	27 (7)	7 (11)	
Nonfermenting	22 (5)	41 (10)	1 (2)	
Virus	18 (4)	12 (3)	6 (9)	.069
SARS-CoV-2	5 (1)	4 (1)	3 (5)	
Influenza	11 (2)	8 (2)	3 (5)	
Fungus	2 (0.4)	1 (0.2)	1 (2)	
Blood leukocytes, g/L (n = 388)	11 (8–15)	11 (8–14)	11 (8–15)	.520
Neutrophils, g/L (n = 369)	8 (5–12)	8 (4–11)	9 (5–12)	.412
Lymphocytes, g/L (n = 366)	1 (0.6–1.6)	1 (0.8–1.7)	0.9 (0.6–1.6)	**.041**
Serum creatinine, μM (n = 328)	92 (65–148)	90 (62–109)	140 (74–160)	**<.001**
Hemoglobin, g/L (n = 388)	12 (11–14)	12 (11–15)	11 (10–14)	**.014**

Data are presented as No. (%), mean ± standard deviation, or median (interquartile range). Two-tailed *P* values come from unadjusted comparisons using χ^2 or Fisher exact tests for categorical variables, and *t* tests or Mann-Whitney tests for continuous variables, as appropriate. No adjustment for multiple comparisons was performed. Bolded *P* values are significant at the *P* < .05 level.

Abbreviations: BMI, body mass index; COPD, chronic obstructive pulmonary disease; HIV, human immunodeficiency virus; HSCT, hematopoietic stem cell transplantation; SAPS II, Simplified Acute Physiology Score II; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

[a]Solid organ transplant: kidney (n = 14), heart (n = 9), lung (n = 7), liver (n = 4).

[b]Other: treatment with antitumor, immunosuppressive drugs (n = 8); treatment with glucocorticoids and disease-modifying antirheumatic drugs (n = 7).

Table 2. Intensive Care Management and Outcomes of Patients With Respiratory Syncytial Virus Infection According to Vital Status at Day 30

Management and Outcome	All Patients (n = 474)	Day 30 Survivors (n = 409)	Day 30 Nonsurvivors (n = 65)	*P* Value
Antibiotics	392 (83)	334 (82)	58 (89)	**.039**
Corticosteroids	245 (52)	212 (52)	33 (51)	.577
Ribavirin	15 (3)	10 (2)	5 (8)	**.034**
Palivizumab	0	0	0	
Intravenous immunoglobulins	7 (1)	4 (1)	3 (5)	.051
Noninvasive mechanical ventilation	263 (55)	221 (54)	42 (65)	.070
Invasive mechanical ventilation	131 (28)	96 (23)	35 (54)	**<.001**
Vasopressor support	113 (24)	79 (19)	34 (52)	**<.001**
Renal replacement therapy	11 (2)	4 (1)	7 (11)	**.001**
Duration of hospital stay, d	13 (8–23)	13 (10–25)	10 (7–20)	**.007**
Pre-ICU hospitalization stay, d	0 (0–0)	0 (0–0)	0 (0–0)	.586
Duration of ICU stay, d	5 (3–11)	5 (3–9)	7 (4–12)	.077
In-ICU mortality	52 (11)	0	52 (80)	
In-hospital mortality	62 (13)	0	62 (95)	

Data are presented as No. (%) or median (interquartile range). Two-tailed *P* values come from unadjusted comparisons using χ^2 or Fisher exact tests for categorical variables, and *t* tests or Mann-Whitney tests for continuous variables, as appropriate. No adjustment for multiple comparisons was performed. Bolded *P* values are significant at the $P < .05$ level.

Abbreviation: ICU, intensive care unit.

Cluster 1 (n = 98/474 [17%]) consisted mainly of men (n = 73/98 [74%]) with the common feature that all were immunosuppressed (100%), mainly with an underlying hematological disease or solid cancer, with a median number of immunosuppressive conditions of 2 (IQR, 1–3), higher than in the other clusters. In addition, several of these patients had underlying cardiovascular and respiratory comorbidities, but to a lesser extent than cluster 2 patients. Cluster 1 patients had the highest day 30 mortality (21% vs 14% and 9% for clusters 2 and 3, respectively).

Cluster 2 was the main cluster, comprising 202 of 474 patients (43%). They were the oldest (median age 71 years) and had several comorbidities, except for immunosuppressive conditions. Indeed, 51% (n = 103/202) of them had obesity, 41% (n = 83/202) had diabetes, 68% (n = 138/202) had hypertension, 51% (n = 103/202) had chronic heart failure, and 60% (n = 121/202) had COPD. Almost all patients in cluster 2 presented with acute respiratory failure (n = 184/202 [91%]) and required noninvasive ventilation more frequently than patients in other clusters (73% [n = 148/202] vs 50% [n = 49/98] and 38% [n = 66/174] for clusters 1 and 3, respectively).

Finally, cluster 3 (n = 174/474 [37%]) included the youngest patients (median age, 57 years), the majority of whom had no or a few comorbidities, except for chronic dialysis in 9% (n = 16/202) and immunosuppression (mainly SOT) in 27% (n = 47/202). This cluster had the lowest mortality rate (n = 16/202 [9%]).

DISCUSSION

The main findings of the current study, which included RSV-infected patients admitted to the ICUs of 17 hospitals in the Greater Paris area between 2017 and 2023, are as follows: (1) All-cause day 30 mortality was 14% (n = 65/474); (2) age >65 years, acute respiratory failure at ICU admission, and underlying immunosuppression were consistently associated with all-cause day 30 mortality in univariable and multivariable logistic regression; and (3) an unsupervised approach identified 3 homogeneous clusters of patients, including a cluster corresponding to immunosuppressed patients who had the highest mortality rate, confirming previous findings.

Patients included in the current study were admitted during 6 consecutive winters, following a typical seasonal pattern for RSV infections. Of note, as previously reported, almost no admissions occurred in 2021 due to the implementation of strict barrier measures in the context of the first waves of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic, then a peak occurred in the winter of 2022–2023, probably explained by the combination of a relaxation of the aforementioned barrier measures and a reduced population immunity following a prolonged period of reduced RSV exposure ("immunity debt" hypothesis) during the 2 preceding years of the SARS-CoV-2 pandemic. Our cohort was consistent with previously published cohorts of critically ill RSV-infected patients in terms of patient demographics and comorbidity features [10, 19, 28]. Underlying chronic cardiovascular and respiratory comorbidities have been previously reported to be frequently associated with RSV infection and to be associated with severity in hospitalized RSV-infected patients [9, 14, 29], including a greater risk of requiring invasive mechanical ventilation [19, 20].

We found a stable and robust association between immunosuppression status and all-cause day 30 mortality. This finding was based not only on the supervised approach (ie, logistic regression modeling), but also on the unsupervised approach, which allowed the identification of a single cluster (cluster 1) that included only immunosuppressed patients and had a

Table 3. Univariable and Multivariable Logistic Regression Models Assessing Risk Factors for Day 30 Mortality

Model[a]	Variable	Day 30 Survivors (n = 409)	Day 30 Nonsurvivors (n = 65)	OR (95% CI)	*P* Value	Adjusted OR (95% CI)	*P* Value
Model 1	Immunosuppression (yes/no)	130 (32)	32 (49)	2.11 (1.24–3.70)	**.006**	2.10 (1.14–3.80)	**.035**
	Age >65 y	219 (53)	49 (75)	2.73 (1.51–5.12)	**.002**	3.05 (1.50–5.56)	**.003**
	Female sex	182 (44)	27 (44)	0.96 (.54–1.64)	.922	1.11 (.63–2.12)	.784
	Chronic respiratory failure	93 (23)	16 (25)	1.06 (.57–1.97)	.243	0.98 (.46–1.86)	.606
	Chronic heart failure	147 (36)	27 (41)	1.12 (.65–1.94)	.730	0.88 (.49–1.56)	.783
	Acute respiratory failure	298 (73)	57 (88)	3.55 (1.60–9.37)	**.010**	3.76 (1.56–9.12)	**.003**
	Bacterial coinfection	87 (21)	8 (12)	0.54 (.23–1.15)	.126	0.41 (.15–1.05)	.063
	Viral coinfection	12 (3)	6 (9)	2.67 (.86–7.45)	.073	2.98 (.87–8.23)	.100
	Serum creatinine (μM)	80 (59–107)	107 (75–155)	1.00 (.99–1.01)	.157	1.02 (.96–1.07)	.453
Model 2	Underlying immunosuppressive condition						
	No immunosuppression	280 (68)	32 (49)	Ref.	**.046**	Ref.	**.038**
	Solid cancer[b]	30 (7)	10 (15)	2.10 (.70–4.84)		1.50 (.42–3.10)	
	Solid organ transplant[b]	27 (7)	7 (11)	1.86 (.52–4.70)		2.33 (.76–8.23)	
	Hematologic malignancy and/or allogeneic HSCT	58 (14)	23 (35)	2.85 (1.36–5.51)		2.68 (1.47–6.43)	
	Other immunosuppressive conditions[c]	34 (8)	9 (14)	1.51 (.49–3.61)		1.12 (.76–5.41)	
	Age >65 y	219 (53)	49 (75)	2.55 (1.45–5.10)	**.001**	2.25 (1.52–5.66)	**.005**
	Female sex	182 (44)	27 (44)	0.91 (.51–1.57)	.923	1.11 (.53–2.20)	.736
	Chronic respiratory failure	93 (23)	16 (25)	1.22 (.63–1.95)	.246	1.13 (.61–1.89)	.506
	Chronic heart failure	147 (36)	27 (41)	1.34 (.71–1.94)	.736	0.97 (.35–1.37)	.822
	Acute respiratory failure	298 (73)	57 (88)	3.76 (1.69–9.39)	**.006**	3.80 (1.57–9.38)	**.001**
	Bacterial coinfection	87 (21)	8 (12)	0.45 (.20–1.12)	.111	0.56 (.24–1.03)	.052
	Viral coinfection	12 (3)	6 (9)	2.72 (.81–7.20)	.076	2.45 (.88–8.23)	.110
	Serum creatinine (μM)	80 (59–107)	107 (75–155)	1.01 (.99–1.02)	.184	1.00 (.98–1.01)	.608
Model 3	Immunosuppressive conditions (No.)	0 (0–0)	0 (0–1)	2.27 (1.37–3.96)	**.003**	1.64 (1.28–2.66)	**.002**
	Age >65 y	219 (53)	49 (75)	2.83 (1.49–5.11)	**.003**	2.58 (1.40–5.87)	**.003**
	Female sex	182 (44)	27 (44)	0.96 (.54–1.45)	.922	1.13 (.64–2.07)	.714
	Chronic respiratory failure	93 (23)	16 (25)	1.18 (.62–1.91)	.231	1.09 (.54–1.83)	.531
	Chronic heart failure	147 (36)	27 (41)	1.05 (.72–1.70)	.741	0.99 (.54–1.49)	.831
	Acute respiratory failure	298 (73)	57 (88)	3.86 (1.43–8.45)	**.007**	3.98 (1.55–9.19)	**.002**
	Bacterial coinfection	87 (21)	8 (12)	0.54 (.28–1.11)	.136	0.67 (.35–.99)	.053
	Viral coinfection	12 (3)	6 (9)	2.57 (.91–6.99)	.059	2.10 (.78–8.43)	.097
	Serum creatinine (μM)	80 (59–107)	107 (75–155)	1.00 (.99–1.01)	.151	1.00 (.99–1.02)	.477

Data are presented as No. (%) or median (interquartile range) unless otherwise indicated. *P* values come from multivariable logistic regression models. Bolded *P* values are significant at the $P < .05$ level. The models showed good discrimination: the area under the curve (95% CI) was 0.74 (.69–.79) for model 1, 0.76 (.69–.80) for model 2, and 0.75 (.71–.78) for model 3.

Abbreviations: CI, confidence interval; HSCT, hematopoietic stem cell transplantation; OR, odds ratio.

[a]Center and period mixed effects were included in the models independently of their explanatory variance (similar results without mixed effects; data not shown).

[b]Solid organ transplant: kidney (n = 14), heart (n = 9), lung (n = 7), liver (n = 4).

[c]Other immunosuppressive conditions: neutropenia, human immunodeficiency virus infection, innate immunodeficiency, exposure to immunosuppressants (antitumor and immunosuppressive drugs, glucocorticoids, and disease-modifying antirheumatic drugs).

higher all-cause day 30 mortality compared with other clusters. The greater severity of RSV infection in immunosuppressed patients has been highlighted in previous reports, although these have reported quite variable rates of hospital mortality (from 7% to 35%), particularly in allogeneic HSCT and SOT recipients [20, 30–32]. However, recent analyses performed in ICU patients have yielded conflicting results, with immunosuppression being inconsistently associated with mortality in multivariable analysis [10]. Our study not only confirms the strong association between immunosuppression and mortality, but also highlights the burden of RSV disease in this specific population, with one-third of patients being immunosuppressed, and shows that patients with onco-hematological malignancies have the highest risk of death among immunosuppressed patients.

Cluster analysis isolated 3 patient phenotypes. First, cluster 1 consisted exclusively of onco-hematological patients with high mortality, in whom ribavirin and intravenous immunoglobulins were rarely administered, despite current recommendations [6], illustrating their low level of evidence. Whether targeted strategies, including the preexposure administration of nirsevimab in addition to RSV vaccination [33], can prevent hospitalization and death in this subset of patients should be assessed prospectively. Antivirals such as ziresovir, which has recently been shown to reduce the symptoms of bronchiolitis in infants hospitalized with RSV infection [34], could be evaluated

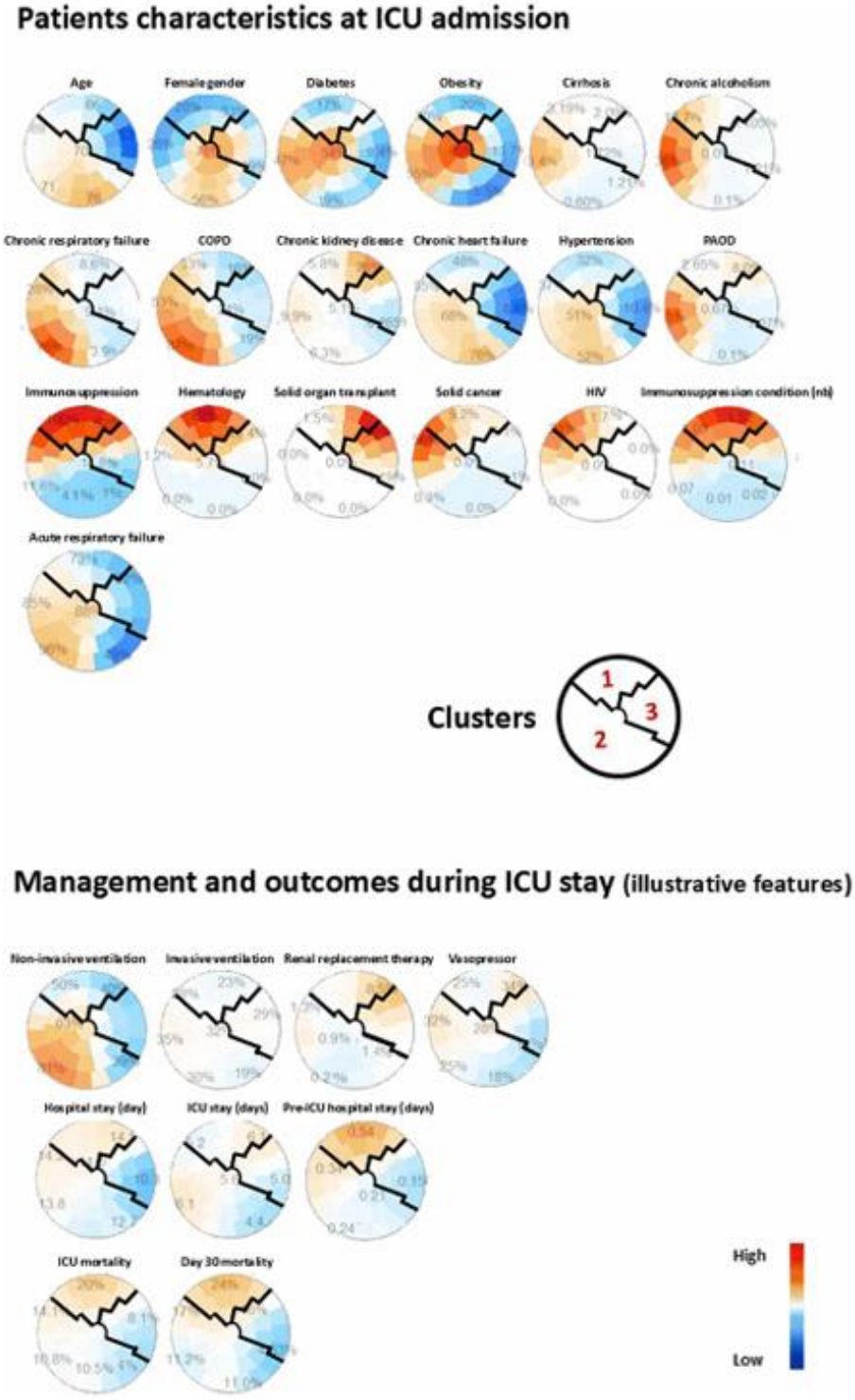


Figure 2. Results from the unsupervised clustering analysis by self-organized maps (SOMs). Unsupervised analysis by SOM automatically located patients with similar clinical and paraclinical parameters within 1 of 40 small groupings ("districts") throughout the map. Patients are in the same location on each map; the more similar the patients, the closer they are on the map (the number of patients per district is not uniformly distributed). Each individual map shows the mean values or proportions per district for each characteristic: blue indicates the lowest average values and red the highest, with numbers shown for a selection of representative districts in each SOM. Only "Patient characteristics at ICU admission" variables have been used to constitute the SOM. "Management and outcomes during ICU stay" variables are only illustrative features. The values in the circles illustrate the value of the displayed variable in a given district (ie, median for quantitative variables, and percentages for qualitative variables). Close districts were combined to provide 3 suitable clusters of patients. Cluster 1 is located in the upper left area and is mainly composed of immunosuppressed patients (all except those with solid organ transplant [SOT]) with a high in-hospital mortality rate. Cluster 2, in the bottom left area, includes older patients who have more cardiorespiratory comorbidities, an intermediate mortality rate, and more noninvasive ventilation requirement. Finally, cluster 3, corresponding to the area located on the right, is made up of younger patients who, apart from a high prevalence of SOT, present fewer comorbidities, with a low mortality rate; Abbreviations: COPD, chronic obstructive pulmonary disease; Hematology, hematologic malignancy and/or allogeneic hematopoietic stem cell transplantation; HIV, human immunodeficiency virus; ICU, intensive care unit; nb, number ; PAOD, peripheral arterial obstructive disease.

Table 4. Main Characteristics of Respiratory Syncytial Virus–Infected Patients According to Clusters Identified by Self-organized Maps

Characteristic	Total (n = 474)	Cluster 1 (n = 98)	Cluster 2 (n = 202)	Cluster 3 (n = 174)
Age, y	65 ± 17	67 ± 13	71 ± 11	57 ± 21
Female sex	209 (44)	25 (26)	113 (56)	71 (41)
BMI ≥30 kg/m²	154 (32)	27 (28)	103 (51)	24 (14)
Diabetes	135 (28)	22 (22)	83 (41)	30 (17)
Hypertension	243 (51)	49 (50)	138 (68)	56 (32)
Peripheral arterial obstructive disease	30 (6)	4 (4)	21 (10)	5 (3)
Chronic heart failure	174 (37)	34 (35)	103 (51)	37 (21)
COPD	189 (40)	36 (37)	122 (60)	31 (18)
Chronic respiratory failure	109 (23)	19 (19)	75 (37)	15 (9)
Cirrhosis	16 (3)	5 (5)	8 (4)	3 (2)
Chronic alcoholism	32 (7)	10 (10)	21 (10)	1 (0.5)
Chronic kidney disease	34 (7)	4 (4)	16 (8)	14 (8)
Chronic hemodialysis	23 (5)	1 (1)	6 (3)	16 (9)
Immunosuppression	162 (34)	98 (100)	17 (8)	47 (27)
Solid cancer	40 (8)	36 (37)	4 (2)	0
Solid organ transplant[a]	34 (7)	0	0	34 (19)
Hematologic disease	61 (13)	52 (53)	0	9 (5)
Allogeneic HSCT	20 (4)	17 (17)	0	3 (2)
Neutropenia	14 (3)	0	3 (1)	11 (6)
HIV infection	14 (3)	14 (14)	0	0
Other[b]	13 (3)	0	10 (5)	3 (2)
No. of immunosuppressive conditions	0 (0–1)	2 (1–3)	0	1 (0–1)
Acute respiratory failure	355 (75)	75 (77)	184 (91)	96 (55)
SAPS II (n = 105)	38 ± 17	37 ± 15	38 ± 16	38 ± 21
Coinfections	107 (23)	21 (21)	44 (22)	42 (24)
Bacteria	95 (20)	19 (19)	40 (20)	36 (21)
Virus	18 (4)	3 (3)	5 (2)	10 (6)
Fungus	2 (0.4)	0	0	2 (1)
Blood leukocytes, g/L (n = 388)	11 (8–15)	10 (5–14)	11 (8–14)	11 (8–16)
Neutrophils, g/L (n = 369)	8.4 (5–12)	5.7 (4–10)	8.4 (6–11)	8.5 (5–14)
Lymphocytes, g/L (n = 366)	1 (0.6–1.6)	0.9 (0.6–1.6)	0.9 (0.6–1.5)	1 (0.6–1.6)
Serum creatinine, μM (n = 328)	92 (65–148)	97 (71–165)	87 (63–130)	90 (63–165)
Hemoglobin, g/L (n = 388)	12 (11–14)	11 (10–13)	12 (11–14)	12 (10–14)
Noninvasive mechanical ventilation	263 (55)	49 (50)	148 (73)	66 (38)
Invasive mechanical ventilation	131 (28)	25 (26)	65 (32)	41 (24)
Vasopressor support	113 (24)	24 (24)	52 (26)	37 (21)
Renal replacement therapy	11 (2)	2 (2)	2 (1)	7 (4)
Antibiotics	392 (83)	83 (85)	173 (86)	136 (78)
Corticosteroids	245 (52)	55 (56)	110 (54)	80 (46)
Ribavirin	15 (3)	6 (6)	1 (0.5)	8 (5)
Intravenous immunoglobulin	7 (1)	2 (2)	0	5 (3)
Duration of hospital stay, d	13 (8–23)	14 (9–28)	14 (8–24)	11 (7–20)
Pre-ICU hospitalization stay, d	0 (0–0)	0 (0–1)	0 (0–0)	0 (0–0)
Duration of ICU stay, d	5 (3–11)	5 (2–11)	6 (3–12)	5 (2; 10)
In-ICU mortality	52 (11)	18 (18)	22 (11)	12 (7)
In-hospital mortality	62 (13)	20 (20)	26 (13)	16 (9)
Day 30 mortality	65 (14)	21 (21)	28 (14)	16 (9)

Data are presented as No. (%), mean ± standard deviation, or median (interquartile range). Two-tailed *P* values come from unadjusted comparisons using χ^2 or Fisher exact tests for categorical variables, and *t* tests or Mann-Whitney tests for continuous variables, as appropriate. No adjustment for multiple comparisons was performed. Bolded *P* values are significant at the $P < .05$ level.

Abbreviations: BMI, body mass index; COPD, chronic obstructive pulmonary disease; HSCT, hematopoietic stem cell transplantation; HIV, human immunodeficiency virus; ICU, intensive care unit; SAPS II, Simplified Acute Physiology Score II.

[a]Solid organ transplant: kidney (n = 14), heart (n = 9), lung (n = 7), liver (n = 4).

[b]Other: treatment with antitumor, immunosuppressive drugs (n = 8); treatment with glucocorticoids and disease-modifying antirheumatic drugs (n = 7).

in this high-risk population. Cluster 2 consisted of older and highly comorbid patients, who should be the target of the next RSV vaccination campaign. Finally, cluster 3 included young patients, with fewer comorbidities and a relatively low mortality rate. The surprising location of SOT patients in this good outcome cluster may be related to the following points: First, these patients had particularly few comorbidities compared with other immunosuppressed patients; second, the excess mortality associated with RSV infection in SOT recipients has mainly been described in lung transplant recipients [35], who represented a small proportion of our patients (n = 7, ie, 20% of SOT recipients and 4% of cluster 3). For the large proportion of patients in this cluster with no apparent risk factor for severe RSV infection, including 9 of the 16 patients who were dead at day 30, we cannot rule out an underlying genetic susceptibility or interferonopathy, as has been reported [36]. Further studies are needed to explore the phenotype of these patients and determine whether they have a specific endotype.

Our study has several limitations. First, the observational and retrospective nature of the study may introduce bias. In particular, patient selection was based on *ICD-10* codes, which may be imperfect. This is because we did not have direct access to the virology database, due to software incompatibility. Second, the coding of some variables, based on the natural language processing method and *ICD-10* codes, may have underestimated certain data, such as coinfections (eg, viral coinfections) or ribavirin administration. The association between "acute respiratory failure" and mortality may reflect a lack of coding of the least severe patients or the fact that some patients were admitted to the ICU with RSV infection and nonrespiratory reasons. The lack of data reflecting patient severity on admission (few SAPS II data and no timing of organ support available) is also contentious. Nevertheless, it should be noted that the median SAPS II, when available, was consistent with recently published data [10] and that the aim of the study was to establish clinical phenotypes based on the intrinsic characteristics of the patients on admission. Also, the lack of details regarding the mechanism of acute respiratory failure (ie, bronchiolitis, pneumonia, or cardiogenic acute pulmonary edema) associated with RSV infection prevents the establishment of a direct causality between RSV infection and the outcomes measured, and we could speculate that some patients had acute respiratory failure or died due to coinfections or underlying diseases. Finally, we acknowledge that care practices (ie, indications for ventilation support, local guidelines for the management of RSV infection) may have differed between centers.

Our study also has several strengths. To our knowledge, this is the largest study to date of patients admitted to the ICU with RSV infection. It was a multicenter study, including a large number of patients during a limited period of time, during which care practices are unlikely to have changed significantly. We combined supervised and unsupervised statistical approaches, which allowed us to identify 3 homogeneous patient phenotypes that are relevant not only for patient care, but also for selecting the population that should be targeted by new anti-RSV therapeutics.

CONCLUSIONS

This large multicenter retrospective study showed that in patients admitted to the ICU with RSV infection, immunosuppression was consistently associated with all-cause day 30 mortality. Three homogeneous patient clusters were created based on different patient comorbidity profiles and were associated with different outcomes. The effect of anti-RSV therapies aimed at preventing ICU admission should be evaluated in these subgroups of patients.

Supplementary Data

Supplementary materials are available at *The Journal of Infectious Diseases* online (http://jid.oxfordjournals.org/). Supplementary materials consist of data provided by the author that are published to benefit the reader. The posted materials are not copyedited. The contents of all supplementary data are the sole responsibility of the authors. Questions or messages regarding errors should be addressed to the author.

Notes

Acknowledgments. The authors thank all nurses and physicians involved in patient care. The authors would also like to thank the Collège des Enseignants de Médecine Intensive Réanimation d'Ile-de-France as well as all the members of the Prono-RSV study group of the Clinical Data Warehouse of Greater Paris University Hospitals: Hafid Ait-Oufella (Saint-Antoine, Paris), Djillali Annane (Raymond Poincaré, Garches), Elie Azoulay (Saint-Louis, Paris), Pierre Bay (Henri Mondor, Créteil), Lila Bouadma (Bichat, Paris), Rania Bouneb (Raymond Poincaré, Garches), Alain Combes (Pitié-Salpêtrière, Paris), Julien Charpentier (Cochin, Paris), Alexandre Demoule (Pitié-Salpêtrière, Paris), Jacques Duranteau (Bicêtre, Le Kremlin-Bicêtre), Stéphane Gaudry (Avicenne, Bobigny), Anatole Harrois (Bicêtre, Le Kremlin-Bicêtre), Marc Humbert (Bicêtre, Le Kremlin-Bicêtre), Charles-Edouard Luyt (Pitié-Salpêtrière, Paris), Bruno Mégarbane (Lariboisière, Paris), Virginie Maxime (Raymond Poincaré, Garches), Nathalie Marin (Cochin, Paris), Jean-Paul Mira (Cochin, Paris), Pierre Moine (Raymond Poincaré, Garches), Nicolas Mongardon (Henri Mondor, Créteil), Xavier Monnet (Bicêtre, Le Kremlin-Bicêtre), Philippe Montravers (Bichat, Paris), Jean-Damien Ricard (Louis Mourier, Colombes), Damien Roux (Louis Mourier, Colombes), Laurent Savale (Bicêtre, Le Kremlin-Bicêtre), Romain Sonneville (Bichat, Paris), Sébastien Tanaka (Bichat, Paris), Jean-François Timsit (Bichat, Paris),

Fabrice Uhel (Louis Mourier, Colombes), Clara Vigneron (Cochin, Paris), Guillaume Voiriot (Tenon, Paris), and Lara Zafrani (Saint-Louis, Paris).

Author contributions. A. G., N. O., R. L., and N. d. P. designed the study. A. G. and R. L. performed statistical analyses. A. G., N. O., R. L., and N. d. P. wrote the first draft of the article. N. d. P. supervised the study. All authors validated the results of the study, reviewed, edited, and approved the article. The corresponding authors attest that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. N. d. P. is the guarantor.

Ethics approval and patient consent. This cohort study was approved by the Institutional Review Board of the AP-HP Data Warehouse (approval number CSE-20-83), and patients included in this study were informed and did not express any opposition to participating in the study. The AP-HP Data Warehouse was approved by the National Committee of Informatics and Liberties.

Financial support. None.

Data availability. The data that support the findings of this study are available on request from the corresponding author (A. G.).

Potential conflicts of interest. S. F. has served as a speaker for GlaxoSmithKline, AstraZeneca, MSD, Pfizer, Cepheid, and Moderna. N. d. P. has served as an advisor or speaker for Moderna and AstraZeneca. All other authors report no potential conflicts of interest.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Collins PL, Crowe JEJ. Respiratory syncytial virus and metapneumovirus. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA, Roizman B, Straus SE, eds. Fields virology. 5th ed. Philadelphie: Lippincott Williams and Wilkins, **2007**:1601–46.
2. Shi T, McLean K, Campbell H, Nair H. Aetiological role of common respiratory viruses in acute lower respiratory infections in children under five years: a systematic review and meta-analysis. J Glob Health **2015**; 5:010408.
3. Falsey AR, Hennessey PA, Formica MA, Cox C, Walsh EE. Respiratory syncytial virus infection in elderly and high-risk adults. N Engl J Med **2005**; 352:1749–59.
4. Nguyen-Van-Tam JS, O'Leary M, Martin ET, et al. Burden of respiratory syncytial virus infection in older and high-risk adults: a systematic review and meta-analysis of the evidence from developed countries. Eur Respir Rev **2022**; 31: 220105.
5. Shi T, Denouel A, Tietjen AK, et al. Global disease burden estimates of respiratory syncytial virus–associated acute respiratory infection in older adults in 2015: a systematic review and meta-analysis. J Infect Dis **2020**; 222:S577–83.
6. Chemaly RF, Shah DP, Boeckh MJ. Management of respiratory viral infections in hematopoietic cell transplant recipients and patients with hematologic malignancies. Clin Infect Dis **2014**; 59(Suppl 5):S344–51.
7. Falsey AR, Formica MA, Hennessey PA, Criddle MM, Sullender WM, Walsh EE. Detection of respiratory syncytial virus in adults with chronic obstructive pulmonary disease. Am J Respir Crit Care Med **2006**; 173:639–43.
8. Walsh EE, Falsey AR, Hennessey PA. Respiratory syncytial and other virus infections in persons with chronic cardiopulmonary disease. Am J Respir Crit Care Med **1999**; 160:791–5.
9. Ackerson B, Tseng HF, Sy LS, et al. Severe morbidity and mortality associated with respiratory syncytial virus versus influenza infection in hospitalized older adults. Clin Infect Dis **2019**; 69:197–203.
10. Coussement J, Zuber B, Garrigues E, et al. Characteristics and outcomes of patients in the ICU with respiratory syncytial virus compared with those with influenza infection: a multicenter matched cohort study. Chest **2022**; 161:1475–84.
11. Falsey AR, Walsh EE, House S, et al. Risk factors and medical resource utilization of respiratory syncytial virus, human metapneumovirus, and influenza-related hospitalizations in adults—a global study during the 2017–2019 epidemic seasons (Hospitalized Acute Respiratory Tract Infection [HARTI] Study). Open Forum Infect Dis **2021**; 8:ofab491.
12. Surie D, Yuengling KA, DeCuir J, et al. Severity of respiratory syncytial virus vs COVID-19 and influenza among hospitalized US adults. JAMA Netw Open **2024**; 7:e244954.
13. Bajema KL, Bui DP, Yan L, et al. Severity and long-term mortality of COVID-19, influenza, and respiratory syncytial virus. JAMA Intern Med **2025**; 185:324–34.
14. Loubet P, Lenzi N, Valette M, et al. Clinical characteristics and outcome of respiratory syncytial virus infection among adults hospitalized with influenza-like illness in France. Clin Microbiol Infect **2017**; 23:253.e9.
15. FDA. U.S. FDA approves ABRYSVO, Pfizer's vaccine for the prevention of respiratory syncytial virus (RSV) in older adults. 2023. https://www.pfizer.com/news/press-release/press-release-detail/us-fda-approves-abrysvotm-pfizers-vaccineprevention#:~:text=(NYSE: PFE) announced today, individuals 60 years and olde. Accessed 31 May 2023.
16. AstraZeneca. Beyfortus approved in the EU for the prevention of RSV lower respiratory tract disease in infants. 2022. https://www.astrazeneca.com/media-centre/pressreleases/2022/beyfortus-approved-in-the-eu-for-the-prevention-of-rsv-lower-respiratorytract-disease-in-infants.html. Accessed 4 November 2022.

17. Mahmood K, Eldeirawi K, Wahidi MM. Association of gender with outcomes in critically ill patients. Crit Care **2012**; 16:R92.
18. Duncan CB, Walsh EE, Peterson DR, Lee FE-H, Falsey AR. Risk factors for respiratory failure associated with respiratory syncytial virus infection in adults. J Infect Dis **2009**; 200:1242–6.
19. Celante H, Oubaya N, Fourati S, et al. Prognosis of hospitalised adult patients with respiratory syncytial virus infection: a multicentre retrospective cohort study. Clin Microbiol Infect **2023**; 29:943.e1–e8.
20. Lee N, Lui GCY, Wong KT, et al. High morbidity and mortality in adults hospitalized for respiratory syncytial virus infections. Clin Infect Dis **2013**; 57:1069–e77.
21. Voiriot G, Visseaux B, Cohen J, et al. Viral-bacterial coinfection affects the presentation and alters the prognosis of severe community-acquired pneumonia. Crit Care **2016**; 20:375.
22. Asner SA, Science ME, Tran D, Smieja M, Merglen A, Mertz D. Clinical disease severity of respiratory viral coinfection versus single viral infection: a systematic review and meta-analysis. PLoS One **2014**; 9:e99392.
23. Vincent J-L, Moreno R, Takala J, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. Intensive Care Med **1996**; 22:707–10.
24. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol **2007**; 165:710–8.
25. Smeden MV, de Groot JAH, Moons KGM, et al. No rationale for 1 variable per 10 events criterion for binary logistic regression analysis. BMC Med Res Methodol **2016**; 16:163.
26. Kohonen T, Somervuo P. How to make large self-organizing maps for nonvectorial data. Neural Netw **2002**; 15:945–52.
27. Gao S, Mutter S, Casey A, Mäkinen V-P. Numero: a statistical framework to define multivariable subgroups in complex population-based datasets. Int J Epidemiol **2019**; 48: 369–74.
28. Kim T, Huh JW, Hong S-B, et al. Epidemiology and characteristics of respiratory syncytial virus pneumonia in critically ill adults. Open Forum Infect Dis **2023**; 10:ofad131.
29. Chen L, Han X, Li Y, Zhang C, Xing X. Comparison of clinical characteristics and outcomes between respiratory syncytial virus and influenza-related pneumonia in China from 2013 to 2019. Eur J Clin Microbiol Infect Dis **2021**; 40:1633–43.
30. Legof J, Zucman N, Lemiale V, et al. Clinical significance of upper airway virus detection in critically ill hematology patients. Am J Respir Crit Care Med **2019**; 199:518–28.
31. Avetisyan G, Mattsson J, Sparrelid E, Ljungman P. Respiratory syncytial virus infection in recipients of allogeneic stem-cell transplantation: a retrospective study of the incidence, clinical features, and outcome. Transplantation **2009**; 88:1222e6.
32. Kim Y-J, Guthrie KA, Waghmare A, et al. Respiratory syncytial virus in hematopoietic cell transplant recipients: factors determining progression to lower respiratory tract disease. J Infect Dis **2014**; 209:1195.e204.
33. Langedijk AC, Bont LJ. Respiratory syncytial virus infection and novel interventions. Nat Rev Microbiol **2023**; 21:734–49.
34. Zhao S, Shang Y, Yin Y, et al. Ziresovir in hospitalized infants with respiratory syncytial virus infection. N Engl J Med **2024**; 391:1096–107.
35. Lee KW, Yoo SG, Cho Y, et al. Characteristics of community-acquired respiratory viruses infections except seasonal influenza in transplant recipients and non-transplant critically ill patients. J Microbiol Immunol Infect **2021**; 54:253–60.
36. Córdova-Dávalos LE, Hernández-Mercado A, Barrón-García CB, et al. Impact of genetic polymorphisms related to innate immune response on respiratory syncytial virus infection in children. Virus Genes **2022**; 58:501–14.

Clinical phenotypes and outcomes associated with respiratory syncytial virus infection in critically ill patients: a retrospective multicentre cohort study in Great Paris area hospitals, 2017-2023

Supplementary appendix

Table of contents

Content	Page Number
Table S1: International Classification of Diseases, Tenth Revision (ICD-10) codes used for the diagnosis of patients with RSV infection in patients (n=474) admitted in intensive care units.	**42**
Table S2: ICD-10 codes used for identification of comorbidities in critically ill patients with RSV infection.	**43**
Table S3: Common Classification of Medical Procedures codes used for identification of organ supports in critically ill patients with RSV infection.	**44**
Figure S1: Hospitalisations in intensive care unit for respiratory syncytial virus infection between July 2017 and May 2023. Distribution histogram representing the number of hospitalisations per month for respiratory syncytial virus infection according to vital status at hospital discharge (alive shown in blue, dead shown in orange).	**45**
Figure S2: In hospital mortality of patients admitted to intensive care unit from July 2017 to May 2023 with a RSV. Distribution histogram representing the number/percent vital status at hospital discharge per year (alive shown in blue, dead shown in orange).	**46**

Table S1. International Classification of Diseases, Tenth Revision (ICD-10) codes used for the diagnosis of patients with RSV infection in patients (n=474) admitted in intensive care units.

ICD-10 code	ICD-10 term	Total	Main diagnosis	Significant associated diagnosis
J121	Respiratory syncytial virus pneumonia	317	128	189
J205	Acute bronchitis due to respiratory syncytial virus	143*	36	106
J210	Acute bronchiolitis due to respiratory syncytial virus	36	13	23
B974	Respiratory syncytial virus as the cause of diseases classified elsewhere	74	1	73

*1 code was not categorized; 83 hospital stays had 2 RSV codes, 5 had 3 codes and 1 had 4 codes.

Table S2: International Classification of Diseases, Tenth Revision (ICD-10) codes used for identification of comorbidities in critically ill patients with RSV infection.

Comorbidities	ICD-10 codes
Chronic respiratory failure	J961
Chronic obstructive pulmonary disease	J44
Peripheral arterial obstructive disease	I7020
Hypertension	I10, I15
Diabetes	E10-E14
Chronic heart failure	I50
Obesity	E66
Cirrhosis	K70.0, K70.3, K71.7, K74, K743, K744, K745, K746, K76.1
Chronic alcoholism	F10, F101, F1024, F1025, F1026, K70, T51
Chronic hemodialysis	Z992
Chronic kidney disease	N18.4, N18.5, Z49.2
Solid cancer	C00-C97, D37-D39, D41-D44
Solid organ transplant	Z940, Z941, Z942, Z943, Z944, Z948
Hematologic malignancy	C81-C96, D45-D47
Allogeneic Hematopoietic Stem Cell Transplantation	Z9480, Z94800, Z94802, Z94803, Z94804, Z94809
Neutropenia	D70
Human Immunodeficiency Virus infection	B20-B24, O98.7, R75, Z21
Other immunosuppression	D80-D89, T380, Y420, Y43.4

Table S3: Common Classification of Medical Procedures codes used for identification of organ supports in critically ill patients with RSV infection.

Endpoints	ICD-10 codes	CCAM
Non-invasive mechanical ventilation	-	GLLD002, GLLD003, GLLD012, GLLD019
Invasive mechanical ventilation	-	GLLD004, GLLD006, GLLD008, GLLD013, GLLD015
Vasopressor support	-	EQLF001, EQLF003, EQLF005
Renal replacement therapy	Z49	JVJB002, JVJF005, JVJF006, JVJF007

Abbreviations: CCAM: Classification Commune des Actes Médicaux or Common Classification of Medical Procedures ; ICD-10: International Classification of Diseases, Tenth Revision

Figure S1: Intensive care unit admissions for respiratory syncytial virus infection between July 2017 and May 2023. Distribution histogram representing the number of hospital admissions per month for respiratory syncytial virus infection according to vital status at hospital discharge (alive patients shown in blue, dead patients shown in orange).

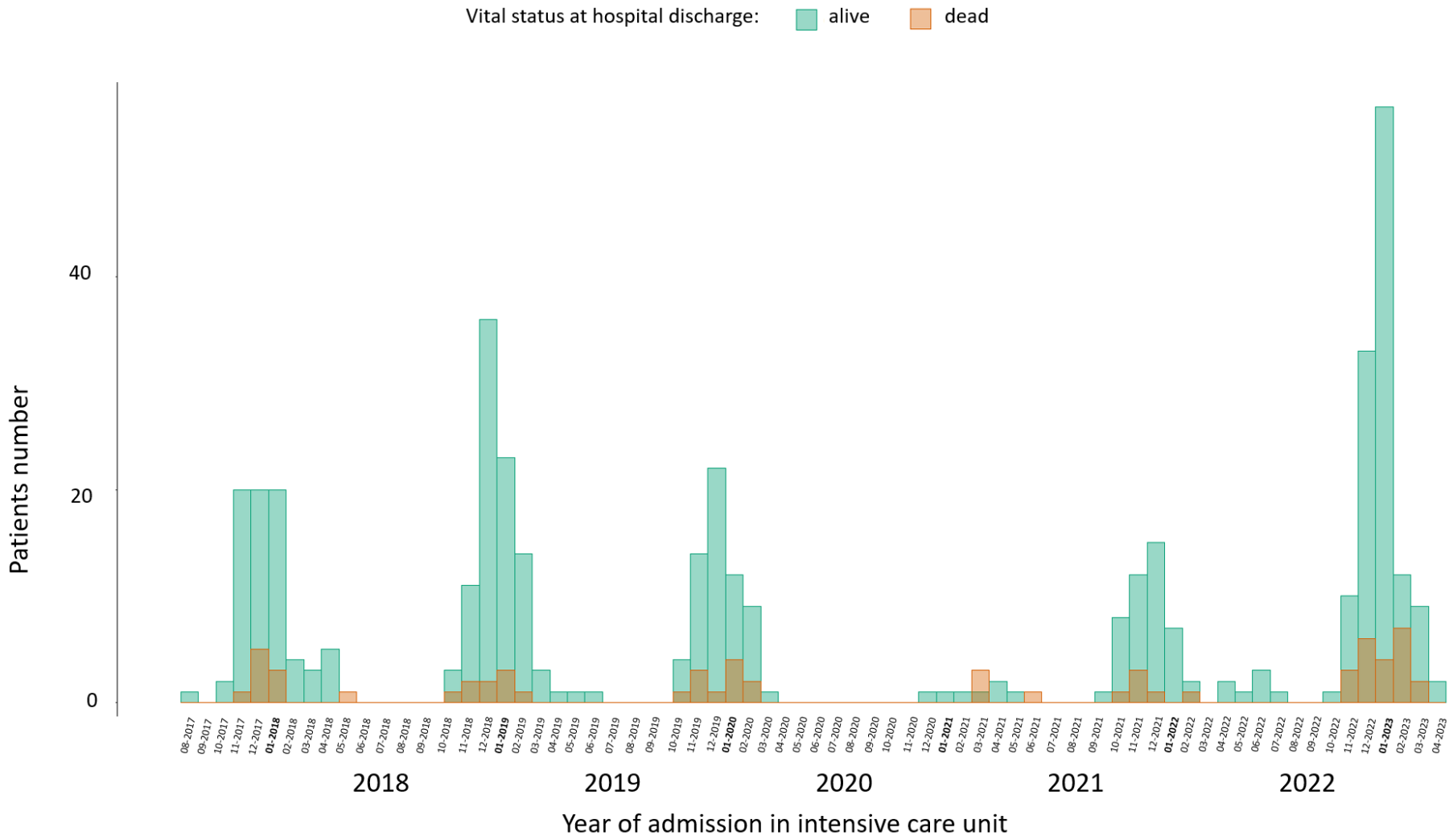
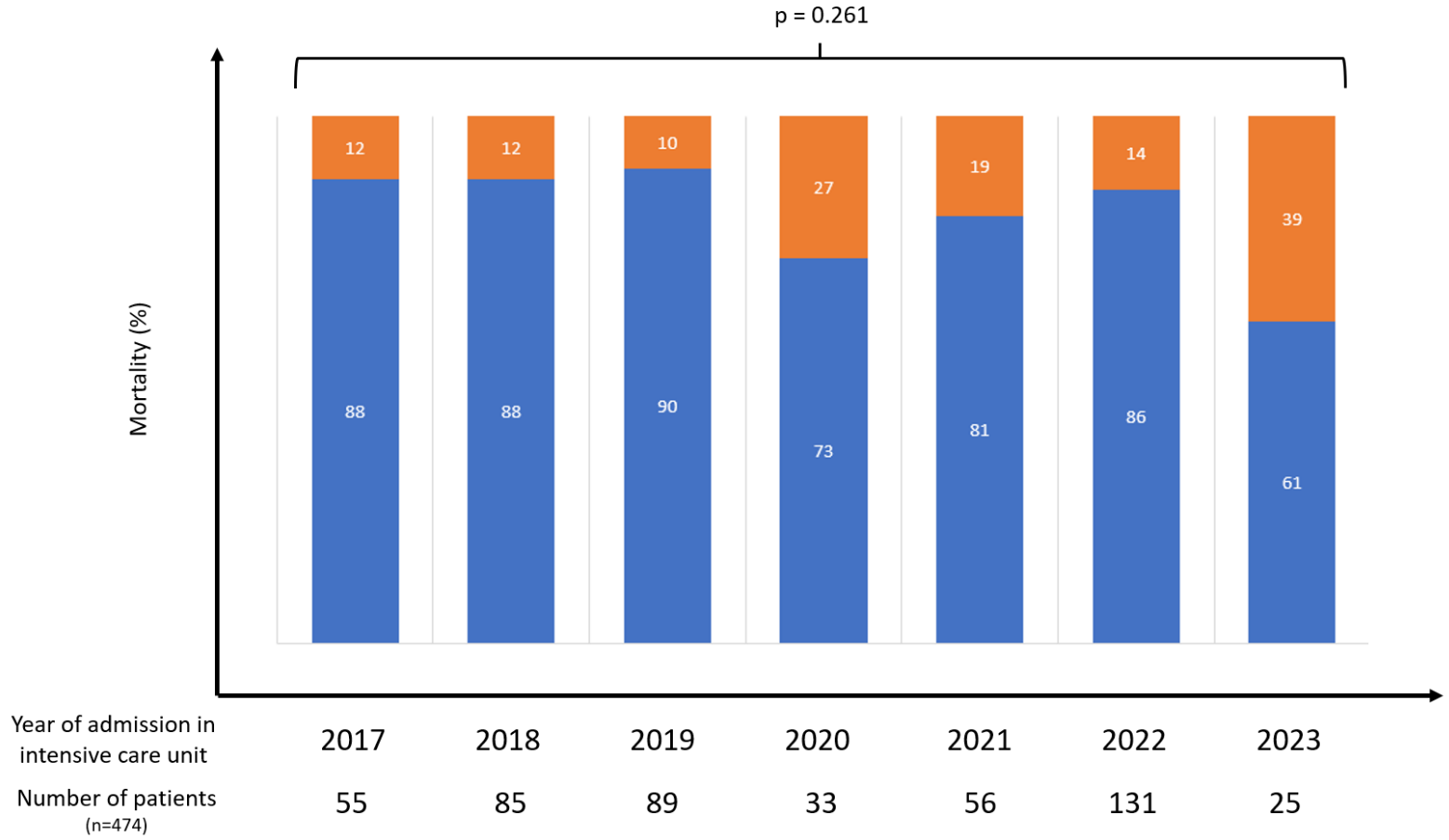


Figure S2: In hospital mortality of patients admitted to intensive care unit from July 2017 to May 2023 with a RSV. Distribution histogram representing the number/percent vital status at hospital discharge per year (alive patients shown in blue, dead patients shown in orange).



B. Article 2

Management, outcomes, and predictors of intensive care unit admission in DRESS: A retrospective multicenter cohort study in Great Paris area hospitals

Antoine Gaillet, Richard Layese, Nadia Oubaya, Emilie Sbidian, Vincent Descamps, Elie Azoulay, Jean-David Bouaziz, Maxime Touron, Nicolas Dupin, Armand Mekontso Dessap, Saskia Ingen-Housz-Oro, Nicolas de Prost for the ICUDRESS study group

En cours de soumission

Rationnel :

Le *Drug Reaction with Eosinophilia and Systemic Symptoms* (DRESS) est une réaction cutanée médicamenteuse sévère à médiation lymphocytaire T, classée parmi les toxidermies graves, dont la prévalence est estimée à environ 2 cas pour 100 000 individus. Elle survient typiquement après une longue période de latence suivant l'exposition médicamenteuse. Les principaux agents en cause sont les anticonvulsivants aromatiques, suivis de l'allopurinol, des antibiotiques de type sulfamides et d'autres molécules comme la vancomycine. Le tableau clinique associe une éruption cutanée diffuse, une fièvre, des anomalies hématologiques (dont une éosinophilie) et une atteinte viscérale variable, allant d'élévations modérées des enzymes hépatiques à des complications cardiaques ou pulmonaires engageant le pronostic vital. La mortalité globale est estimée entre 1,2 % et 7,1 %, mais elle peut atteindre 10 à 47 % en réanimation selon les séries. L'hétérogénéité des présentations rend nécessaire une prise en charge individualisée, d'autant qu'aucun critère validé de gravité (admission en réanimation, risque de décès) n'est actuellement disponible. Un score de sévérité a récemment été proposé,

mais il reste à valider en pratique clinique. Dans ce contexte, une étude multicentrique rétrospective a été menée afin d'identifier les facteurs de risque associés à l'admission en réanimation pour DRESS, avec pour objectifs secondaires l'évaluation de la mortalité hospitalière et la description des stratégies thérapeutiques intra et extra-réanimation.

Matériels et méthodes :

Nous avons mené une étude de cohorte multicentrique, rétrospective, incluant tous les patients ≥18 ans hospitalisés dans les établissements de l'AP-HP entre juillet 2017 et janvier 2023 avec un diagnostic confirmé de DRESS, identifiés via l'EDS AP-HP (39 hôpitaux). En l'absence de code CIM-10 spécifique du DRESS, l'identification reposait sur une stratégie en trois étapes : sélection de codes CIM-10 pertinents (D72.1 (*Eosinophilia*), T78.4 (*Allergy*), T88.7 (*Adverse drug reaction*), ou L27.0 (*Generalized rash due to drugs*), n=13 881), recherche textuelle du terme « dress » dans les comptes rendus (n=724), puis validation experte avec un score RegiSCAR ≥3 (n=235). Les données recueillies à partir du dossier médical informatisé comprenaient les caractéristiques démographiques et les comorbidités, les scores cliniques (SOFA, RegiSCAR, score de sévérité DRESS), les paramètres biologiques, les atteintes viscérales confirmées histologiquement ou par un faisceau clinique, biologique et radiologique, les traitements (notamment corticothérapie) et les modalités de suppléance d'organe (ventilation, épuration extrarénale, vasopresseurs). Le critère de jugement principal était l'admission en réanimation. Les critères secondaires incluaient la mortalité hospitalière ainsi que les interventions thérapeutiques réalisées en réanimation ou en service conventionnel. Les facteurs de risque d'admission en réanimation liée au DRESS ont été évalués à l'aide de modèles de régression logistique univariée et multivariée, en ajustant sur les facteurs de confusion de base sélectionnés a priori, et en excluant de l'analyse les patients ayant développé un DRESS acquis en réanimation. Une analyse de sensibilité a été réalisée en

incluant uniquement les cas probables ou certains de DRESS (c'est-à-dire présentant un score RegiSCAR ≥ 4). Cette étude de cohorte a été approuvée par le comité scientifique et éthique de l'EDS AP-HP (numéro d'approbation CSE-22-22_ICUDRESS). L'Entrepôt de Données de Santé de l'AP-HP a par ailleurs été validé par la Commission Nationale Informatique et Libertés (CNIL).

Résultats :

Au total, 235 cas de DRESS issus de 39 centres ont été inclus, parmi lesquels 37 patients (16 %) ont été admis en réanimation, auxquels s'ajoutaient 7 cas (3 %) de DRESS acquis en réanimation. L'âge médian des patients était de 60 ans [IQR 54–66], avec une prédominance féminine (n = 139, 59 %) et un score médian de comorbidités de Charlson de 3 [2–4]. Les antiépileptiques constituaient la classe pharmacologique la plus souvent en cause (n = 47, 20 %). L'atteinte hépatique représentait la complication viscérale la plus fréquente (n = 161, 69 %), suivie de l'atteinte rénale (n = 65, 28 %), de l'instabilité hémodynamique (n = 14, 6 %), de l'atteinte cardiaque (n = 11, 5 %) et pulmonaire (n = 6, 3 %). La sévérité du DRESS était classée comme légère, modérée et sévère chez 40 %, 35 % et 24 % des patients respectivement. Tous les patients ont reçu des corticoïdes, administrés localement dans deux tiers des cas et par voie systémique dans un tiers des cas. Globalement, 12 patients (5 %) sont décédés au cours de leur hospitalisation, dont 7 (16 %) parmi ceux admis en réanimation. En analyse multivariée, les deux seuls facteurs indépendamment associés à un risque accru d'admission en réanimation étaient le caractère nosocomial du DRESS (OR ajusté, ORa = 4,09 ; IC95 % [1,66–10,2] ; p = 0,002) et la sévérité de la maladie (formes modérées : ORa = 2,77 ; IC95 % [0,81–11,1] ; formes sévères : ORa = 3,2 ; IC95 % [4,15–52,3] ; p < 0,001). Parmi les 44 patients pris en charge en réanimation, un tiers (34 %) a nécessité un traitement par vasopresseurs, un quart (27 %) une épuration extrarénale, et 20 % (n=9) une ventilation mécanique invasive. Une transplantation hépatique n'a été réalisée que chez un seul patient. Les rechutes étaient rares

(3 %, n=6/235) et aucune séquelle auto-immune n'a été rapportée. Malgré une surveillance virologique fréquente, les réactivations virales documentées restaient limitées (EBV : 24 %, HHV-6 : 15 %, CMV : 7 %). Les patients admis en réanimation présentaient des durées d'hospitalisation plus longues (31 jours [12–76] vs. 10 jours [6–17], p<0,001) et une mortalité hospitalière plus élevée (16 % vs. 3 %, p<0,001). Les patients décédés étaient significativement plus âgés (75 ans vs. 58 ans), porteurs de plus de comorbidités (score de Charlson 6 vs. 3), et plus souvent atteints de DRESS nosocomial (50 % vs. 17 %), fréquemment lié à l'allopurinol ou aux bêta-lactamines. Ils présentaient moins souvent une atteinte hépatique (42 % vs. 70 %) mais étaient davantage admis en réanimation (67 % vs. 16 %), nécessitant plus fréquemment un soutien d'organes

Conclusion du travail n°2 :

Concernant le DRESS, cette étude rétrospective multicentrique a permis de caractériser la plus large cohorte de patients hospitalisés pour un DRESS, contribuant ainsi à enrichir les connaissances sur cette pathologie rare. Les principaux résultats sont les suivants : (1) les facteurs de risque d'admission en réanimation identifiés en analyse multivariée sont le caractère nosocomial du DRESS et le score de sévérité associé ; (2) 16 % des patients ont nécessité une admission en réanimation, avec une mortalité hospitalière globale de 5 % ; (3) les corticoïdes représentaient le traitement immunosuppresseur quasi exclusif (96 %), privilégiant une administration topique pour les formes légères à modérées et systémique pour les formes sévères, en particulier chez les patients admis en réanimation.

Concernant l'EDS, cette étude cible une situation clinique non couverte par le codage PMSI isolément. Il s'agit par ailleurs d'une maladie rare pour laquelle il existe un intérêt certain à exploiter un entrepôt aussi vaste que l'EDS AP-HP. Ainsi, le screening initial des patients a été fait via les codages PMSI les plus proches (sur concertation d'experts de la pathologie, en

diagnostic principal ou relié, au sein d'un séjour hospitalier). Ensuite, l'accès aux comptes-rendus a permis la recherche textuelle du terme « dress ». Les dossier médicaux électroniques des 724 patients ainsi sélectionnés ont finalement été revus individuellement afin de valider ou non le diagnostic de DRESS (c'est-à-dire ne retenant que les patients avec un score RegiSCAR ≥ 3), aboutissant à la cohorte finale de 235 patients. Etant donné la subtilité des données à recueillir (médicament imputable, délais d'introduction et d'arrêt des médicaments, imputabilité des atteintes viscérales au DRESS), l'accès aux dossier médical électronique et la taille « raisonnable » de la cohorte, la stratégie de recherche textuelle n'a pas été plus développée dans cette étude. Ce projet EDS a donc permis de constituer « simplement » la plus grande cohorte de DRESS publiée jusque-là (141 DRESS étaient inclus dans une cohorte multicentrique européenne publiée récemment [25]), permettant d'avoir une puissance suffisante pour identifier rigoureusement des facteurs de risque de sévérité.

Management, outcomes, and predictors of intensive care unit admission in DRESS: A retrospective multicenter cohort study in Great Paris area hospitals

Antoine Gaillet[1], Richard Layese[2,3], Nadia Oubaya[2,4], Emilie Sbidian[5,6,7], Vincent Descamps[5,8], Elie Azoulay[9], Jean-David Bouaziz[10], Maxime Touron[11], Nicolas Dupin[12], Armand Mekontso Dessap[1,2,3,13], Saskia Ingen-Housz-Oro[5,6,7], Nicolas de Prost[1,2,3,4,5,13] for the ICUDRESS study group

Affiliations

1. Service de Médecine Intensive Réanimation, Hôpitaux Universitaires Henri Mondor - Albert Chenevier, Assistance Publique-Hôpitaux de Paris (AP-HP), 94010 Créteil, France

2. Univ Paris Est Créteil, INSERM, IMRB, F-94010 Créteil, France

3. Assistance Publique-Hopitaux de Paris (AP-HP), Henri Mondor Hospital, Clinical Research Unit, F-94010 Créteil, France

4. INSERM U955, Team « Viruses, Hepatology, Cancer », Créteil, France

5. Centre de Référence des Dermatoses Bulleuses Toxiques et toxidermies graves TOXIBUL, Hôpitaux Universitaires Henri Mondor, Assistance Publique – Hôpitaux de Paris (AP-HP), Créteil

6. Service de Dermatologie, Hôpitaux Universitaires Henri Mondor, Assistance Publique – Hôpitaux de Paris (AP-HP), Créteil, France

7. EpiDermE, Université Paris Est Créteil Val de Marne, Créteil, France.

8. Service de Dermatologie, Hôpital Bichat, AP-HP, Paris, France.

9. Service de Médecine Intensive Réanimation, Hôpital Saint-Louis, Assistance Publique-Hôpitaux de Paris (AP-HP), Paris, France

10. Service de Dermatologie, Hôpital Saint-Louis, Assistance Publique-Hôpitaux de Paris (AP-HP), Paris, France

11. Service de Médecine Intensive Réanimation, Hôpital Cochin, Assistance Publique-Hôpitaux de Paris (AP-HP), Paris, France

12. Service de Dermatologie, Hôpital Cochin, Assistance Publique-Hôpitaux de Paris (AP-HP), Paris, France

13. Groupe de Recherche Clinique CARMAS, Université Paris Est-Créteil, 94010 Créteil, France

Corresponding author

Dr Antoine Gaillet

Service de Médecine Intensive Réanimation, Hôpitaux Universitaires Henri Mondor - Albert Chenevier, Assistance Publique-Hôpitaux de Paris (AP-HP), 94010 Créteil, France

Email : antoine.gaillet@aphp.fr : Phone : 0140257951

ABSTRACT

Background: Drug Reaction with Eosinophilia and Systemic Symptoms (DRESS) is a rare but potentially life-threatening hypersensitivity reaction to drugs. Typically, around 15% of patients require admission to an intensive care unit (ICU).

Objective: Identifying factors associated with ICU admission could help optimize the management of this condition.

Methods: Multicenter retrospective cohort study of adult patients with possible to definite DRESS (RegiSCAR score ≥ 3) hospitalized in the Great Paris area hospitals between 01/07/2017 and 01/01/2023. Data were extracted from the Assistance Publique – Hôpitaux de Paris Health Data Warehouse. The primary objective was to identify risk factors associated with DRESS-related ICU admission. Our secondary objectives were to assess the in-hospital mortality rate and report on therapeutic interventions inside and outside the ICU.

Results: A total of 235 DRESS cases from 39 centers were included, of whom 37 patients (16%) were admitted to the ICU, in addition to seven (3%) in-ICU acquired DRESS cases. The median age was 60 years [IQR 54–66], with a female predominance (n=139, 59%) and a median Charlson comorbidity score of 3 [2–4]. Antiepileptic drugs were the most frequently implicated drug class (n=47, 20%). Hepatic involvement was the most frequent visceral complication (n=161, 69%), followed by renal involvement (n=65, 28%), hemodynamic instability (n=14, 6%), cardiac involvement (n=11, 5%), and pulmonary involvement (n=6, 3%). DRESS severity was classified as mild, moderate, and severe in 40%, 35%, and 24% of patients, respectively. All patients received corticosteroids (two-thirds topically, and one-third systemically). Overall, 12 patients (5%) died during their hospitalization, including seven (16%) of those admitted to the ICU. In multivariable analysis, the only two independent factors significantly associated with

an increased risk of ICU admission were hospital-acquired DRESS (adjusted odds ratio, aOR=4.09; 95% CI [1.66–10.2]; p=0.002) and DRESS severity (moderate forms: aOR=2.77; 95% CI [0.81–11.1]; severe forms: aOR=3.2; 95% CI [4.15–52.3]; p<0.001).

Conclusions: This study confirms the high rate of patients managed in the ICU (nearly one in five) among hospitalized patients with DRESS. It also highlights the prognostic impact of hospital-acquired DRESS and the specific severity of the disease.

KEYWORDS

DRESS; hospital-acquired; intensive care unit (ICU)

WORD COUNTS

Abstract: 345 words

Manuscript: 3192 words

Tables: 3

Figure: 1

Supplementary Appendix: 5 Tables and 2 Figures

References: 24

INTRODUCTION

Drug Reaction with Eosinophilia and Systemic Symptoms (DRESS) is a rare T cell–mediated severe cutaneous adverse drug reaction (SCAR) with an estimated prevalence of two cases per 100,000 individuals [1]. It typically occurs after a prolonged latency period following drug exposure [2]. The most strongly implicated drugs class are aromatic anticonvulsants, followed by allopurinol, sulfonamide antibiotics, and other agents such as vancomycin [3].

DRESS is characterized by a variety of symptoms, including a widespread skin rash, fever, hematologic abnormalities, including eosinophilia, and involvement of internal organs. The severity and presentation of organ involvement can vary widely, ranging from mild and transient elevations in hepatic enzymes to severe, life-threatening cardiac or pulmonary complications [4]. Estimated mortality rates range from 1.2% to 7.1% globally [3, 5, 6], but can reach 10% to 47% in intensive care unit (ICU) settings [3, 7, 8], depending on the study. The heterogeneity of this patient population necessitates an individualized management. Notably, reliable criteria for identifying severity (e.g., need for ICU admission, high risk of mortality) remain elusive. A severity score has recently been proposed, but it has yet to be validated in real-life settings and it is unclear whether it could be helpful in identifying patients at higher risk of ICU admission [4]. The primary objective of this multicenter, retrospective cohort study was to identify risk factors associated with DRESS-related ICU admission. Our secondary objectives were to assess the in-hospital mortality rate and report on therapeutic interventions inside and outside the ICU.

METHODS

Study design and population

We conducted a multicenter, retrospective cohort study including all patients aged 18 years and older, admitted to the Assistance Publique – Hôpitaux de Paris (AP-HP) hospitals between July 1st, 2017, and January 1st, 2023, with a diagnosis of DRESS, using the AP-HP Health Data Warehouse. This database aggregates all data from inpatients and outpatients across the 39 hospitals of the AP-HP network in the Greater Paris area (population of 12 million), including socio-demographic information, clinical summaries, laboratory results, medication prescriptions, and medical records with diagnoses and procedures related to hospital stays, as well as vital status at discharge.

Since there is no specific ICD-10 (International Classification of Diseases, Tenth Revision) code for DRESS, we proceeded in three steps to identify DRESS patients (**Figure 1**): 1) we first selected patients coded with D72.1 (Eosinophilia), T78.4 (Allergy), T88.7 (Adverse drug reaction), or L27.0 (Generalized rash due to drugs), in any hospital stay during the study period, using the French national hospital database (Programme de Médicalisation des Systèmes d'Information), as these codes were considered the most relevant for DRESS (n = 13,881); 2) we then performed a textual search for the term "DRESS" in the clinical reports (n = 724); and 3) finally, the diagnosis of DRESS was confirmed by expert review of the medical records and using a RegiSCAR score ≥3 (n = 235).

Ethical considerations

This cohort study was approved by the Institutional Review Board of the AP-HP Data Warehouse (approval # CSE-22-22_ICUDRESS). Patients received information during their hospital stay that data abstracted from their medical charts could be used for research

purposes. The AP-HP Data Warehouse was approved by the "*Commission Nationale Informatique et Libertés*".

Data recorded

The following data were collected on admission and throughout the hospital stay: epidemiological data (age, sex, comorbidities), clinical scores (Sequential Organ Failure Assessment (SOFA) [9], RegiSCAR [10], and DRESS severity scores [4]), and biological parameters. Visceral involvement related to DRESS was confirmed histologically whenever possible; otherwise, it was established in each case based on the combination of clinical symptoms (i.e., dyspnea, electrocardiographic repolarization abnormalities, oliguria), laboratory abnormalities (i.e., liver function tests, serum troponin and creatinine levels), and imaging findings (i.e., pulmonary infiltrates on chest CT scan, echocardiographic cardiac dysfunction), after exclusion of differential diagnoses. Information regarding the culprit drug(s), management setting (including the need for invasive or non-invasive ventilatory support, renal replacement therapy, or vasopressors), treatment modalities (use of corticosteroids), and follow-up (ICU length of stay, relapses, sequelae, and mortality) was extracted from medical records using a standardized data collection form.

Endpoints

The primary endpoint was ICU admission. Secondary endpoints included in-hospital mortality and therapeutic interventions inside and outside the ICU.

Statistical analyses

Categorical variables were presented as numbers (percentages), and quantitative variables as mean ± standard deviation (SD) or median [interquartile range (IQR)], as appropriate.

Characteristics of patients who died in hospital and those who did not were compared using Chi-squared tests or Fisher's exact tests for categorical variables and Student t-tests or Mann-Whitney tests for quantitative variables, as appropriate.

Risk factors for DRESS-related ICU admission were assessed using univariable and multivariable logistic regression models, adjusting for baseline confounders selected *a priori*, and excluding from the analysis patients with ICU-acquired DRESS. Variables affecting too few patients (prevalence ≤1%) and those with uncertain or debated definitions (such as DRESS-related hemodynamic involvement) were not included in the multivariable analysis. A sensitivity analysis was performed including only probable or definite DRESS cases (i.e., having a RegiSCAR score ≥ 4). Adjusted odds ratios (aOR) with their 95% confidence intervals (CI) were calculated. We followed the recommended "5 events per adjustment variable" ratio [11, 12]. There were no missing data among the variables included in the regression model. Likelihood ratio tests were used to explore clinically relevant interactions between variables. Log-linearity, absence of multicollinearity, and presence of outliers were assessed. All statistical tests were two-tailed, and a p-value < 0.05 was considered statistically significant.

All analyses were performed with Python software (version 3.7) and RStudio software, version 4.0.5 (https://www.R-project.org/The R project for Statistical Computing, Vienna, Austria).

RESULTS

Characteristics of patients hospitalized with DRESS

A total of 235 patients were identified as having DRESS. The median age was 60 years [54–66], and 59% (n=139) were female (**Table 1**). The median Charlson comorbidity index was 3 [2–4].

The median time between initiation of the suspected drug and onset of DRESS was 25 days [18–28], and 18% (n=43) of cases were hospital-acquired. A total of 44 patients were managed in the ICU, including 37 DRESS-related ICU admissions and seven cases of DRESS that occurred in patients who were already in the ICU (**Figure 1)**.

The four main drug classes implicated were anticonvulsants (n=47, 20%), allopurinol (n=32, 14%), sulfonamides (n=30, 13%), and beta-lactams (n=29, 12%) (**Supplementary Appendix, Figure S1)**. The main reason for hospital admission was exanthema (n=165, 70%), whereas ICU admission (in 16% of cases) was mainly prompted by sepsis (n=15, 40%) or non-infectious respiratory failures (n=13, 35%). In addition to nearly constant exanthema (97% of cases), facial edema was observed in 60% (n=140) of cases. The liver and kidneys were the most commonly affected visceral organs, at 69% (n=161) and 28% (n=65), respectively. Meanwhile, cardiac and pulmonary involvement were reported in only 5% (n=11) and 3% (n=6) of cases. Histological confirmation supported the diagnosis of the visceral involvement in 13 patients (kidney (n=1); liver (n=11); heart (n=1), see **Supplementary Appendix, Figure S2)**. The median RegiSCAR score was 6 [4-7], meaning that the majority of patients (n=214, 91%) had probable/definite DRESS (i.e., RegiSCAR scores >3) (**Table 1 and Supplementary Appendix, Table S1**). Patients were almost evenly distributed across the three DRESS severity score categories (**Supplementary Appendix, Table S2**).

Factors associated with intensive care unit admission

Univariable logistic regression analysis revealed that hospital-acquired DRESS (OR 4.15, 95% CI [1.87;9.12], p<0.001), DRESS induced by beta-lactam antibiotics (OR 2.64, 95% CI [1.01;6.40], p=0.037), renal (OR 2.76, 95% CI [1.30;5.83], p=0.008) or hemodynamic involvement (OR 20.5, 95% CI [6.34;79.6], p<0.001), higher eosinophil counts (per

eosinophil/mm³, OR 1.01, 95% CI [1.00; 1.02], p=0.041), and higher DRESS severity scores (p<0.001) were significantly associated with ICU admission (**Table 3**). Conversely, hospital admission for exanthema (OR 0.06, 95% CI [0.02; 0.15], p<0.001) and the presence of facial edema (OR 0.42, 95% CI [0.19; 0.87], p=0.020) were protective factors.

In multivariable analysis, only hospital-acquired DRESS (aOR 4.09, 95% CI [1.66; 10.2], p=0.002) and the DRESS severity score (moderate grade: aOR 2.77, 95% CI [0.81; 11.1]; severe grade: aOR 13.2, 95% CI [4.15; 52.3], p<0.001) remained independently associated with ICU admission. Risk factors for ICU admission remained consistent even when patients with a RegiSCAR score of 3 (i.e., possible DRESS) were excluded (**Supplementary Appendix, Table S3**).

Hospital mortality, therapeutic interventions and other outcomes

Twelve patients (5%) died during hospitalization, including seven of the 44 patients admitted to the ICU (16%).

Nearly all patients were treated with corticosteroids (n=225/235, 96%), with a predominance of topical corticosteroids for mild (n=75/95, 79%) and moderate cases (n=50/83, 60%), and systemic corticosteroids for severe forms (n=34/57, 60%), notably among patients admitted to the ICU (n=22/44, 50% vs. n=55/191, 29%, for non-ICU patients, p <0.001) (**Supplementary Appendix, Table S2**). The median duration of treatment was 30 days [30; 90]. Other therapeutic strategies were only used anecdotally (i.e., one patient received intravenous immunoglobulins, none received anti-IL-5 monoclonal antibodies or cyclosporine, and six patients received antivirals, **Table 2**).

Among the 44 patients managed in ICU, one third (n=15/44, 34%) required vasopressors, one quarter (n=12/44, 27%) underwent renal replacement therapy, and 20% (n=9/44) required invasive mechanical ventilation support. Only one patient required a liver

transplantation. Finally, very few relapses were observed (n=6/235, 3%), and no autoimmune sequelae were reported. Notably, 16 cases of bacteremia were diagnosed during the acute phase of DRESS, mainly in ICU (n=12) and of cutaneous origin (n=11; including 6 catheter-associated bacteremias), involving Enterobacterales (n=12, *E. cloacae* (n=5), *E. cloacae* (n=3, all ESBL), *P. aeruginosa* (n=2) and *E. coli* (n=2)) and Gram-positive cocci (n=10, *S. aureus* (n=3), *coagulase-negative staphylococci* (n=2), *E. faecalis* (n=4), *S. oralis* (n=1)). Despite frequent testing, few viral reactivations were documented during the acute phase (EBV: 24%, HHV-6: 15%, CMV: 7%).

Notably, ICU-admitted patients had significantly longer hospital stays (31 days [12; 76] vs. 10 days [6; 17], p<0.001) and higher in-hospital mortality rates (6/44 [16%] vs. 6/191 [3%], p<0.001) than others.

Patients who died during their hospital stay were older (75 years [68; 78] vs. 58 years [50; 63], p=0.008), had more comorbidities (Charlson score 6 [4; 7] vs. 3 [2; 3], p=0.010), and had more frequently developed a hospital-acquired DRESS (50% vs. 17%, p=0.003), most often following exposure to allopurinol (42% vs. 12%) or beta-lactams (33% vs. 11%) (**Supplementary Appendix, Table S4)**. They were less likely to present with hepatitis (42% vs. 70%, p=0.039) and were more frequently admitted to the ICU (67% vs. 16%, p<0.001), with a higher need for organ support (**Supplementary Appendix, Table S5**).

DISCUSSION

This multicenter retrospective study led to the description of a large cohort of hospitalized patients with DRESS, thereby expanding current knowledge on this rare condition. The main findings of the current work are as follows: 1) Risk factors of ICU admission identified by multivariable logistic regression analysis include hospital-acquired DRESS and the DRESS

severity score; 2) Patients with DRESS required ICU admission in 16% of cases and had an in-hospital mortality of 5%; and 3) Corticosteroids were the predominant, and nearly exclusive, immunosuppressive therapy administered (96%), with a predominance of topical corticosteroids for mild and moderate cases, and systemic corticosteroids for severe forms, notably for ICU admitted patients.

The baseline characteristics of our cohort (age, sex, and comorbidities) were similar to those reported in the literature [6]. The main drug classes implicated were also comparable, namely anticonvulsants, sulfonamides, beta-lactams, and allopurinol, with a median latency period of approximately 25 days between drug initiation and DRESS onset, consistent with prior findings [6]. The DRESS-related ICU admission rate in our cohort (16%) was similar to that reported in recent studies. However, a key finding of our study is the identification of two independent factors that are associated with ICU admission: hospital-acquired DRESS and the recently proposed DRESS severity score [4]. To our knowledge, hospital-acquired onset has not previously been examined as a prognostic factor, and despite being derived from expert consensus, the severity score has not yet been validated in clinical practice. Our study reinforces these criteria by that ICU admission is more likely to be required for the most severe cases.

A noteworthy and previously underreported finding is the significant proportion (18%) of hospital-acquired DRESS cases [2]. This peculiar subgroup of patients may have been indirectly captured in earlier cohorts through the variable beta-lactam antibiotics (identified as prognostic in the study by Kridin et al. [6]), which in our analysis showed collinearity with hospital-acquired cases.

Although comparative data remains limited due to the recent development of the severity score, its distribution in our cohort mirrored that of Kridin et al. [6], with a similar ICU admission rate, suggesting a potential correlation. However, due to its collinearity with visceral involvement, adjustment for individual organ dysfunctions was not feasible. However, ICU admission was predominantly driven by visceral complications rather than cutaneous involvement.

Pulmonary complications were often multifactorial in origin, including bronchospasm (with or without eosinophilic pneumonia), pulmonary edema (cardiogenic or nephrogenic), and cardiovascular dysfunction, such as myocarditis or vasoplegic hypotension that mimicked septic shock. Causality with DRESS was difficult to establish in such cases due to the absence of documented infection and the presence of fluid-refractory hypotension. This presentation aligns with the findings of Kimmoun et al. [8], in which 71% of patients required vasopressors. Although pulmonary and cardiac involvements have historically been reported as frequent (approximately 30% and 20%, respectively) [13, 14], their prevalence appears to be lower in more recent cohorts [6, 15], including ours. Notably, such complications are classically associated with minocycline exposure, a drug that is absent from our dataset, probably due to very restrictive indications nowadays. Renal and hepatic involvements, while predominant as in previous reports [5, 6], were mostly subclinical and rarely justified ICU admission on their own. This aligns with prior findings indicating that hepatic injury in DRESS is typically mild, although it can occasionally exceed the upper limit of normal by up to tenfold and, in rare cases, require liver transplantation, as observed in our series [16, 17]. Furthermore, a literature review from 2023 by Dagnon da Silva et al. [18], involving 71 patients, reported favorable renal outcomes despite a high rate of renal replacement therapy (30%), which is similar to that experienced by ICU patients in our cohort.

Although viral reactivation has been suggested as a prognostic marker [19, 20] and is now recommended for screening at diagnosis [4], we excluded it from multivariate models due to missing data and heterogeneity in the timing and frequency of measurements. The timing of sampling and asynchronous viral kinetics (e.g., HHV-6 typically peaking 2–3 weeks after onset) likely contribute to variability in reported prevalence (34% in the study of Kridin et al. vs. 77% in the study of Marcombes et al. [6, 15]).

The mortality observed in our cohort falls within the previously reported range [3, 5-8], reaching 33% (n=5/15) among vasopressor-dependent patients. Notably, nearly half of the deaths occurred outside the ICU, often in the context of withholding life-sustaining therapies. Surprisingly, the DRESS severity score was not associated with an increased risk of death. This may be explained by three factors that are partially overlapping: the low overall mortality rate of DRESS, the effective management of patients admitted to the ICU, and the limited statistical power. In contrast, hospital-acquired DRESS was identified as a significant risk factor for in-hospital death, further underscoring the vulnerability of this subgroup. The classic risk factors were also confirmed, including age, burden of comorbidities, use of allopurinol or beta-lactams, and ICU admission with need for organ support [2, 6].

Few clinical trials have been conducted in DRESS due to the rarity of the condition. Nevertheless, an international expert consensus was recently published, providing recommendations on the route, dose, and duration of corticosteroids—which remain the cornerstone of immunosuppressive therapy—as well as on the role of adjunctive treatments [4]. These guidelines were broadly followed in our cohort, with topical corticosteroids used in mild to moderate cases, with the option of escalating to systemic therapy if the condition worsened [21], and systemic corticosteroids administered in severe cases. For the most severe

or corticosteroid-resistant forms of DRESS, second-line therapies may include intravenous immunoglobulins, immunosuppressants such as cyclosporine, antivirals in the context viral replication, and anti–IL-5 or anti–IL-5R monoclonal antibodies [4]. However, as in other recent cohorts, very few patients in our study received these alternative treatments. This is probably due to the high responsiveness of DRESS to corticosteroids and, in the case of anti–IL-5 agents, the recent introduction of these agents to the market, with no current approved indication for DRESS despite promising preliminary data [22, 23].

Our study has several limitations. Firstly, the observational and retrospective design of the study may introduce bias. Especially, the absence of a specific ICD code for DRESS required indirect patient identification. However, diagnoses were confirmed by an expert and supported by the RegiSCAR score. While patients with a RegiSCAR score of 3 (possible DRESS) were included, despite often being excluded from other studies [6], the majority of patients had definite (n=74) or definite/probable (n=140) DRESS (n=214/235, 91% overall), and the variables associated with ICU admission remained consistent when patients with a RegiSCAR score = 3 were excluded from the analysis (**Supplementary Appendix, Table S3**). It is worth highlighting that the RegiSCAR score aims to provide diagnostic guidance. Clinical judgement should remain central to the diagnosis process, particularly following a multidisciplinary review involving dermatologists, pathologists, and organ specialists, as demonstrated in this study. We acknowledge that the lack of standardized follow-up in our study has probably resulted in an underestimation of the relapse and sequelae rates. Indeed, although DRESS may have a prolonged course with potential relapses despite appropriate drug withdrawal and treatment [24], only a few cases were identified in our cohort. Similarly, sequelae and autoimmune complications appeared rare compared to the findings of Kridin et al. [6] (3% vs.

33% for sequelae, and 0% vs. 9% for autoimmune complications), likely reflecting recall or follow-up bias. At last, immunosuppressive management and ICU care were not standardized by nature. Secondly, as with most studies on DRESS, very few visceral biopsies were performed. Consequently, visceral involvement was predominantly inferred based on clinical context, which may introduce subjectivity. Thirdly, we did not have systematic access to allergological investigations, which could have influenced the identification of the culprit drug in some cases. Fourthly, almost no genetic data was reported in the medical records, precluding any analysis of the potential prognostic implications for ICU admission. Finally, despite the relatively large sample size, the number of in-hospital deaths was too low to enable robust assessment of risk factors for mortality using regression analysis.

Our study also has several strengths. It reinforces previous epidemiological findings on the largest cohort of hospitalized DRESS patients to date, including those admitted to the ICU. It also identifies new vulnerability factors, such as hospital-acquired DRESS and the severity score, which could be practically implemented in the management strategy of hospitalized DRESS patients to help optimize care.

Conclusion

The in-hospital mortality rate of patients with DRESS was 5%. This study confirms the high rate of patients managed in the ICU (nearly 20%) among hospitalized DRESS patients and highlights the prognostic relevance of hospital-acquired onset and disease severity. These findings underscore the need for enhanced clinical surveillance and individualized therapeutic approaches in this high-risk population.

REFERENCES

1. Wolfson AR, Zhou L, Li Y, Phadke NA, Chow OA, Blumenthal KG. Drug reaction with eosinophilia and systemic symptoms (DRESS) syndrome identified in the electronic health record allergy module. J Al lergy Clin Immunol Pract 2019; 7: 633-40

2. Kroshinsky D, Cardones ARG, Blumenthal KG. Drug Reaction with Eosinophilia and Systemic Symptoms. N Engl J Med. 2024 Dec 12;391(23):2242-2254.

3. Cacoub P, Musette P, Descamps V, et al. The DRESS syndrome: a literature re view. Am J Med 2011; 124: 588-97.

4. Brüggen MC, Walsh S, Ameri MM, Anasiewicz N, Maverakis E, French LE et al. Management of Adult Patients With Drug Reaction With Eosinophilia and Systemic Symptoms: A Delphi-Based International Consensus. JAMA Dermatol. 2024 Jan 1;160(1):37-44.

5. Kardaun SH, Sekula P, Valeyrie-Allanore L, et al. Drug reaction with eosinophilia and systemic symptoms (DRESS): an original multisystem adverse drug reaction: results from the prospective RegiSCAR study. Br J Dermatol 2013; 169: 1071-80. 6.

6. Kridin K, Brüggen M-C, Walsh S, et al. Management and treatment outcome of DRESS patients in Europe: an international multicentre retrospective study of 141 cases. J Eur Acad Dermatol Venereol 2023;37:753-62.

7. Eshki M, Allanore L, Musette P, Milpied B, Grange A, Guillaume JC et al. Twelve-year analysis of severe cases of drug reaction with eosinophilia and systemic symptoms: a cause of unpredictable multiorgan failure. Arch Dermatol. 2009 Jan;145(1):67-72.

8. Kimmoun A, Dubois E, Perez P, Barbaud A, Levy B. Shock state: an unrecognized and underestimated presentation of drug reaction with eosinophilia and systemic symptoms. Shock. 2013 Nov;40(5):387-91.

9. Vincent JL, Moreno R, Takala J, Willatts S, De Mendonça A, Bruining H et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. Intensive Care Med. 1996 Jul;22(7):707-10.

10. Kardaun SH, Sekula P, Valeyrie-Allanore L, Liss Y, Chu CY, Creamer D et al. Drug reaction with eosinophilia and systemic symptoms (DRESS): an original multisystem adverse drug reaction. Results from the prospective RegiSCAR study. Br J Dermatol. 2013 Nov;169(5):1071-80.

11. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol. 2007 Mar 15;165(6):710-8.

12. Smeden MV, de Groot JAH, Moons KGM, Collins GS, Altman DG, Eijkemans MJC et al. No rationale for 1 variable per 10 events criterion for binary logistic regression analysis. BMC Med Res Methodol. 2016 Nov 24;16(1):163.

13. Taweesedt PT, Nordstrom CW, Stoeckel J, Dumic I. Pulmonary manifestations of drug reaction with eosinophilia and systemic symptoms (DRESS) syn drome: a systematic review. Biomed Res Int 2019; 2019:7863815.

14. Thongsri T, Chularojanamontri L, Pichler WJ. Cardiac involvement in DRESS syndrome. Asian Pac J Allergy Immunol 2017;35:3-10

15. Marcombes C, Ingen-Housz-Oro S, Dezoteux F, Staumont-Sallé D, Milpied B, Tetart F et al. Retrospective study on the association of human herpesvirus reactivation with severe DRESS: A description of blood and skin reactivations. J Eur Acad Dermatol Venereol. 2023 Dec;37(12):2550-2557.

16. Martinez-Cabriales SA, Shear NH, Gonzalez-Moreno EI. Liver involvement in the drug reaction, eosinophilia, and systemic symptoms syndrome. World J Clin Cases 2019;7:705-16.

17. Lin I-C, Yang H-C, Strong C, et al. Liver injury in patients with DRESS: a clinical study of 72 cases. J Am Acad Dermatol 2015;72:984-91.

18. Dagnon da Silva M, Domingues SM, Oluic S, et al. Renal manifestations of drug reaction with eosinophilia and sys temic symptoms (DRESS) syndrome: a systematic review of 71 cases. J Clin Med 2023;12:4576.

19. Mizukawa Y, Hirahara K, Kano Y, Shiohara T. Drug-induced hypersensitivi ty syndrome/drug reaction with eosinophilia and systemic symptoms severity score: a useful tool for assessing disease severity and predicting fatal cytomegalovirus dis ease. J Am Acad Dermatol 2019; 80(3): 670 678.e2.

20. Tohyama M, Hashimoto K, Yasukawa M, et al. Association of human herpesvirus 6 reactivation with the flaring and severity of drug-induced hypersensitivity syndrome. Br J Dermatol 2007; 157: 934 40

21. Ingen-Housz-Oro S, Guichard E, Milpied B, Bensaid B, Collet E, Barbaud A et al. Topical versus oral corticosteroids in moderate drug reaction with eosinophilia and systemic

symptoms: A multicenter randomized clinical trial. J Am Acad Dermatol. 2024 Sep;91(3):544-547.

22. Gschwend A, Helbling A, Feldmeyer L, et al. Treatment with IL5-/IL-5 receptor antagonists in drug reaction with eosinophilia and systemic symptoms (DRESS). Allergo J Int 2023; 32: 104-11.

23. Mesli F, Dumont M, Soria A et al, (2021) Benralizumab: A potential tailored treatment for life-threatening DRESS in the COVID-19 era. J Allergy Clin Immunol Pract. 2021 Sep;9(9):3529-3531.e1.

24. Sharma AN, Shwe S, Ravi V, et al. Characterizing DRESS syndrome recurrence: a systematic review. Arch Dermatol Res 2022;314:721-8.

DECLARATIONS

Ethics approval and consent to participate

This cohort study was approved by the Institutional Review Board of the AP-HP Data Warehouse (approval # CSE-22-22_ICUDRESS), and patients included in this study did not express their opposition to participating to the study. The AP-HP Data Warehouse was approved by the National Committee of Informatics and Liberties.

Consent for publication

Not applicable

Availability of data and materials

The dataset used and analyzed during the current study are available from the AP-HP Clinical Data Warehouse on reasonable request [AG].

Competing interests

N.D.P has served as an advisor or speaker for Moderna and AstraZeneca. V.D. has received consulting fees by Abbvie, Boehringer, Celgene, Janssen, Leo Pharma, Lilly, MSD, Novartis, Sanofi and support for attending meetings and/or travel for JANSSEN. Other authors have no conflict of interest to disclose.

Funding

There was no specific source of funding for this study

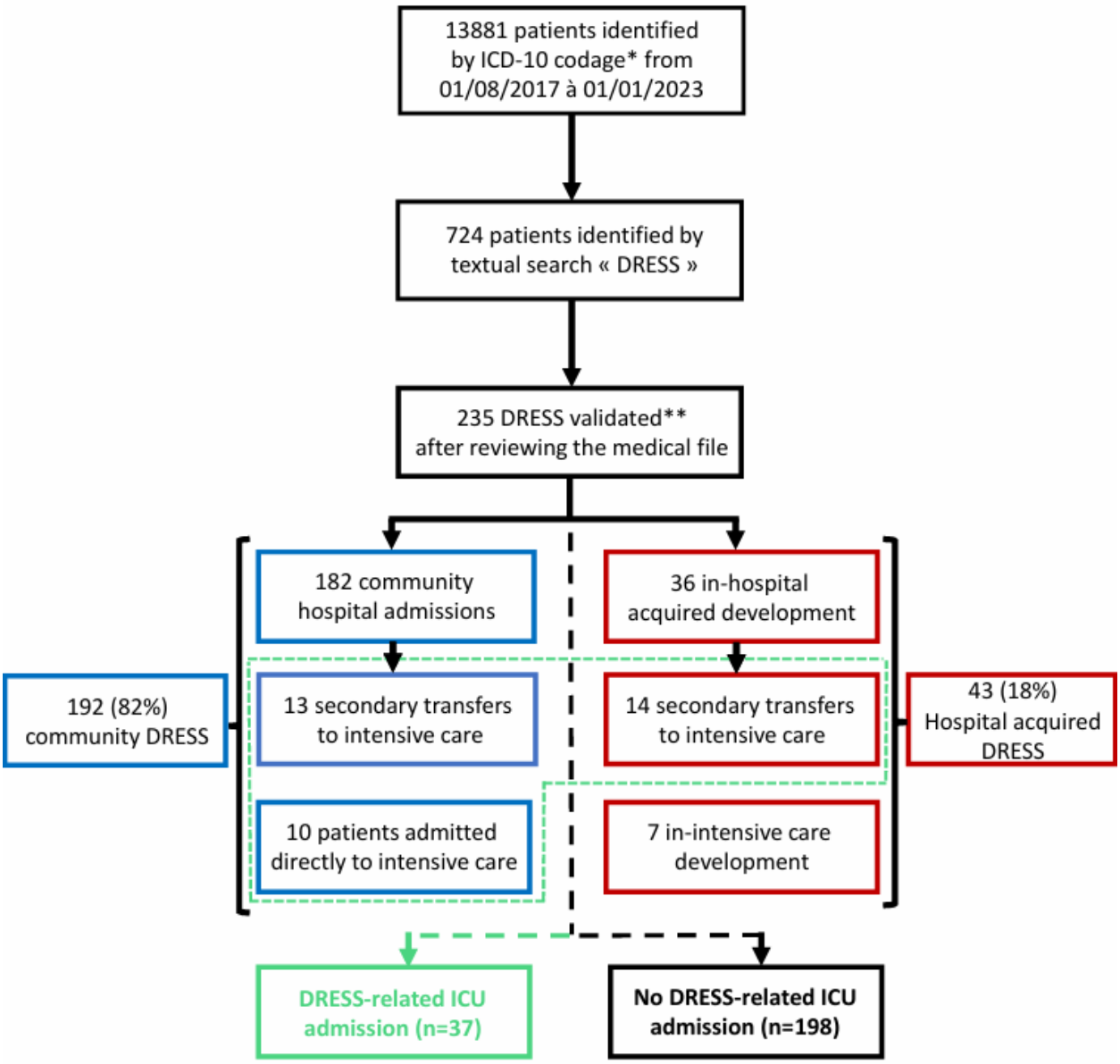
Authors' contributions

G.A., N.O., R.L., E.S., S.IHO. and N.D.P. designed the study; G.A. and R.L. performed statistical analyses; G.A., S.IHO. and N.D.P. wrote the first draft of the article; N.D.P. supervised the study; All authors validated the results of the study, reviewed, edited and approved the article. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. N.D.P. is the guarantor.

Acknowledgements

The authors would like to thank all nurses and physicians involved in patient care. The authors would also like to thank the « Collège des Enseignants de Médecine Intensive Réanimation d'Ile-de-France » (MIRIF) as well as all the members of the ICUDRESS study group of the clinical data warehouse of Greater Paris University Hospitals: Nicolas de Prost, Armand Mekontso Dessap, Saskia Ingen-Housz-Oro, Emilie Sbidian, Pierre Wolkenstein (Henri Mondor, Créteil), Elie Azoulay, Jean-David Bouaziz (Saint-Louis, Paris), Maxime Touron, Jean-Paul Mira, Nicolas Dupin (Cochin, Paris), Alexandre Demoule, Alain Combes (Pitié-Salpêtrière, Paris), Antoine Gaillet, Jean-François Timsit, Lila Bouadma, Vincent Descamps (Bichat, Paris), Bruno Mégarbane (Lariboisière, Paris), Antoine Vieillard-Baron, Tu-Anh Duong (Ambroise Paré, Boulogne), Stéphane Gaudry, Frédéric Caux (Avicenne, Bobigny), Hafid Ait-Oufella, Eric Maury (Saint-Antoine, Paris), Xavier Monnet (Bicêtre, Le Kremlin-Bicêtre).

Figure 1: Flow-chart of the ICUDRESS study



Abbreviations: DRESS: Drug Reaction with Eosinophilia and Systemic Symptoms; ICD-10: International Classification of Diseases, Tenth Revision

The green dashed line identifies DRESS-related ICU admissions (n=37) corresponding to patients who were thus included in the regression analysis identifying risk factors for ICU admission. Blue boxes represent community-acquired DRESS, while red boxes correspond to hospital-acquired DRESS.

*D72.1 (Eosinophilia), T78.4 (Allergy), T88.7 (Adverse drug reaction), or L27.0 (Generalized rash due to drugs), in any hospital stay during the study period, using the French national hospital database (Programme de Médicalisation des Systèmes d'Information),

**confirmed by expert review of the medical records (first expert (A.G.) then a second opinion if there is any doubt about the diagnosis (N.D.P.)) and using a RegiSCAR score ≥3

Table 1: Characteristics of hospitalized patients with DRESS (n=235)

	All patients (n=235)
Age (year) (n=235)	60 [54 ; 66]
Female sex (n=235)	139 (59%)
BMI ≥ 30 kg/m² (n=235)	26 (11%)
Diabetes (n=235)	41 (17%)
Solid tumor (n=235)	32 (14%)
Lymphoma/leukemia (n=235)	43 (18%)
Charlson score (n=235)	3 [2 ; 4]
DRESS timing (n=235):	
Culprit drug introduction – DRESS onset (day)	25 [18 ; 28]
DRESS diagnosis – culprit drug suspension time lag (day)	2 [1 ; 2]
Duration under treatment (day)	26 [20 ; 29]
Hospital-acquired development	43 (18%)
Culprit drug (n=235):	
Anticonvulsants	47 (20%)
Allopurinol	32 (14%)
Beta-lactams	29 (12%)
Sulfonamides	30 (13%)
Rifampicin	8 (3%)
Revlimid	14 (6%)
Other	64 (27%)
Main reason for hospital admission (n=235):	
Exanthema	165 (70%)
Related to DRESS	205 (87%)
Any ICU stay (n=235)*	44 (19%)
ICU admission main reason (n=37):**	
Exanthema	5 (13%)
Acute respiratory failure	13 (35%)
Acute renal failure	2 (5%)
Sepsis	15 (40%)
Other	9 (24%)
Related to DRESS	37 (100%)
SOFA at DRESS diagnosis (n=40)	4 [2 ; 5]
Cutaneous presentation (n=235):	
Exanthema	228 (97%)
Face oedema	140 (60%)
Blister	4 (2%)
Pustules	13 (6%)
Systemic involvement (n=235):	
Fever	184 (78%)
Lymphadenopathy	137 (58%)
Hepatic	161 (69%)
Renal	65 (28%)
Pulmonary	6 (3%)
Cardiac	11 (5%)
Hemodynamic failure	14 (6%)
Haemophagocytosis	8 (3%)
Multivisceral involvement (≥2 organs)	60 (26%)
Laboratory analyses:	
Eosinophilia (G/L, n=225)	1.7 [1.2 ; 2.0]
Lymphocytosis (n=235)	77 (33%)
CMV reactivation (n=201)	16 (7%)
EBV reactivation (n=200)	56 (24%)
HHV6 reactivation (n=188)	35 (15%)
RegiSCAR† scoring system for classifying DRESS cases (n=235):	6 [4 ; 7]
3 (possible)	21 (9%)
4-5 (probable)	74 (31%)
≥6 (certain)	140 (60%)
DRESS severity grading†† (n=235):	2 [1 ; 3]
Mild (1)	95 (40%)
Moderate (2)	83 (35%)
Severe (3)	57 (24%)

Results are N (%), means (±standard deviation) or medians (interquartile range). BMI: Body mass index; CMV: Cytomegalovirus; DRESS: Drug rash with eosinophilia and systemic symptoms; EBV: Epstein–Barr virus; HHV6: Human herpesvirus 6; ICU: Intensive care unit; SOFA: Sequential Organ Failure Assessment.* including direct ICU admission (n=10); ICU onset (n=7); secondary transfers to ICU from hospital (n=27). ** DRESS cases admitted in ICU (n=37), excluding cases with ICU onset (n=7). † Kardaun SH et al. Br J Dermatol. 2013 Nov;169(5):1071-80. †† Brüggen MC et al. JAMA Dermatol. 2024

Table 2: Management and outcomes of hospitalized patients with DRESS

	All patients (n=235)
Specific therapeutics (n=235):	
Topical corticosteroids only	148 (63%)
Systemic corticosteroids	77 (33%)
Intravenous immunoglobulins	1 (0.5%)
Anti-viral drugs	6 (3%)
Cyclosporine	0
Anti-IL5 monoclonal antibodies	0
Organ support for ICU patients (n=44):	
Invasive mechanical ventilation	9 (20%)
Vasopressor support	15 (34%)
Renal replacement therapy	12 (27%)
Extracorporeal membrane oxygenation	5 (11%)
Transplantation	1 (2%)
Duration of hospital stay (day, n=233)	12 [5 ; 16]
Duration of ICU stay (day, n=44)	6 [4 ; 14]
In-hospital mortality (n=235)	12 (5%)
In-ICU mortality (n=44)	7 (16%)
Relapse (n=235)	6 (3%)
Autoimmune sequelae (n=235)	0

Results are N(%), means (±standard deviation) or medians (interquartile range); ICU (intensive care unit)

Table 3: Uni/multivariable logistic regression assessing risk factors for ICU admission in hospitalized DRESS

	No DRESS-related ICU admission (n=198)	DRESS-related ICU admission (n=37)	OR	p	aOR	p-value
Age (year) (n=235)	60 [43 ; 72]	59 [36 ; 74]	0.99 [0.99 ; 1.00]	0.711	1.01 [0.98 ; 1.03]	0.577
Female gender (n=235)	123 (62%)	16 (43%)	0.56 [0.27 ; 1.17]	0.124		
BMI ≥ 30 kg/m² (n=235)	20 (10%)	6 (16%)	1.94 [0.66 ; 5.01]	0.192		
Diabetes (n=235)	37 (19%)	4 (11%)	0.59 [0.17 ; 1.61]	0.350		
Solid tumor (n=235)	28 (14%)	4 (11%)	0.82 [0.23 ; 2.29]	0.734		
Lymphoma/leukemia (n=235)	40 (20%)	3 (8%)	0.39 [0.09 ; 1.16]	0.134		
Charlson score (n=235)	3 [1 ; 6]	3 [1 ; 6]	0.99 [0.98 ; 1.01]	0.628	1.02 [0.84 ; 1.20]	0.856
DRESS timing (n=235):						
Introduction – DRESS delay (day)	25 [13 ; 35]	20 [13 ; 29]	0.99 [0.99 ; 1.00]	0.680		
DRESS - suspension delay (day)	2 [0 ; 7]	2 [0 ; 5]	0.99 [0.98 ; 1.00]	0.170		
Timing under treatment (day)	27 [11 ; 38]	22 [14 ; 33]	0.99 [0.99 ; 1.00]	0.312		
Hospital-acquired onset	29 (15%)	14 (38%)	**4.15 [1.87 ; 9.12]**	**<0.001**	**4.09 [1.66 ; 10.2]**	**0.002**
Culprit drug (n=235):						
Anticonvulsants	43 (22%)	4 (11%)	0.49 [0.13 ; 0.28]	0.202		
Allopurinol	29 (15%)	3 (8%)	0.57 [0.13 ; 1.74]	0.383		
Beta-lactams	21 (11%)	8 (22%)	2.64 [1.01 ; 6.40]	**0.037**		
Sulfonamides	29 (15%)	1 (3%)	0.18 [0.01 ; 0.89]	0.097		
Rifampicin	5 (3%)	3 (8%)	3.79 [0.75 ; 16.3]	0.078		
Revlimid	12 (6%)	2 (5%)	0.98 [0.15 ; 3.83]	0.984		
Other	52 (26%)	12 (32%)	1.56 [0.70 ; 3.33]	0.256		
Main reason for hospital admission (n=235):	159 (80%)	7 (19%)	0.06 [0.02 ; 0.15]	**<0.001**		
Exanthema Related to DRESS	171 (86%)	12 (32%)	0.10 [0.04 ; 0.21]	**<0.001**		
Cutaneous presentation (n=235):						
Exanthema	197 (99%)	31 (84%)	0.21 [0.04 ; 1.11]	0.079		
Face oedema	126 (64%)	14 (38%)	**0.42 [0.19 ; 0.87]**	**0.020**	0.43 [0.18 ; 1.01]	0.055
Blister	1 (1%)	3 (8%)	**19.3 [2.39 ; 398]**	**0.011**		
Pustulosis	13 (7%)	0	-	-		
Systemic involvement (n=235):						
Fever	156 (79%)	28 (76%)	1.35 [0.56 ; 3.78]	0.536		
Lymphadenopathy	116 (59%)	21 (57%)	1.18 [0.57 ; 1.18]	0.658		
Hepatic	139 (70%)	22 (59%)	0.82 [0.39 ; 1.80]	0.606		
Renal	49 (25%)	16 (43%)	**2.76 [1.30 ; 5.83]**	**0.008**		
Pulmonary	4 (2%)	2 (5%)	3.08 [0.42 ; 16.4]	0.205		
Cardiac	8 (4%)	3 (8%)	2.33 [0.49 ; 8.57]	0.229		
Hemodynamic failure	4 (2%)	10 (27%)	**20.5 [6.34 ; 79.6]**	**<0.001**		
Haemophagocytosis	7 (4%)	1 (3%)	0.87 [0.04 ; 5.10]	0.895		
Multivisceral involvement (≥2 organs)	44 (22%)	16 (43%)	**3.17 [1.49 ; 6.75]**	**0.003**		
Laboratory analyses:						
Eosinophilia (/mm3, n=225)	1700 [900 ; 3000]	2000 [1100 ; 3500]	1.01 [1.00 ; 1.02]	**0.041**	1.00 [0.99 ; 1.01]	0.190
Lymphocytosis (n=235)	63 (32%)	14 (38%)	1.53 [0.72 ; 3.21]	0.261		
CMV reactivation (n=201)	12 (6%)	4 (11%)	2.10 [0.56 ; 6.49]	0.224		
EBV reactivation (n=200)	44 (22%)	12 (32%)	1.95 [0.87 ; 4.19]	0.094		
HHV6 reactivation (n=188)	31 (16%)	4 (11%)	0.73 [0.21 ; 2.02]	0.581		
RegiSCAR† scoring system (n=235):	6 [5 ; 7]	6 [4 ; 7]	0.99 [0.96 ; 1.02]	0.726		
3 (possible DRESS)	19 (10%)	2 (5%)	Ref.	0.393		
4-5 (probable DRESS)	60 (30%)	14 (38%)	2.22 [0.55 ; 14.9]			
≥6 (certain DRESS)	122 (62%)	18 (49%)	1.40 [0.36 ; 9.26]			
DRESS severity grading (n=235):	2 [1 ; 2]	3 [2 ; 3]	**1.16 [1.10 ; 1.22]**	**<0.001**		
Minor (1)	91 (46%)	4 (11%)	**Ref.**	**<0.001**	**Ref.**	**<0.001**
Moderate (2)	73 (37%)	10 (27%)	**3.12 [1.00 ; 11.7]**		**2.77 [0.81; 11.1]**	
Severe (3)	37 (19%)	20 (54%)	**12.3 [4.32 ; 44.5]**		**13.2 [4.15; 52.3]**	

Results are N (%), means (±standard deviation) or medians (interquartile range). BMI: Body mass index; CMV: Cytomegalovirus; DRESS: Drug rash with eosinophilia and systemic symptoms; EBV: Epstein–Barr virus; HHV6: Human herpesvirus 6; ICU: Intensive care unit; SOFA: Sequential Organ Failure Assessment; Two-tailed p-values come from unadjusted comparisons using Chi square or Fisher's exact tests for categorical variables, and t-tests or Mann-Whitney tests for continuous variables, as appropriate. No adjustment for multiple comparisons was performed; aOR (CI 95%): adjusted Odds Ratio (95% confidence interval). p-values come from multivariable logistic regression models; **Bolded** p-values are significant at the $p < 0.05$ level. † Kardaun SH et al. Br J Dermatol. 2013 Nov;169(5):1071-80.

Management, outcomes, and predictors of intensive care unit admission in DRESS: A retrospective multicenter cohort study in Great Paris area hospitals

Supplementary appendix

Table of contents

Content	Page Number
Table S1: RegiSCAR scoring system for classifying DRESS cases	**80**
Table S2: DRESS severity score detail	**81**
Table S3: Uni/multivariable logistic regression assessing risk factors for ICU admission in hospitalized probable or definite DRESS cases (RegiSCAR score ≥ 4).	**82**
Table S4: Characteristics of hospitalized patients with DRESS, according to vital status at hospital discharge	**83**
Table S5: Management and outcomes of patients with DRESS according to vital status at hospital discharge	**84**
Figure S1: Culprit drug histogram	**85**
Figure S2: Visceral involvement histogram, according to the histological evidence	**86**

Table S1: RegiSCAR† scoring system for classifying DRESS cases

Score	-1	0	1	2
Fever ≥ 38.5°C	51 (22%)	184 (78%)		
Enlarged lymph nodes		98 (42%)	137 (58%)	
Atypical lymphocytes		158 (67%)	77 (33%)	
Eosinophilia - **0-699 cells or <10%** - **700-1,499 cells or 10-19.9%** - **≥1,500 cells or ≥20%**		14 (6%)	69 (29%)	142 (60%)
Skin rash extent >50%		7 (3%)	228 (97%)	
At least two of: edema, infiltration, purpura, scaling	5 (2%)		230 (98%)	
Biopsy suggesting DRESS	25 (11%)	210 (89%)		
Internal organ involved		48 (20%)	127 (54%)	60 (26%)
Resolution in >15 days	5 (2%)	230 (98%)		
Alternative diagnoses excluded (by ≥3 biological investigations)		12 (5%)	223 (95%)	
Total RegiSCAR score	≤2	3	4-5	≥6
Patients (number)	0	21 (9%)	74 (31%)	140 (60%)

Results are N (%), means (±standard deviation). DRESS: Drug rash with eosinophilia and systemic symptoms. † Kardaun et al. Br J Dermatol. 2013 Nov;169(5):1071-80.

Table S2: DRESS severity score detail††

Severity	Total classification	Blood	Kidney	Liver	Other organs*		Topical corticosteroids only	Systemic corticosteroids
Mild	95 (40%)	210 (89%)	165 (70%)	129 (55%)	0		75 (79%)	16 (17%)
Moderate	83 (35%)	25 (11%)	43 (18%)	68 (29%)	0		50 (60%)	27 (33%)
Severe	57 (24%)	0	11 (5%)	36 (15%)	19 (8%)		23 (40%)	34 (60%)

Results are N (%), means (±standard deviation). * pulmonary n=6, cardiac n=11, hemodynamic failure n=14, haemophagocytosis n=8. †† Brüggen et al. JAMA Dermatol. 2024 Jan 1;160(1):37-44.

Table S3: Uni/multivariable logistic regression assessing risk factors for ICU admission in hospitalized probable or definite DRESS cases (RegiSCAR score ≥ 4).

	No DRESS-related ICU admission (n=182)	DRESS-related ICU admission (n=32)	OR	p	aOR	p-value
Age (year) (n=214)	58 [43 ; 72]	59 [35 ; 74]	0.99 [0.99 ; 1.00]	0.691	1.01 [0.98 ; 1.04]	0.659
Female gender (n=214)	114 (63%)	15 (47%)	0.53 [0.24 ; 1.12]	0.096		
BMI ≥ 30 kg/m² (n=214)	18 (10%)	6 (19%)	2.10 [0.71 ; 5.55]	0.150		
Diabetes (n=214)	37 (20%)	4 (13%)	0.56 [0.16 ; 1.53]	0.305		
Solid tumor (n=214)	26 (14%)	4 (13%)	0.86 [0.24 ; 2.41]	0.789		
Lymphoma/leukemia (n=214)	37 (20%)	3 (9%)	0.40 [0.09 ; 1.22]	0.154		
Charlson score (n=214)	3 [1 ; 6]	3 [2 ; 6]	0.99 [0.98 ; 1.01]	0.824	1.02 [0.85 ; 1.22]	0.790
DRESS timing (n=214):						
Introduction – DRESS delay (day)	25 [13 ; 35]	20 [13 ; 31]	0.99 [0.99 ; 1.00]	0.675		
DRESS - suspension delay (day)	2 [0 ; 7]	2 [0 ; 5]	0.99 [0.98 ; 1.00]	0.230		
Timing under treatment (day)	26 [10 ; 38]	22 [14 ; 33]	0.99 [0.99 ; 1.00]	0.434		
Hospital-acquired onset	26 (14%)	13 (41%)	**4.10 [1.79 ; 9.30]**	**<0.001**	**4.02 [1.57 ; 10.4]**	**0.004**
Culprit drug (n=214):						
Anticonvulsants	42 (23%)	4 (13%)	0.48 [0.13 ; 1.30]	0.187		
Allopurinol	29 (16%)	3 (9%)	0.54 [0.12 ; 1.67]	0.344		
Beta-lactams	19 (10%)	7 (22%)	2.40 [1.11 ; 6.10]	**0.044**		
Sulfonamides	23 (13%)	1 (3%)	0.22 [0.01 ; 1.12]	0.149		
Rifampicin	4 (2%)	3 (9%)	3.42 [0.70 ; 15]	0.084		
Revlimid	11 (6%)	2 (6%)	1.04 [0.15 ; 4.11]	0.964		
Other	44 (24%)	11 (34%)	1.64 [0.71 ; 3.62]	0.227		
Main reason for hospital admission (n=214):						
Exanthema	147 (81%)	6 (19%)	0.05 [0.02 ; 0.15]	**<0.001**		
Related to DRESS	165 (91%)	10 (31%)	0.10 [0.04 ; 0.21]	**<0.001**		
Cutaneous presentation (n=214):						
Exanthema	178 (98%)	29 (91%)	0.22 [0.06 ; 1.15]	0.054		
Face oedema	123 (68%)	14 (44%)	**0.37 [0.17 ; 0.80]**	**0.011**	0.40 [0.16 ; 1.05]	0.061
Blister	1 (1%)	3 (9%)	**18.7 [2.31 ; 386]**	**0.012**		
Pustulosis	13 (7%)	0	-	-		
Systemic involvement (n=214):						
Fever	147 (81%)	26 (81%)	1.03 [0.42 ; 2.94]	0.949		
Lymphadenopathy	113 (62%)	21 (66%)	1.16 [0.54 ; 2.64]	0.703		
Hepatic	131 (72%)	22 (69%)	0.86 [0.39 ; 2.00]	0.709		
Renal	49 (27%)	15 (47%)	**2.39 [1.10 ; 5.17]**	**0.025**		
Pulmonary	2 (1%)	2 (6%)	5.55 [0.70 ; 52]	0.079		
Cardiac	8 (4%)	3 (9%)	2.25 [0.47 ; 8.30]	0.251		
Hemodynamic failure	4 (2%)	10 (31%)	**20 [6.21 ; 79.1]**	**<0.001**		
Haemophagocytosis	7 (4%)	1 (3%)	0.83 [0.04 ; 4.92]	0.867		
Multivisceral involvement (≥2 organs)	44 (24%)	16 (50%)	**3.14 [1.44 ; 6.83]**	**0.004**		
Laboratory analyses:						
Eosinophilia (/mm3, n=225)	1700 [900 ; 3000]	1950 [1000 ; 3800]	1.01 [1.00 ; 1.02]	**0.042**	1.00 [0.99 ; 1.01]	0.196
Lymphocytosis (n=214)	63 (35%)	14 (44%)	1.47 [0.68 ; 3.14]	0.323		
CMV reactivation (n=195)	11 (6%)	4 (13%)	2.22 [0.58 ; 7.02]	0.197		
EBV reactivation (n=193)	40 (22%)	11 (34%)	1.86 [0.80 ; 4.12]	0.133		
HHV6 reactivation (n=180)	30 (16%)	4 (13%)	0.72 [0.20 ; 2.02]	0.571		
RegiSCAR† scoring system (n=214):	6 [5 ; 7]	6 [4 ; 7]	0.63 [0.29 ; 1.38]	0.239		
4-5 (probable DRESS)	60 (33%)	14 (44%)				
≥6 (certain DRESS)	122 (67%)	18 (56%)				
DRESS severity grading†† (n=214):	2 [1 ; 2]	3 [2 ; 3]	**1.15 [1.08 ; 1.22]**	**<0.001**		
Minor (1)	78 (43%)	4 (13%)	**Ref.**	**<0.001**	**Ref.**	**<0.001**
Moderate (2)	68 (37%)	10 (31%)	**2.87 [0.91 ; 10.8]**		**2.22 [0.64; 8.91]**	
Severe (3)	36 (20%)	18 (56%)	**9.75 [3.36 ; 35.6]**		**9.32 [2.91; 36.9]**	

Results are N (%), means (±standard deviation) or medians (interquartile range). BMI: Body mass index; CMV: Cytomegalovirus; DRESS: Drug rash with eosinophilia and systemic symptoms; EBV: Epstein–Barr virus; HHV6: Human herpesvirus 6; ICU: Intensive care unit; SOFA: Sequential Organ Failure Assessment; Two-tailed p-values come from unadjusted comparisons using Chi square or Fisher's exact tests for categorical variables, and t-tests or Mann-Whitney tests for continuous variables, as appropriate. No adjustment for multiple comparisons was performed; aOR (CI 95%): adjusted Odds Ratio (95% confidence interval). p-values come from multivariable logistic regression models; **Bolded** p-values are significant at the p < 0.05 level. † Kardaun SH et al. Br J Dermatol. 2013 Nov;169(5):1071-80. †† Brüggen MC et al. JAMA Dermatol. 2024

Table S4: Characteristics of hospitalized patients with DRESS, according to vital status at hospital discharge

	All patients (n=235)	Hospital survivors (n=223)	Hospital non survivors (n=12)	p-value
Age (year) (n=235)	60 [54 ; 66]	58 [50 ; 63]	75 [68 ; 78]	**0.008**
Female gender (n=235)	139 (59%)	129 (58%)	10 (83%)	0.080
BMI ≥ 30 kg/m² (n=235)	26 (11%)	23 (10%)	3 (25%)	0.114
Diabetes (n=235)	41 (17%)	37 (17%)	4 (33%)	0.136
Solid tumor (n=235)	32 (14%)	31 (14%)	1 (8%)	0.584
Lymphoma/leukemia (n=235)	43 (18%)	43 (19%)	0	0.092
Charlson score (n=235)	3 [2 ; 4]	3 [2 ; 3]	6 [4 ; 7]	**0.010**
DRESS timing (n=235):				
Introduction – DRESS delay (day)	25 [18 ; 28]	24 [14 ; 29]	32 [20 ; 37]	0.201
DRESS - suspension delay (day)	2 [1 ; 2]	2 [1 ; 3]	0 [0 ; 0]	0.162
Timing under treatment (day)	26 [20 ; 29]	26 [18 ; 28]	26 [22 ; 30]	0.489
Hospital acquired development	43 (18%)	37 (17%)	6 (50%)	**0.003**
Culprit drug (n=235):				
Anticonvulsants	47 (20%)	47 (21%)	0	**0.016**
Allopurinol	32 (14%)	27 (12%)	5 (42%)	
Beta-lactams	29 (12%)	25 (11%)	4 (33%)	
Sulfonamides	30 (13%)	30 (13%)	0	
Revlimid	14 (6%)	14 (6%)	0	
Rifampicin	8 (3%)	7 (3%)	1 (8%)	
Other	64 (27%)	62 (28%)	2 (17%)	
Hospital admission main reason (n=235):				
Exanthema	165 (70%)	161 (72%)	4 (33%)	**<0.001**
Related to DRESS	205 (87%)	195 (87%)	10 83%)	0.978
Any ICU stay (n=235)*	44 (19%)	36 (16%)	8 (67%)	**<0.001**
ICU admission main reason (n=37) **:				
Exanthema	5 (13%)	3 (10%)	2 (28%)	0.654
Non infectious respiratory failure	13 (35%)	11 (37%)	2 (28%)	
Acute renal failure	2 (5%)	2 (7%)	0	
Sepsis	15 (40%)	12 (40%)	3 (43%)	
Other	9 (24%)	8 (27%)	1 (14%)	
Related to DRESS	37 (100%)	30 (100%)	7 (100%)	-
Direct admission	10 (27%)	10 (33%)	0	**<0.001**
SOFA at DRESS diagnosis (n=40)	4 [2 ; 5]	4 [2 ; 4]	6 [4 ; 7]	0.130
Cutaneous presentation (n=235):				
Exanthema	228 (97%)	216 (97%)	12 (100%)	0.533
Face oedema	140 (60%)	134 (60%)	6 (50%)	0.488
Blister	4 (2%)	4 (2%)	0	0.640
Pustulosis	13 (6%)	13 (6%)	0	0.390
Systemic involvement (n=235):				
Fever	184 (78%)	174 (78%)	10 (83%)	0.664
Lymphadenopathy	137 (58%)	133 (60%)	4 (33%)	0.072
Hepatic	161 (69%)	156 (70%)	5 (42%)	**0.039**
Renal	65 (28%)	60 (27%)	5 (42%)	0.265
Pulmonary	6 (3%)	5 (2%)	1 (8%)	0.192
Cardiac	11 (5%)	10 (4%)	1 (8%)	0.539
Hemodynamic failure	14 (6%)	13 (6%)	1 (8%)	0.721
Haemophagocytosis	8 (3%)	7 (3%)	1 (8%)	0.289
Multivisceral involvement (≥2 organs)	60 (26%)	57 (26%)	3 (25%)	0.965
Laboratory analyses:				
Eosinophilia (/mm3, n=225)	1700 [1200 ; 2000]	1700 [1100 ; 1900]	1900 [1400 ; 2300]	0.535
Lymphocytosis (n=235)	77 (33%)	73 (33%)	4 (33%)	0.966
CMV reactivation (n=201)	16 (7%)	14 (6%)	2 (17%)	0.164
EBV reactivation (n=200)	56 (24%)	54 (24%)	2 (17%)	0.550
HHV6 reactivation (n=188)	35 (15%)	34 (15%)	1 (8%)	0.512
RegiSCAR† scoring system for classifying DRESS (n=235):	6 [4 ; 7]	6 [4 ; 7]	6 [5 ; 6]	0.060
3 (possible)	21 (9%)	19 (9%)	2 (17%)	0.590
4-5 (probable)	74 (31%)	70 (31%)	4 (33%)	
≥6 (certain)	140 (60%)	134 (60%)	6 (50%)	
DRESS severity grading (n=235):	2 [1 ; 3]	2 [1 ; 3]	2 [1 ; 3]	0.976
Mild (1)	95 (40%)	90 (40%)	5 (42%)	0.989
Moderate (2)	83 (35%)	79 (35%)	4 (33%)	
Severe (3)	57 (24%)	54 (24%)	3 (25%)	

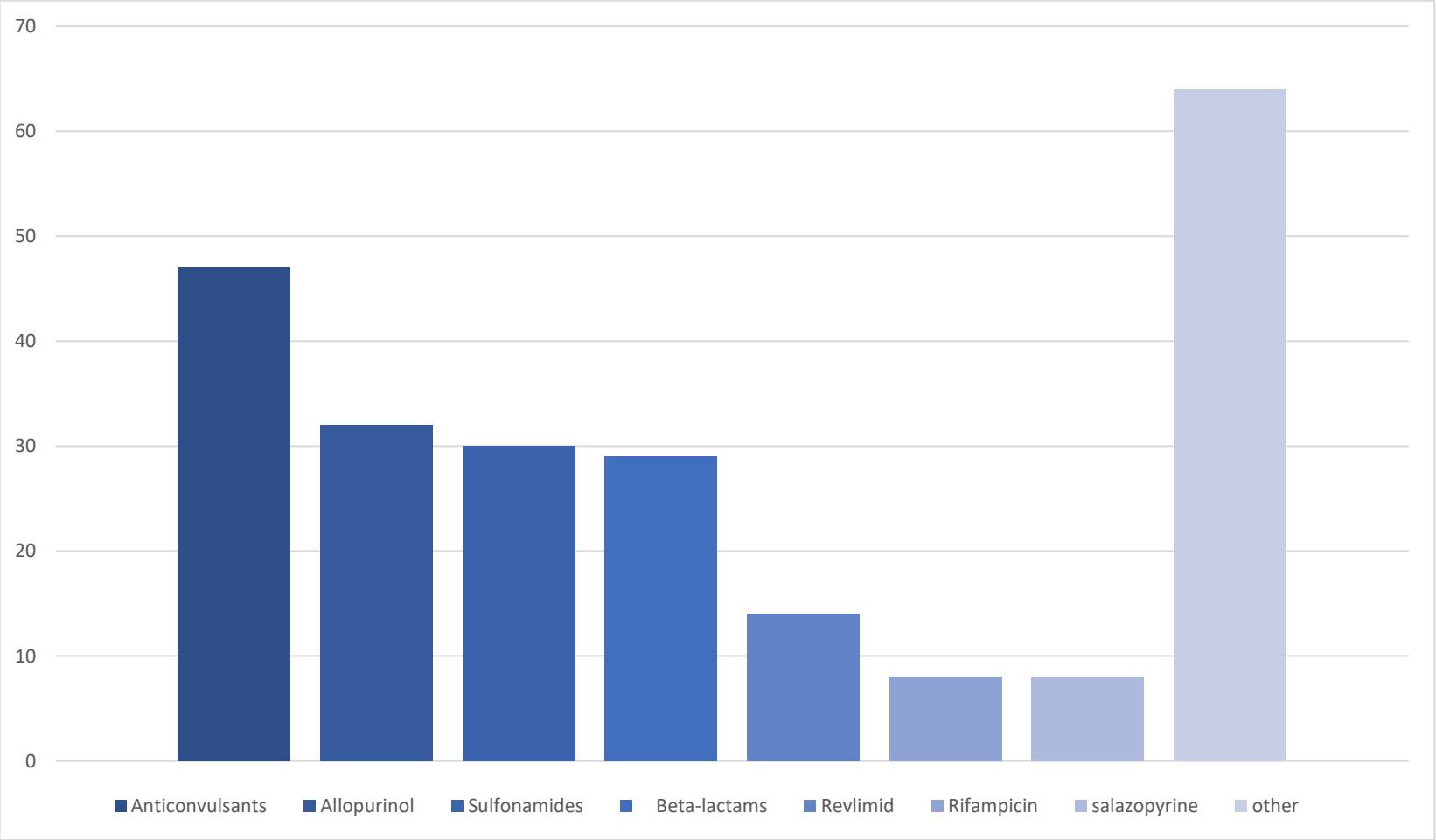
Results are N (%), means (±standard deviation) or medians (interquartile range). BMI: Body mass index; CMV: Cytomegalovirus; DRESS: Drug rash with eosinophilia and systemic symptoms; EBV: Epstein–Barr virus; HHV6: Human herpesvirus 6; ICU: Intensive care unit; SOFA: Sequential Organ Failure Assessment; Two-tailed p-values come from unadjusted comparisons using Chi square or Fisher's exact tests for categorical variables, and t-tests or Mann-Whitney tests for continuous variables, as appropriate. No adjustment for multiple comparisons was performed; **Bolded** p-values are significant at the p<0.05 level. *: including n=10 direct ICU admission, n=7 ICU development, n=27 secondary transfers to ICU from hospital; **: n=37 DRESS admitted in ICU excluding n=7 ICU development; † Kardaun et al. Br J Dermatol. 2013 Nov;169(5):1071-80.

Table S5: Management and outcomes of patients with DRESS according to vital status at hospital discharge

	All patients (n=235)	Hospital survivors (n=223)	Hospital non survivors (n=12)	p
Specific therapeutics (n=235):				
Topical corticosteroids only	148 (63%)	141 (63%)	7 (58%)	0.828
Systemic corticosteroids	77 (33%)	72 (32%)	5 (42%)	
Intravenous immunoglobulins	1 (0.5%)	0	1 (8%)	-
Anti-viral drugs	6 (3%)	5 (2%)	1 (8%)	0.267
Cyclosporine	0	0	0	-
Anti-IL5 drugs	0	0	0	-
Organ support for ICU patients (n=44):				
Invasive mechanical ventilation	9 (20%)	8 (22%)	1 (14%)	0.404
Vasopressor support	15 (34%)	10 (27%)	5 (71%)	**<0.001**
Renal replacement therapy	12 (27%)	10 (27%)	2 (28%)	0.062
Extracorporeal membrane oxygenation	5 (11%)	3 (8%)	2 (28%)	**<0.001**
Transplantation	1 (2%)	1 (0.4%)	0	-
Duration of hospital stay (day, n=233)	12 [5 ; 16]	11 [4 ; 14]	73 [14 ; 82]	**0.013**
Duration of ICU stay (day, n=44)	6 [4 ; 14]	6 [3 ; 12]	9 [7 ; 22]	0.927
In-ICU mortality (n=235)	7 (3%)	0	7 (58%)	-
Relapse (n=235)	6 (3%)	6 (3%)	0	-
Autoimmune sequelae (n=235)	0	0	0	-

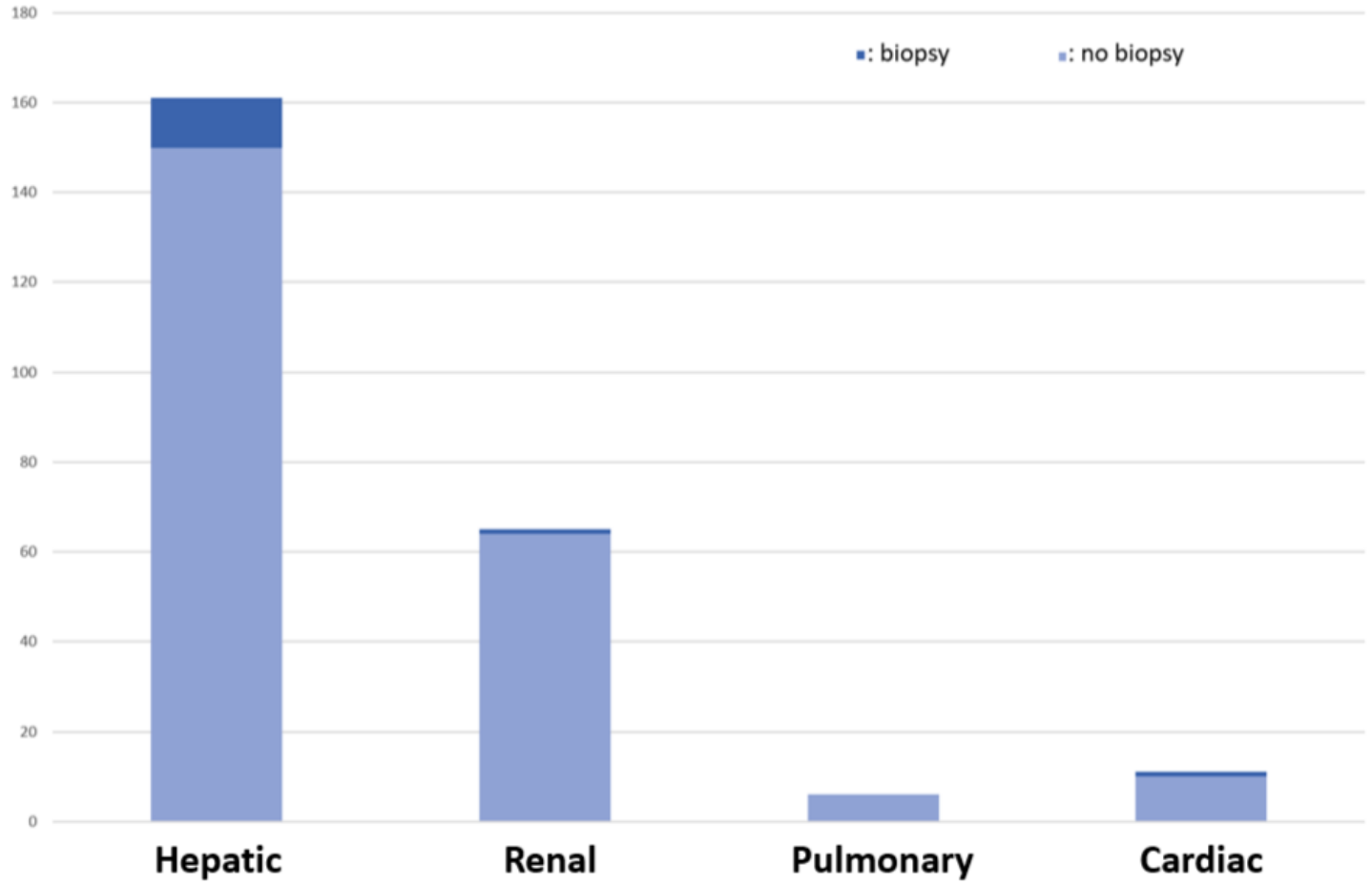
Results are N(%), means (±standard deviation) or medians (interquartile range); ICU (intensive care unit); Two-tailed p-values come from unadjusted comparisons using Chi square or Fisher's exact tests for categorical variables, and t-tests or Mann-Whitney tests for continuous variables, as appropriate. No adjustment for multiple comparisons was performed; **Bolded** p-values are significant at the p<0.05 level.

Figure S1 : Culprit drug histogram



Other : Bortezomib (n=3); Ciprofloxacin (n=2); Dalacine (n=3); Disulone (n=5); Enoxaparin (n=2); Eslicarbazepine (n=2); Everolimus (n=2); Idelalisib (n=2); Iomeron (n=3); Lansoprazole (n=4); Non-steroidal anti-inflammatory (n=5); Pembrolizumab (n=2); Terbinafine (n=3); Vancomycine (n=5); Zonisamide (n=2)

Figure S2 : Visceral involvement histogram, according to the histological evidence



V. Discussion

On montre à travers ces deux situations épidémiologiques quels peuvent être les apports de l'EDS AP-HP pour explorer des conditions cliniques contrastées chez des patients pris en charge dans les services de soins critiques (i.e., réanimations et unités de soins intensifs polyvalentes) des hôpitaux de l'AP-HP. Dans la première étude, le codage CIM-10, couplé à une analyse par recherche textuelle, a permis d'extraire automatiquement des informations essentielles sur les infections sévères à VRS (co-infections, traitements). Validé sur 100 dossiers, l'algorithme a facilité la constitution rapide d'une large cohorte, tout en permettant l'inclusion de variables fines dans la base de données. Dans la seconde étude, dédiée au DRESS, le PMSI seul s'avérait insuffisant. Un pré-screening basé sur les codes PMSI les plus pertinents a été complété par une recherche textuelle du terme « dress » permettant de confirmer le diagnostic et d'affiner efficacement le nombre de dossiers sélectionnés. La complexité des données (confirmation du diagnostic de DRESS, médicament imputable, délais, atteintes viscérales) ne permettait pas de recourir plus largement à la recherche textuelle et une relecture manuelle des 724 dossiers précédemment identifiés s'est avérée indispensable pour valider les 235 cas (RegiSCAR ≥3). Ce projet EDS a ainsi permis de constituer la plus vaste cohorte de DRESS à ce jour, offrant la puissance nécessaire pour identifier avec rigueur les facteurs de risque de sévérité.

A. Revue de la littérature

Ces deux travaux ne reflètent cependant pas la diversité des recherches réalisées grâce à l'EDS AP-HP. Une revue de la littérature menée via le site de l'EDS AP-HP et PUBMED - [Entrepôt de Données de Santé] OR [AP-HP Health Data Warehouse] OR [AP-HP Data Warehouse] OR [Systeme National des Donnees de Sante] OR [Clinical Data Warehouse (CDW) of Greater Paris]

– a permis d'identifier 258 projets et 108 publications, dont les caractéristiques sont résumées dans les tables 2 et 3.

Table 2 : caractéristiques des projets et publications issus de l'EDS AP-HP.

Variable	N (%) / médiane [IQR]
Statut projets (n=258) :	
- Publié	55 (21%)
- En cours	165 (64%)
- Terminé, non publié	38 (15%)
Publications (n=108) :	
- Délais depuis début du projet publié, années	3 [2 ; 3.5]
- Nombre par projet	1 [1 ;1]
Principaux domaines de publication (n=108) :	
- Méthodologie	24 (23%)
- Infectiologie	19 (18%)
- Radiologie	10 (9%)
- Oncologie	8 (7%)
- Gériatrie	8 (7%)
- **Soins critiques**	**8 (7%)**
Objectifs des études (n=108) :	
- Pronostique	55 (52%)
- Diagnostique	12 (11%)
- Méthodologie	24 (23%)
Nombre de patients/publication	2030 [255 ; 8892]
Méthode d'exploitation des données (n=83*) :	
- PMSI	76 (91%)
- Recueil manuelle	23 (28%)
- Recherche textuelle automatisée	22 (26%)

* : 108 – (1 essai contrôlé randomisé + 24 articles de méthodologie)

Table 3 : détail des domaines et sujets des publications issues de l'EDS AP-HP.

Domaine	Sujet	Référence
Méthodologie (n=24)	Qualité (n=14) Sécurité (n= 9) Ethique (n=6) Recherche textuelle (n=15) Accessibilité (n=10)	[18-21, 26-45]
Infectiologie (n=19)	COVID-19 (n=18) Bactériémie (n=1)	[46-64]
Radiologie (n=10)	IRM cérébrale (n=5) IRM prostatique (n=1) IRM cardiaque (n=1) Ostéoporose (n=1) Cancer pulmonaire (n=1) Embolie pulmonaire (n=1)	[65-74]
Oncologie (n=8)	COVID-19 (n=3) Iatrogénie (n=2) Cancer du sein (n=1) Cancer du pancréas (n=1) Urgence diagnostique (n=1)	{75-82]
Gériatrie (n=8)	COVID-19 (n=2) Urgences (n=2) Chute (n=1) Parcours de soins (n=1) Fragilité (n=1) Iatrogénie (n=1)	[83-90]
Soins critiques (n=8)	COVID-19 (n=3) Infection VRS (n=2) Grippe (n=1) Fasciites nécrosantes (n=1) Anémie hémolytique (n=1)	[25, 72-78]
Rhumatologie (n=5)	COVID-19 (n=2) Lupus (n=1) Biothérapie (n=1) Iatrogénie (n=1)	[98-102]
Pharmacologie (n=3)	Prescription (n=3)	[103-105]
Pédiatrie (n=3)	COVID-19 (n=2) Maltraitance (n=1)	[106-108]
Drépanocytose (n=3)	COVID-19 (n=1) Hyponatrémie (n=1) Psychiatrie (n=1)	[109-111]
Vasculaire (n=2)	COVID-19 (n=1) Thrombose (n=1)	[112,113]
Psychiatrie (n=2)	COVID-19 (n=1) Parcours de soins (n=1)	[114,115]

Table 3 (suite) : détail des domaines et sujets des publications issues de l'EDS AP-HP.

Domaine	Sujet	Référence
Pneumologie (n=2)	Asthme (n=1) Cancer du poumon (n=1)	[116,117]
Gastroentérologie (n=2)	Œsophagite (n=1) Hématémèse (n=1)	[118,119]
Diabétologie (n=2)	COVID-19 (n=1) Parcours de soins (n=1)	[120,121]
Dermatologie (n=2)	Biothérapie (n=1) Fasciite éosinophilique (n=1)	[122,123]
Neurologie (n=1)	COVID-19 (n=1)	[124]
Néphrologie (n=1)	Amylose AL (n=1)	[125]
Hématologie (n=1)	Lymphome (n=1)	[126]
Chirurgie (n=1)	COVID-19 (n=1)	[127]
Cardiologie (n=1)	COVID-19 (n=1)	[128]

On dénombre 75 publications directement liées à 55 projets, soit 21% des projets, avec un délai médian entre l'initiation du projet et la publication de 3 ans (un seul des 67 projets initiés depuis 2023 a été directement publié), et en médiane une publication par projet (atteignant un maximum de 7 publications pour un même projet) [24, 27-32, 46-63, 65, 67-71, 76-80, 83, 84, 86-88, 91-120, 122, 123, 125, 126]. Les 33 autres publications associées à l'EDS AP-HP [18-21, 26, 33-45, 64, 66, 72-75, 81, 82, 85, 89, 90, 121, 124, 127, 128] correspondent pour la moitié d'entre elles (54%, n=18/33) à des travaux de méthodologie.

L'infectiologie est une discipline majoritairement abordée, probablement du fait de la pandémie de COVID-19, associée à un pic de publication entre 2021 et 2023 (avec respectivement 85% (n=17/20), 33% (n=9/27) et 44% (n=7/16) des publications associées au sujet en 2021, 2022 et 2023, Figure 2). La méthodologie est également une des disciplines les plus étudiées : on retrouve essentiellement des études d'optimisation de la recherche textuelle (62%, n=15/24) et de la qualité (58%, n=14/24). Viennent ensuite parmi les 21 thématiques couvertes, la radiologie, l'oncologie, la gériatrie et les soins critiques (Figure 3).

La majeure partie des études de radiologie exploitent des données d'IRM (70%, n=7/10), avec un recours à des modèles automatisés d'IA dans 50% des cas. Les images proviennent des PACS (*Picture Archiving and Communication Systems*) des hôpitaux de l'AP-HP et sont accessibles via un processus de fédération et d'archivage des images brutes (DICOM), dans un cadre réglementaire et technique stricte. Les études d'oncologie traitent de sujets variés (iatrogénie, cancer isolé, COVID-19) et ont la particularité de provenir en majorité (62%) de la même équipe, familière de la recherche textuelle. Les études de gériatrie s'intéressent notablement au parcours patient, impliquant des cohortes de grande taille (médiane 3645 [670 ; 7232]). Enfin, on compte également huit publications ciblant la population de soins critiques. Hormis notre étude sur les infections à VRS, on compte trois études sur les COVID-19 sévères, une autre sur les patients VRS hospitalisés, une sur les infections grippales, une sur les fasciites nécrosantes et une sur les anémies hémolytiques auto-immunes. La génétique n'est, elle, directement abordée que dans une seule étude de cohorte de cancer du poumon BRCA muté, mais via une recherche textuelle au sein du dossier patient, et donc sans accès réel aux données génétiques. En effet, les données génétiques brutes (séquençage, génotypage) ne sont pas encore intégrées à l'EDS AP-HP. Le projet ACCES AP-HP est en cours de développement et prévoit l'intégration des données de génétique dans l'EDS, depuis les plateformes de séquençage de l'AP-HP (SeqOIA / MOABI) et la base de données Maladies Rares (BaMaRa). L'accès à ce type de données sera strictement encadré, uniquement après avis favorable du Comité Scientifique et Éthique (CSE) et dans des environnements sécurisés et pseudonymisés.

Enfin, hormis une étude cout-efficacité [84], et les études de radiologie principalement diagnostiques, la majorité des études cliniques avaient pour but d'identifier des facteurs pronostiques.

Figure 2 : Histogramme vertical présentant la distribution des publications par année

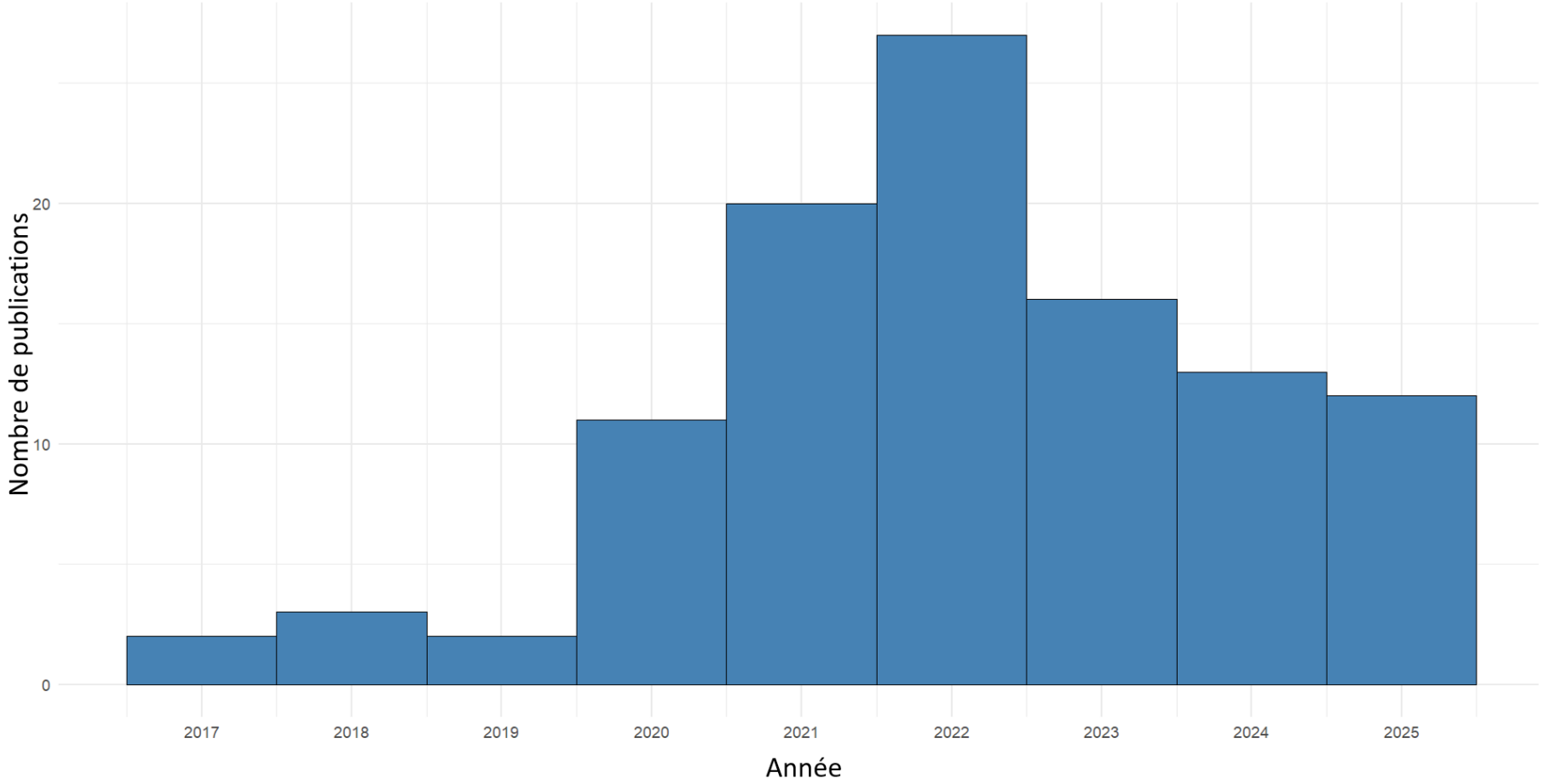
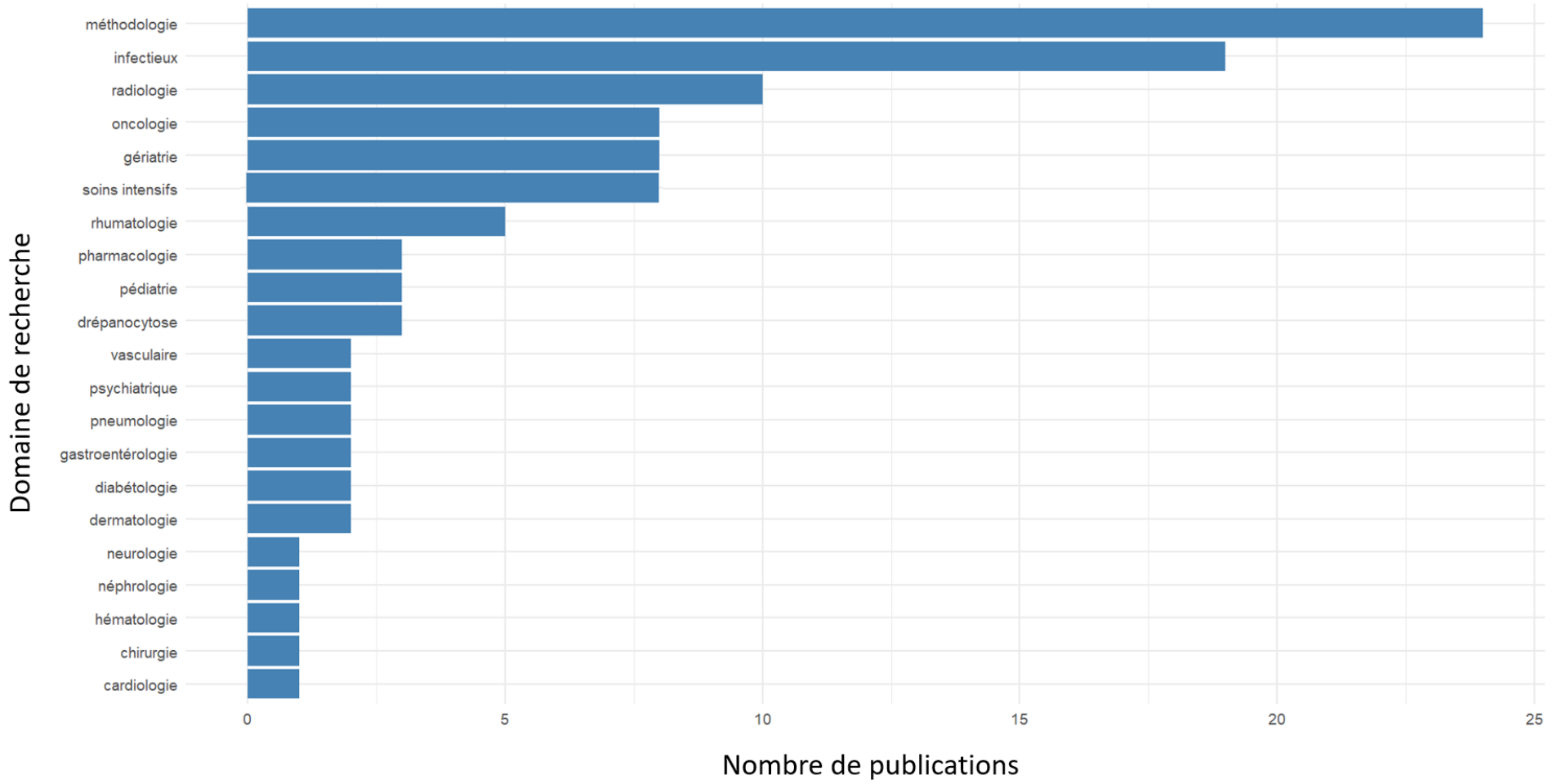


Figure 3 : Histogramme horizontal présentant la distribution des publications par domaine de recherche



La figure 4 illustre des corrélations entre les thèmes des études, le mode d'exploitation et le nombre de patients inclus. Etant donné le caractère à part des études de méthodologie, celles-ci n'ont pas été incluses dans la matrice de corrélation. On remarque qu'il y a de plus en plus d'études utilisant la recherche textuelle automatisée, notamment en oncologie, et le recueil manuel des notes, comme en soins critiques. Ces deux dernières méthodes d'exploitation (automatisée et manuelle) sont d'ailleurs inversement corrélées à la taille des cohortes. La grande majorité (78%, n=18/23) des études ayant eu manuellement recours aux notes médicales avaient un effectif < 500 patients. La figure 5 illustre les différentes stratégies d'exploitation des cohortes issues de l'EDS AP-HP. On remarque le recours quasi systématique au codage PMSI, associé en proportion similaire aux deux autres stratégies (recherche textuelle automatisée et recueil manuel). Dans les rares cas (n=4) où les trois méthodes étaient associées [24, 98, 118, 126], la recherche textuelle automatisée était employée essentiellement pour affiner la sélection d'une cohorte présélectionnée sur le codage PMSI (VRS, myosite sous inhibiteur de check point, œsophagite de Barrett).

Figure 4 : Matrice de corrélation des études de l'EDS AP-HP

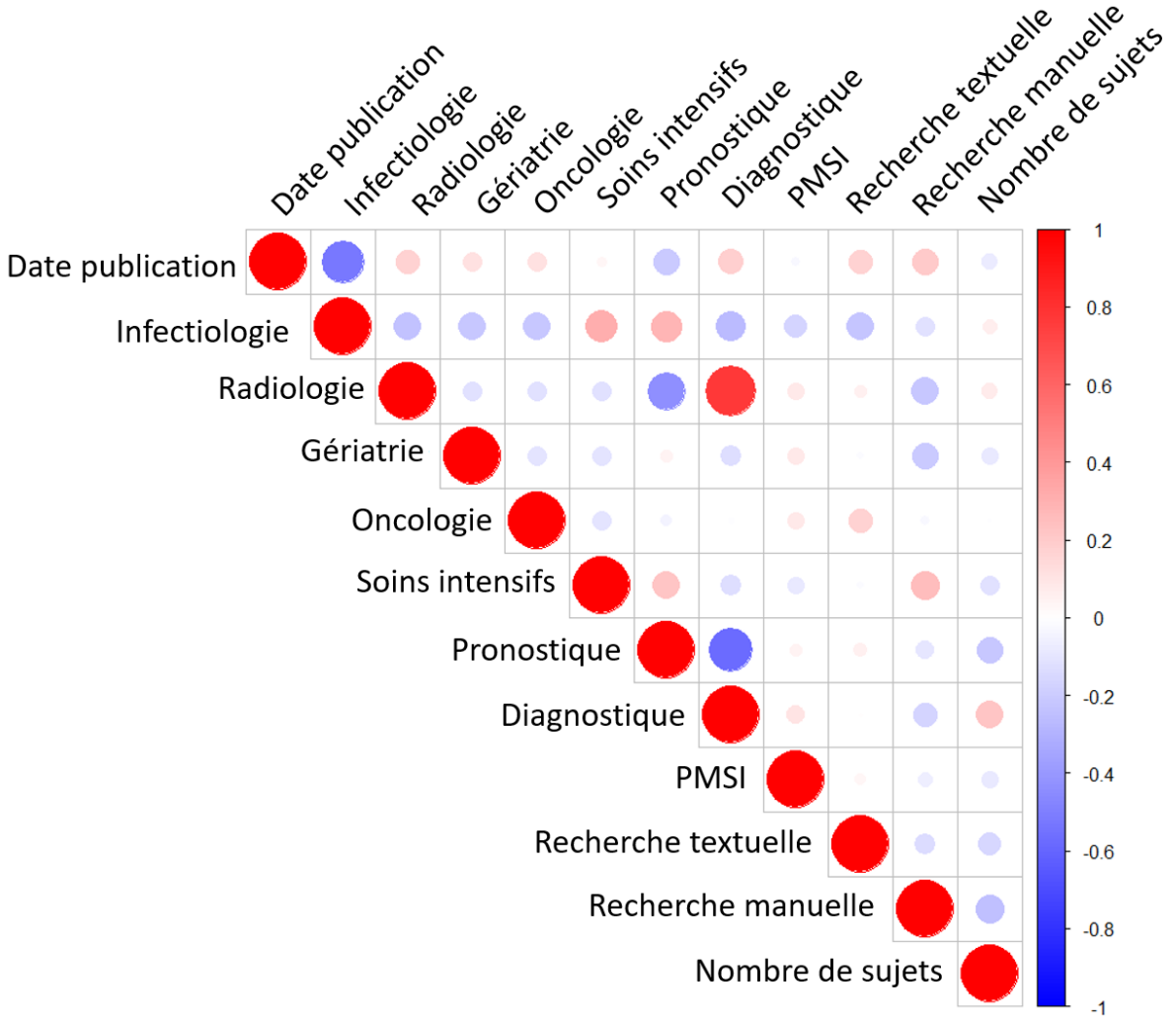


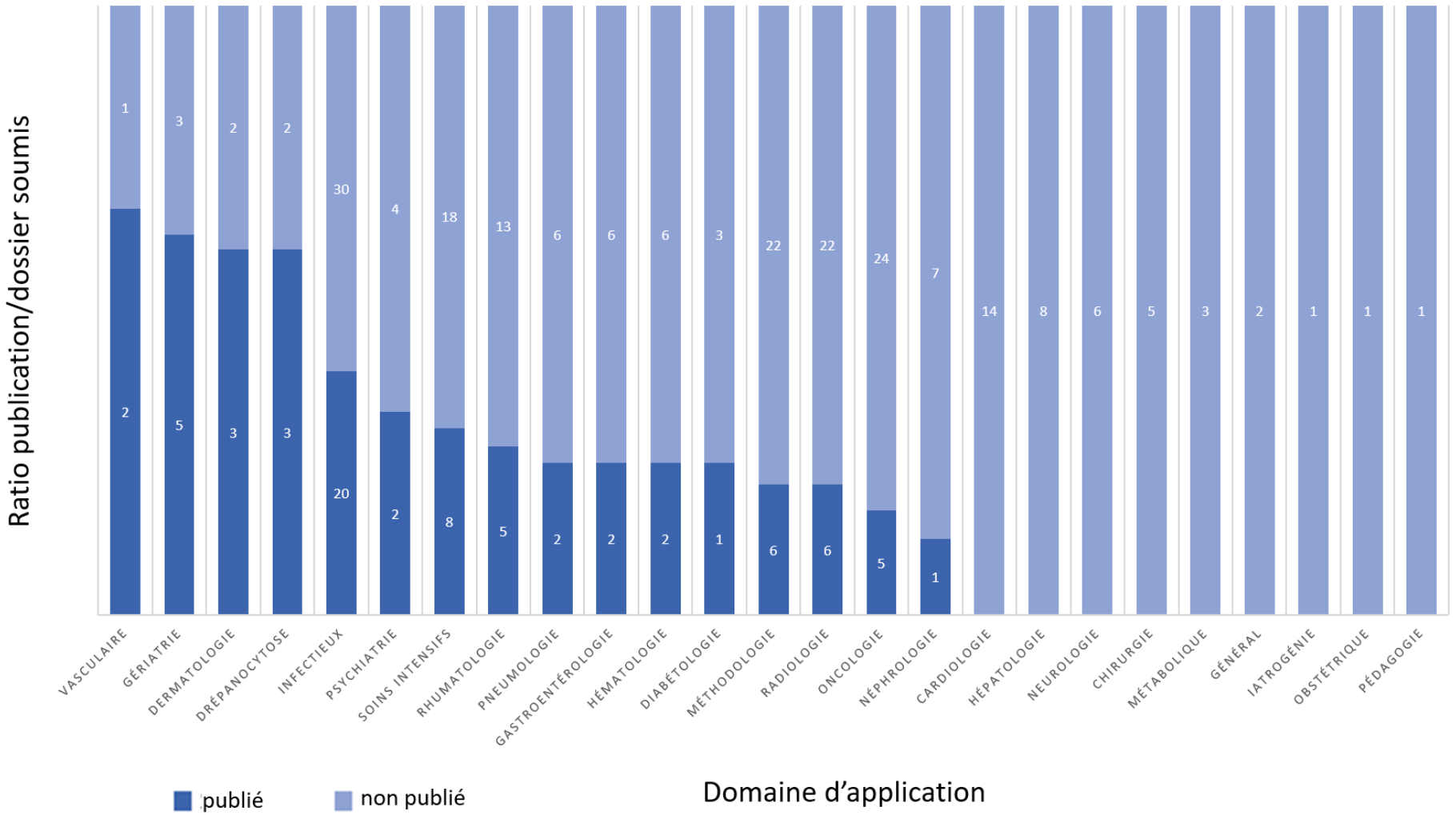
Figure 5 : Diagramme de Venn illustrant les différentes stratégies d'exploitation de l'EDS AP-HP dans 83 études réalisées entre 2017 et 2025.

CRH : compte rendu d'hospitalisation ; NLP : Natural langage processing ; PMSI : Programme de Médicalisation des Systèmes d'Information.



Etant donné le taux faible de publication des projets initiés (21%, n=55/258 projets), il est intéressant d'analyser les facteurs influençant cette probabilité. Comme abordé plus haut, la période a évidemment un impact. En effet, avec un délai médian de publication de 3 ans, les projets datant d'après 2022 ont moins de chance d'avoir abouti. La probabilité de publication des dossiers déposés avant 2023 est d'ailleurs logiquement plus élevée, de l'ordre de 28% (54/191). L'autre paramètre important est la discipline d'application du projet (25 domaines distincts). On observe effectivement visuellement une grande hétérogénéité en fonction de chacune des disciplines (Figure 6), tout en sachant que l'attribution par discipline peut être critiquée (notamment en cas de projet transversal). Cette influence suspectée de la discipline du projet n'est toutefois pas confirmée lorsqu'on effectue un ajustement sur l'année de soumission (via un modèle nul avec seulement un intercept aléatoire imbriqué sur le domaine (Median Odds Ratio (MOR) à 0) et l'année (MOR = 6, pour tenir compte de la dépendance intra-groupe)). Enfin, on constate que les projets mentionnant l'utilisation d'intelligence artificielle dans leur méthodologie n'aboutissaient quasiment jamais (3%, n=1/37).

Figure 6 : Histogramme vertical présentant la distribution des publications issues de projets EDS AP-HP par domaine d'application



Cette revue de la littérature révèle la diversité du champ d'action de l'EDS (avec donc une réserve sur l'applicabilité en génétique) mais appelle à plusieurs questions compte tenu de sa nature numérique, interconnectée et multimodale : accessibilité et sécurité des données., exploitation des données non structurées, contrôle qualité.

B. Accessibilité des données

Comme abordé dans l'étude qualitative de Priou et al. publiée en 2024 [40], l'accessibilité des données cliniques constitue l'un des piliers fondamentaux du développement de la recherche biomédicale contemporaine. Elle désigne la capacité à obtenir, exploiter et partager des données médicales issues des systèmes d'information hospitaliers, tout en respectant les principes éthiques, réglementaires et de sécurité qui encadrent leur utilisation.

Les entrepôts de données cliniques (EDC) reposent sur des architectures robustes combinant serveurs hospitaliers, solutions de stockage en cloud sécurisé et systèmes de requête performants. Les *data marts* — sous-ensembles thématiques des EDC largement développés à l'occasion de la pandémie à SARS-Cov-2 [19] avec la création de >200 *data marts* secondaires — facilitent l'extraction ciblée d'informations pour des besoins spécifiques (études épidémiologiques, essais cliniques, audits de qualité). Ces infrastructures doivent ainsi répondre à des exigences de scalabilité, de résilience et de sécurité, garantissant la disponibilité des données tout en limitant les risques de brèche de confidentialité.

La capacité des systèmes à communiquer entre eux constitue un facteur critique d'accessibilité. L'adoption de standards internationaux tels que l'OMOP ou les protocoles FHIR (*Fast Healthcare Interoperability Resources*) permet d'harmoniser les formats et les terminologies, assurant ainsi la comparabilité et l'intégration des données issues de multiples

établissements ou registres. Cette interopérabilité est un prérequis à la recherche multicentrique et à l'analyse à grande échelle.

L'accessibilité ne dépend pas uniquement des systèmes techniques, mais également de la capacité des utilisateurs à exploiter efficacement les outils. Le développement d'interfaces intuitives et de portails de requête ergonomiques permet aux cliniciens et chercheurs de formuler des questions complexes sans expertise informatique avancée. Ces outils réduisent la dépendance vis-à-vis des ingénieurs de données et accélèrent le processus d'analyse. Ainsi, le système Cohort360, déployé au sein de l'AP-HP depuis 2019, s'appuie directement sur l'infrastructure logicielle i2b2 et son extension i2b2-tranSMART. Il constitue la couche applicative et ergonomique de l'EDS AP-HP, permettant aux cliniciens et chercheurs de formuler des requêtes, identifier des cohortes de patients, et explorer des données cliniques massives issues des dossiers médicaux électroniques (DME).

Toute utilisation secondaire de données de santé doit s'inscrire dans un cadre strict de conformité réglementaire (RGPD, lois nationales sur la protection des données) et être validée par un comité d'éthique compétent. La gouvernance des données implique des processus clairs de gestion des accès, de traçabilité des requêtes et de contrôle des usages. Cette gouvernance garantit non seulement la protection de la vie privée des patients, mais aussi la confiance des institutions et des chercheurs dans la légitimité des projets.

C. Sécurité des données

La réutilisation des données médicales impose de garantir la confidentialité, la protection des informations sensibles et la conformité réglementaire (RGPD, directives locales). La sécurité des EDC n'est pas uniquement une question technique, mais elle repose également sur des

politiques organisationnelles claires et une culture de la protection des données au sein des hôpitaux.

Les stratégies identifiées pour assurer la sécurité incluent la pseudonymisation ou anonymisation systématique des données pour la recherche, la mise en place de contrôles d'accès stricts et traçables, la surveillance continue des systèmes, notamment par de nouveaux outils de traçabilité [26] pour détecter les intrusions ou usages non autorisés, et la formation des utilisateurs à la sensibilisation à la sécurité et aux obligations réglementaires [40].

L'articulation entre sécurité et accessibilité est particulièrement critique : un accès trop restreint limite la recherche, tandis qu'un accès trop laxiste augmente les risques pour les patients. En somme, l'accessibilité des données en santé repose sur un équilibre délicat entre ouverture scientifique et maîtrise du risque. Elle exige une approche intégrée combinant innovations technologiques, cadre réglementaire rigoureux et culture partagée de la sécurité des données.

D. Recherche textuelle

La recherche textuelle dans les EDC représente un pilier fondamental pour exploiter le contenu non structuré des DME. Les textes cliniques — notes de suivi, comptes rendus hospitaliers, prescriptions, observations et rapports opératoires — contiennent une richesse informationnelle considérable, mais leur hétérogénéité lexicale, syntaxique et sémantique constitue un défi majeur pour la réutilisation automatisée à des fins de recherche ou d'aide à la décision clinique.

Le recours au NLP permet de structurer ces informations, d'extraire des entités médicales pertinentes et de construire des variables exploitables dans des analyses

épidémiologiques et cliniques. La méthode la plus généralement utilisée est celle du pipeline. Elle désigne un ensemble structuré d'étapes visant à identifier, parmi un corpus documentaire, les informations les plus pertinentes en réponse à une requête exprimée en langage naturel. Ce processus repose sur l'intégration de techniques issues à la fois de la recherche d'information classique et des approches modernes d'apprentissage profond en traitement automatique du langage. La première étape consiste à préparer les données textuelles afin de faciliter leur exploitation ultérieure [37,39]. Ce prétraitement linguistique inclut généralement la tokenisation (segmentation du texte en unités lexicales), la lemmatisation ou la racinisation (réduction des mots à leur forme canonique), ainsi que la suppression des mots vides. D'autres opérations de normalisation, telles que la conversion en minuscules, la gestion des accents ou la suppression des caractères spéciaux, peuvent également être appliquées. L'objectif est d'obtenir une représentation textuelle standardisée, réduisant la variabilité linguistique et améliorant la cohérence des représentations. La seconde phase vise à extraire, parmi l'ensemble du corpus, un sous-ensemble de documents candidats susceptibles de contenir l'information recherchée. Deux familles d'approches sont couramment employées : les méthodes lexicales, fondées sur la fréquence et la pondération des termes (recherche booléenne, TF-IDF), permettent une recherche rapide et efficace fondée sur la correspondance exacte des mots-clés ; les approches vectorielles, quant à elles, représentent les textes et les requêtes sous forme d'*embeddings* (vecteurs numériques denses) et mesurent leur similarité sémantique à l'aide d'indicateurs tels que la similarité cosinus. L'enrichissement des *embeddings* par des terminologies médicales (ICD-10, SNOMED CT, HPO) a été développée dans des travaux issus de l'EDS [28, 31] améliorant la détection des entités rares ou ambiguës et permettant d'accroître la robustesse des modèles. Les documents identifiés lors de la récupération initiale sont ensuite réévalués selon des critères sémantiques plus avancés. Cette

étape repose sur des modèles de langage contextuels tels que BERT, RoBERTa ou CamemBERT, capables de modéliser les relations entre mots au sein d'un contexte donné. La dernière étape du pipeline consiste à synthétiser ou reformuler l'information extraite notamment via la fusion hybride de scores lexicaux et sémantiques afin d'optimiser la pertinence globale ou par la génération de résumés automatiques mettant en évidence les éléments essentiels des documents. Au-delà de l'identification d'entités, certains pipelines détectent les relations patient–condition ou condition–traitement, permettant la construction de profils cliniques multilabel et la création de cohortes homogènes à partir des similarités phénotypiques [31, 34, 35].

E. Qualité des données

La qualité des données cliniques constitue un déterminant essentiel de la fiabilité des travaux de recherche, de la pertinence des décisions cliniques et de la performance des modèles prédictifs. Elle regroupe un ensemble de dimensions complémentaires visant à garantir la validité, la cohérence et la traçabilité des informations issues des DME et des EDS.

Parmi les dimensions fondamentales de la qualité des données, la complétude correspond à la proportion de champs correctement renseignés et à la limitation des valeurs manquantes. L'exactitude, ou plausibilité, renvoie à la cohérence des valeurs observées avec les standards attendus ou les connaissances cliniques établies. La conformité évalue le respect des formats, des codifications et des terminologies standards telles que le CIM-10 ou les unités de mesure. La temporalité mesure la stabilité et la qualité des données dans le temps, permettant de détecter les ruptures associées à des modifications protocolaires ou à des évolutions de systèmes. Enfin, la robustesse des pipelines de traitement concerne la capacité des chaînes automatisées, telles que les processus ETL et les outils de NLP, à absorber les

variations dans les flux de données, à corriger les erreurs de saisie et à s'adapter aux mises à jour des systèmes cliniques.

L'évaluation de la qualité des données repose sur la combinaison d'indicateurs quantitatifs, tels que les taux de complétude et de cohérence inter-variable, d'analyses statistiques visant à détecter des anomalies ou des ruptures temporelles, ainsi que d'approches correctives incluant la révision des processus ETL, la formation des utilisateurs et l'automatisation des contrôles. Dans les environnements hospitaliers complexes, l'intégration de workflows collaboratifs et de pipelines reproductibles s'avère indispensable pour assurer la pérennité de la qualité des données dans le temps.

Un exemple notable d'initiative institutionnelle est le programme de *Data Quality* conduit à l'AP-HP par Daniel et al. [19]. Ce programme visait à renforcer le contrôle et l'amélioration de la qualité des données cliniques à grande échelle. Deux campagnes pilotes ont été menées sur des données de patients et de services hospitaliers. Un profilage semi-automatisé reposant sur dix-sept indicateurs de qualité, incluant la complétude, la plausibilité et la conformité, a permis d'identifier onze problèmes majeurs, principalement liés aux processus ETL et aux limitations structurelles des DME. La mise en œuvre de plans d'action correctifs et de boucles de rétroaction a conduit à une amélioration mesurable de la fiabilité des données. Cette étude illustre l'efficacité d'une approche itérative, interdisciplinaire et gouvernée pour renforcer la qualité de l'information hospitalière.

VI. Conclusion

L'émergence des EDS s'inscrit dans une transformation profonde de l'épidémiologie, désormais façonnée par la donnée numérique et la puissance analytique des outils informatiques. L'EDS AP-HP, par son envergure, son intégration et son ambition, constitue une illustration emblématique de cette mutation. Il offre une plateforme unique d'agrégation, de standardisation et d'exploitation des données hospitalières, permettant d'aborder des problématiques épidémiologiques complexes avec une puissance statistique et une granularité inédites.

Les travaux menés dans le cadre de cette thèse démontrent concrètement la valeur scientifique et opérationnelle de l'EDS. En mobilisant ses composantes structurées (codages PMSI, actes, biologie) et non structurées (textes cliniques), il a permis la constitution rapide et fiable de cohortes contrastées : l'une portant sur une affection fréquente et répondant à plusieurs codes PMSI (infection par le VRS), l'autre sur une affection rare et non codée (DRESS). Ces études illustrent la capacité de l'EDS à soutenir des projets de nature épidémiologique très différente, tout en exploitant des méthodes variées – du codage standardisé à l'extraction textuelle automatisée – pour produire des résultats scientifiquement robustes.

Cependant, cette promesse s'accompagne de limites concrètes. Le taux de publication encore faible des projets déposés (21%) témoigne de la persistance d'obstacles techniques, organisationnels et humains : centralisation de l'accès via des environnements sécurisés complexes (tel que JupyterHub), hétérogénéité et qualité variable des données, ou encore compétences inégales des utilisateurs face aux outils analytiques. Si la richesse du contenu constitue un atout majeur, la difficulté d'extraction et de standardisation des données non structurées demeure un frein à leur pleine exploitation.

L'analyse des productions scientifiques issues de l'EDS souligne également une forte disparité thématique : un poids prédominant des études méthodologiques et infectiologiques, notamment durant la pandémie de COVID-19, contraste avec la faible représentation d'autres champs comme la génétique. Cela reflète à la fois la maturité inégale des domaines d'application et la nécessité de renforcer les passerelles entre recherche clinique, data science et infrastructures techniques.

À l'échelle institutionnelle, le futur de l'EDS repose sur trois défis majeurs. Le premier est l'accessibilité, à la fois technique et cognitive : simplifier les environnements d'accès, tout en développant des interfaces intuitives et des formations adaptées pour les chercheurs. Le second est la qualité et la standardisation des données : leur fiabilité conditionne la reproductibilité scientifique et la comparabilité internationale. Enfin, le troisième est la sécurité et la gouvernance : la confiance des patients et des institutions repose sur la maîtrise des enjeux éthiques, juridiques et techniques liés à la donnée de santé.

En définitive, l'EDS AP-HP se positionne comme une infrastructure stratégique pour la recherche biomédicale française, capable d'unir puissance numérique et rigueur scientifique. Sa consolidation – par l'amélioration de l'interopérabilité, la simplification des accès et l'intégration de nouvelles dimensions comme la génétique – conditionnera sa capacité à devenir un outil de référence pour l'épidémiologie moderne et la médecine de précision.

VII. Bibliographie

1. Schlienger JL. A brief history of epidemiology, from ancient times up to the present day. Médecine des Maladies Métaboliques. Volume 16, Issue 2, March 2022, Pages 191-199

2. Louis P.-Ch.-A. (1835). Recherches sur les effets de la saignée dans quelques maladies inflammatoires, et sur l'action de l'émétique et des vésicatoires dans la pneumonie. Paris : J.-B. Baillière. https://gallica.bnf.fr/ark:/12148/bpt6k103986t/f62.double

3. La Berge AF. Medical Statistics at the Paris School: What was at Stake? In: Jorland G, Opinel A, Weisz G, eds. Body Counts: Medical Quantification in Historical and Sociological Perspectives. Montreal, Canada: McGill-Queen's University Press; 2005:89-108

4. Armstrong GL, Conn LA, Pinner RW. Trends in infectious disease mortality in the United States during the 20th century. JAMA. 1999 Jan 6;281(1):61-6.

5. Omran AR. The epidemiologic transition: a theory of the epidemiology of population change. 1971. Milbank Q. 2005;83(4):731-57.

6. Lehmann EL (2011). Fisher, Neyman, and the Creation of Classical Statistics. New York: Springer.

7. Doll R, Hill AB. Smoking and carcinoma of the lung; preliminary report. Br Med J. 1950 Sep 30;2(4682):739-48.

8. Sackett DL. Bias in analytic research. J Chronic Dis 1979; 32:51-63

9. Zuccolo L, Holmes MV. Commentary: Mendelian randomization-inspired causal inference in the absence of genetic data. Int J Epidemiol. 2017 Jun 1;46(3):962-965

10. Francis T, Korns RF, Voight RB and al. An evaluation of the 1954 poliomyelitis vaccine trials. Am J Public Health Nations Health. 1955 May;45(5 Pt 2):1-63.

11. De Mauro A, Greco M, Grimaldi M. (2016). A formal definition of Big Data based on its essential features. Library Review, 65(3), 122–135.

12. Hilbert M, López P (2011). The World's Technological Capacity to Store, Communicate, and Compute Information. Science, 332(6025), 60 –65.

13. Institut Pasteur. Comment le Big data révolutionne la recherche en santé. Retrieved October 28, 2025. https://www.pasteur.fr/fr/journal-recherche/dossiers/comment-big-data-revolutionne-recherche-sante

14. Kliemann N, Rauber F, Levy RB. Food processing and cancer risk in Europe: results from the prospective EPIC cohort study. Lancet Planet Health. 2023 Mar;7(3):e219-e232.

15. Sauer CM, Dam TA, Celi LA et al. Systematic Review and Comparison of Publicly Available ICU Data Sets-A Decision Guide for Clinicians and Data Scientists. Crit Care Med. 2022 Jun 1;50(6):e581-e588.

16. Kallout J, Lamer A, Grosjean J et al. Contribution of Open Access Databases to Intensive Care Medicine Research: Scoping Review. J Med Internet Res. 2025 Jan 9:27:e57263. doi: 10.2196/57263.

17. Ke Y, Yang R, Liu N. Comparing Open-Access Database and Traditional Intensive Care Studies Using Machine Learning: Bibliometric Analysis Study. J Med Internet Res. 2024 Apr 17:26:e48330. doi: 10.2196/48330.

18. Daniel C, Serre P, Orlova N, Bréant S, Paris N, Griffon N. Initializing a hospital-wide data quality program. The AP-HP experience. Comput Methods Programs Biomed. 2019 Nov:181:104804.

19. Daniel C, Paris N, Pierre O et al. AP-HP Health Data Space (AHDS) to the Test of the Covid-19 Pandemic. Stud Health Technol Inform. 2022 May 25:294:28-32.

20. Murphy SN, Weber G, Mendis M et al. Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). J Am Med Inform Assoc. 2010 Mar-Apr;17(2):124-30.

21. Paris N, Mendis M, Daniel C, Murphy S, Tannier X, Zweigenbaum P. i2b2 implemented over SMART-on-FHIR. AMIA Jt Summits Transl Sci Proc. 2018 May 18:2017:369-378. eCollection 2018.

22. Hripcsak G, Duke JD, Shah NH et al. Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers. Stud Health Technol Inform. 2015:216:574-8.

23. Health Data Hub. Health Data Hub – La plateforme nationale des données de santé en France. Retrieved October 28, 2025, from https://www.health-data-hub.fr/

24. Gaillet A, Layese R, Fourati S et al. Clinical Phenotypes and Outcomes Associated With Respiratory Syncytial Virus Infection in Critically Ill Patients: A Retrospective Multicenter Cohort Study in Greater Paris Area Hospitals, 2017-2023. J Infect Dis. 2025 Sep 15;232(3):679-690.

25. Kridin K, Brüggen MC, Walsh S et al. Management and treatment outcome of DRESS patients in Europe: An international multicentre retrospective study of 141 cases. J Eur Acad Dermatol Venereol. 2023 Apr;37(4):753-762.

26. Wack M, Coulet A, Burgun A, Rance B. Enhancing clinical data warehousing with provenance data to support longitudinal analyses and large file management: The gitOmmix approach for genomic and image data. J Biomed Inform. 2025 Mar:163:104788.

27. Tannier X, Wajsbürt P, Callige A et al. Development and Validation of a Natural Language Processing Algorithm to Pseudonymize Documents in the Context of a Clinical Data Warehouse. Methods Inf Med. 2024 May;63(1-02):21-34.

28. Gérardin C, Mageau A, Mékinian A, Tannier X, Carrat F. Construction of Cohorts of Similar Patients From Automatic Extraction of Medical Concepts: Phenotype Extraction Study. JMIR Med Inform. 2022 Dec 19;10(12):e42379.

29. Gérardin C, Xiong Y, Wajsbürt P, Carrat F, Tannier X. Impact of Translation on Biomedical Information Extraction: Experiment on Real-Life Clinical Notes. JMIR Med Inform. 2024 Apr 4:12:e49607.

30. Lerner I, Paris N, Tannier X. Terminologies augmented recurrent neural network model for clinical named entity recognition. J Biomed Inform. 2020 Feb:102:103356.

31. Gérardin C, Wajsbürt P, Vaillant P, Bellamine A, Carrat F, Tannier X. Multilabel classification of medical concepts for patient clinical profile identification. Artif Intell Med. 2022 Jun:128:102311.

32. Wajsbürt P, Sarfati A, Tannier X. Medical concept normalization in French using multilingual terminologies and contextual embeddings. J Biomed Inform. 2021 Feb:114:103684.

33. Looten V, Kong Win Chang L, Neuraz A, Rance B et al. What can millions of laboratory test results tell us about the temporal aspect of data quality? Study of data spanning 17 years in a clinical data warehouse. Comput Methods Programs Biomed. 2019 Dec;181:104825. doi:10.1016/j.cmpb.2018.12.030.

34. Garcelon N, Neuraz A, Salomon R et al. Next generation phenotyping using narrative reports in a rare disease clinical data warehouse. Orphanet J Rare Dis. 2018;13:85. doi:10.1186/s13023-018-0830-6.

35. Garcelon N, Neuraz A, Salomon R et al. A clinician friendly data warehouse oriented toward narrative reports: Dr. Warehouse. J Biomed Inform. 2018 Mar;80:52-63. doi:10.1016/j.jbi.2018.02.019.

36. Garcelon N, Neuraz A, Benoit V et al. Finding patients using similarity measures in a rare diseases-oriented clinical data warehouse: Dr. Warehouse and the needle in the needle stack. J Biomed Inform. 2017 Jul;73:121-129. doi:10.1016/j.jbi.2017.07.016.

37. Karunakaran S, Van Gysel D, Guinemer S, Mahe L, Sallah K. Reconstruction de variables structurées à partir des données textuelles d'un entrepôt de données de santé, à des fins de recherche clinique, Paris. Revue d'Épidémiologie et de Santé Publique. Volume 68, Supplement 1, March 2020, Page S28

38. Garcelon N, Neuraz A, Benoit V, Salomon R, Burgun A. Improving a full-text search engine: the importance of negation detection and family history context to identify cases in a biomedical data warehouse. J Am Med Inform Assoc. 2017 May 1;24(3):607-613.

39. Petit-Jean T, Gérardin E, Berthelot E et al. Collaborative and privacy-enhancing workflows on a clinical data warehouse: an example developing natural language processing pipelines to

detect medical conditions. J Am Med Inform Assoc. 2024 Jun;31(6):1280-1290. doi:10.1093/jamia/ocae069.

40. Priou S, Kempf E, Jankovic M, Lamé G. “Goldmine” or “big mess”? An interview study on the challenges of designing, operating, and ensuring the durability of Clinical Data Warehouses in France and Belgium. J Am Med Inform Assoc. 2024;31(11):2699-2707. doi:10.1093/jamia/ocae244.

41. Kohane IS, Aronow BJ, Avillach P et al. What Every Reader Should Know About Studies Using Electronic Health Record Data but May Be Afraid to Ask. J Med Internet Res. 2021 Mar 2;23(3):e22219.

42. Kempf E, Vaterkowski M, Leprovost D et al. How to Improve Cancer Patients ENrollment in Clinical Trials From rEal-Life Databases Using the Observational Medical Outcomes Partnership Oncology Extension: Results of the PENELOPE Initiative in Urologic Cancers. JCO Clin Cancer Inform. 2023 May:7:e2200179.

43. Petit-Jean T, Gérardin C, Berthelot E et al. Collaborative and privacy-enhancing workflows on a clinical data warehouse: an example developing natural language processing pipelines to detect medical conditions. J Am Med Inform Assoc. 2024 May 20;31(6):1280-1290.

44. Neuraz A, Lerner I, Digan W et al. Natural Language Processing for Rapid Response to Emergent Diseases: Case Study of Calcium Channel Blockers and Hypertension in the COVID-19 Pandemic. J Med Internet Res. 2020 Aug 14;22(8):e20773.

45. Bannour N, Wajsbürt P, Rance B, Tannier X, Névéol A. Privacy-preserving mimic models for clinical named entity recognition in French. J Biomed Inform. 2022;130:104073. doi:10.1016/j.jbi.2022.104073.

46. Bataille P, Amiot A, Claudepierre P et al. Letter: severe COVID-19 infection and biologic therapies – a cohort study of 7 808 patients in France. Aliment Pharmacol Ther. 2020 Oct;52(7):1245-1248. doi:10.1111/apt.16040 PubMed+1

47. Pourcher V, Tubach F, Rozes A, Boussouar S, Estellat C. Outcomes of coronavirus disease 2019-related hospitalization among people with HIV: historical cohort from the Greater Paris area multicenter hospital data warehouse. AIDS. 2023 Dec;37(12):?–?. doi:10.1097/QAD.0000000000003655

48. Abbara S, Guillemot D, El Oualydy S et al. Antimicrobial Resistance and Mortality in Hospitalized Patients with Bacteremia in the Greater Paris Area from 2016 to 2019. Clin Epidemiol. 2022;14:1547-1560. doi:10.2147/CLEP.S385555 Research

49. Benderra MA, Aparicio A, Leblanc J et al. Clinical Characteristics, Care Trajectories and Mortality Rate of SARS-CoV-2 Infected Cancer Patients: A Multicenter Cohort Study. Cancers (Basel). 2021 Sep 23;13(19):4749.

50. Czernichow S, Beeker N, Rives-Lange C et al. Obesity Doubles Mortality in Patients Hospitalized for Severe Acute Respiratory Syndrome Coronavirus 2 in Paris Hospitals, France: A Cohort Study on 5,795 Patients. Obesity (Silver Spring). 2020 Dec;28(12):2282-2289.

51. Sánchez-Rico M, Limosin F, Vernet R et al. Hydroxyzine Use and Mortality in Patients Hospitalized for COVID-19: A Multicenter Observational Study. J Clin Med. 2021 Dec 15;10(24):5891.

52. Hoertel N, Sánchez-Rico M, Vernet R et al. Observational Study of Chlorpromazine in Hospitalized Patients with COVID-19. Clin Drug Investig. 2021 Mar;41(3):221-233.

53. Hoertel N, Sánchez-Rico M, Gulbins E et al. Association Between FIASMAs and Reduced Risk of Intubation or Death in Individuals Hospitalized for Severe COVID-19: An Observational Multicenter Study. Clin Pharmacol Ther. 2021 Dec;110(6):1498-1511.

54. Hoertel N, Sánchez-Rico M, Herrera-Morueco JJ et al. Comorbid medical conditions are a key factor to understand the relationship between psychiatric disorders and COVID-19-related mortality: Results from 49,089 COVID-19 inpatients. Mol Psychiatry. 2022 Mar;27(3):1278-1280.

55. Hoertel N, Sánchez-Rico M, Vernet R et al. Observational study of haloperidol in hospitalized patients with COVID-19. PLoS One. 2021 Feb 19;16(2):e0247122.

56. Hoertel N, Sánchez-Rico M, Vernet R et al. Dexamethasone use and mortality in hospitalized patients with coronavirus disease 2019: A multicentre retrospective observational study. Br J Clin Pharmacol. 2021 Oct;87(10):3766-3775.

57. Hoertel N, Sánchez-Rico M, Vernet R et al. Association between antidepressant use and reduced risk of intubation or death in hospitalized patients with COVID-19: results from an observational study. Mol Psychiatry. 2021 Sep;26(9):5199-5212.

58. Mustafic H, Fayssoil A, Josseran L et al. Impact of Angiotensin-Converting Enzyme Inhibitors and Angiotensin II Receptor Blockers in Hypertensive Patients with COVID-19 (COVIDECA Study). Am J Cardiol. 2021 May 15:147:58-60.

59. Weber GM, Zhang HG, L'Yi S et al. International Changes in COVID-19 Clinical Trajectories Across 315 Hospitals and 6 Countries: Retrospective Cohort Study. J Med Internet Res. 2021 Oct 11;23(10):e31400.

60. Estiri H, Strasser ZH, Brat GA, Semenov YR; Consortium for Characterization of COVID-19 by EHR (4CE); Patel CJ, Murphy SN. Evolving phenotypes of non-hospitalized patients that indicate long COVID. BMC Med. 2021 Sep 27;19(1):249.

61. Bourgeois FT, Gutiérrez-Sacristán A, Keller MS et al. International Analysis of Electronic Health Records of Children and Youth Hospitalized With COVID-19 Infection in 6 Countries. JAMA Netw Open. 2021 Jun 1;4(6):e2112596.

62. Jannot AS, Countouris H, Straaten AV, Burgun A, Katsahian S, Rance B; AP-HP/Universities/Inserm COVID-19 research collaboration. Low-income neighbourhood was a key determinant of severe COVID-19 incidence during the first wave of the epidemic in Paris. J Epidemiol Community Health. 2021 Dec;75(12):1143-1146.

63. Leblanc J, Dusserre-Telmon L, Chauvin A et al. Intensified screening for SARS-CoV-2 in 18 emergency departments in the Paris metropolitan area, France (DEPIST-COVID): A cluster-randomized, two-period, crossover trial. PLoS Med. 2023 Dec 7;20(12):e1004317.

64. Neuraz A, Lerner I, Diganet W et al. Natural Language Processing for Rapid Response to Emergent Diseases: Case Study of Calcium Channel Blockers and Hypertension in the COVID-19 Pandemic. J Med Internet Res. 2020 Aug 14;22(8):e20773.

65. Bottani S, Burgos N, Maire A, Wild A, Ströer S, Dormont D, Colliot O; APPRIMAGE Study Group. Automatic quality control of brain T1-weighted magnetic resonance images for a clinical data warehouse. Med Image Anal. 2021 Jul;75:102219.

66. Loizillon S, Bottani S, Maire A et al. Automatic quality control of brain 3D FLAIR MRIs for a clinical data warehouse. Med Image Anal. 2025 May;103:103617.

67. Roux C, Rozes A, Reizine D et al. Fully automated opportunistic screening of vertebral fractures and osteoporosis on more than 150,000 routine computed tomography scans. Rheumatology (Oxford). 2022 Aug 3;61(8):3269-3278.

68. Hamzaoui D, Montagne S, Granger B et al. Prostate volume prediction on MRI: tools, accuracy and variability. Eur Radiol. 2022 Jul;32(7):4931-4941.

69. Revel MP, Beeker N, Porcher R et al. What level of D-dimers can safely exclude pulmonary embolism in COVID-19 patients presenting to the emergency department? Eur Radiol. 2022 Apr;32(4):2704-2712.

70. Wang P, Martel P, El Hajjam M, Grimaldi L, Giroux Leprieur E; 'AP-HP / Universities / Inserm COVID-19 research collaboration and AP-HP Covid CDW Initiative. Incidental diagnosis of lung cancer on chest CT scan performed for suspected or documented COVID-19 infection Respir Med Res. 2024 Jun:85:101084.

71. Suc G Dewavrin T, Mesnier J et al. Cardiac magnetic resonance imaging-derived right ventricular volume and function, and association with outcomes in isolated tricuspid regurgitation. Arch Cardiovasc Dis. 2025 Jan;118(1):43-51.

72. Bottani S, Elina Thibeau-Sutre 1, Aurélien Maire et al. Contrast-enhanced to non-contrast-enhanced image translation to exploit a clinical data warehouse of T1-weighted brain MRI. BMC Med Imaging. 2024 Mar 20;24(1):67.

73. Loizillon S, Bottani S, Maire A et al. Automatic motion artefact detection in brain T1-weighted magnetic resonance images from a clinical data warehouse using synthetic data. Med Image Anal. 2024;93:103073.

74. Bottani S, Burgos N, Maire A et al. Evaluation of MRI-based machine learning approaches for computer-aided diagnosis of dementia in a clinical data warehouse. Med Image Anal. 2023;89:102903.Haut du formulaire

75. Barclay ME, Cohen A, Priou S et al. Emergency cancer diagnosis in Paris: A cross-sectional study using AP-HP data. Int J Cancer. 2025 Dec 1;157(11):2283-2293.

76. Aparicio T, Layese R, Hemery F et al. Effect of lockdown on digestive system cancer care amongst older patients during the first wave of COVID-19: The CADIGCOVAGE multicentre cohort study. Dig Liver Dis. 2022 Jan;54(1):10-18.

77. Kempf E, Lamé G, Layese R et al. New cancer cases at the time of SARS-Cov2 pandemic and related public health policies: A persistent and concerning decrease long after the end of the national lockdown. Eur J Cancer. 2021;150:260-267.

78. Rohr O, Priou S, Chatellier G et al. Prevalence and risks of intravenous chemotherapy-induced severe neutropenia in solid cancers: a multicenter retrospective cohort study on real-life data. Support Care Cancer. 2024 Sep 13;32(10):657.

79. Guével E, Priou S, Lamé G et al. Impact of the COVID-19 pandemic on clinical presentation, treatments, and outcomes of new breast cancer patients: A retrospective multicenter cohort study. Cancer Med. 2023 Nov;12(22):20918-20929.

80. Guével E, Priou S, Flicoteaux R et al. Development of a natural language processing model for deriving breast cancer quality indicators : A cross-sectional, multicenter study. Rev Epidemiol Sante Publique. 2023 Dec;71(6):102189.

81. Fenioux C, Abbar B, Boussouar S et al. Thymus alterations and susceptibility to immune checkpoint inhibitor myocarditis. Nat Med. 2023 Dec;29(12):3100-3110.

82. Kempf E, Priou S, Lamé G et al. No changes in clinical presentation, treatment strategies and survival of pancreatic cancer cases during the SARS-COV-2 outbreak: A retrospective multicenter cohort study on real-world data. Int J Cancer. 2023 Dec 15;153(12):1988-1996.

83. Barclay ME, Cohen A, Priou S et al. Emergency cancer diagnosis in Paris: A cross-sectional study using AP-HP data. Int J Cancer. 2025 Aug;157(11):2283-2293.

84. Naouri D, Yordanov Y, Lapidus N, Pelletier-Fleury N. Cost-effectiveness analysis of direct admission to acute geriatric unit versus admission after an emergency department visit for elderly patients. BMC Geriatr. 2023 May 10;23(1):283. doi: 10.1186/s12877-023-03985-0.

85. El Abd A, Schwab C, Clementz A, Fernandez C, Hindlet P. Safety of Elderly Fallers: Identifying Associated Risk Factors for 30-Day Unplanned Readmissions Using a Clinical Data Warehouse. J Patient Saf. 2022 Apr 1;18(3):230-236.

86. Jean C, Paillaud E, Boudou-Rouquette P et al. Hospital care trajectories of older adults with cancer and the associated clinical profiles. Oncologist. 2025 Mar 10;30(3):oyae301.

87. Jean C, Paillaud E, Boudou-Rouquette P et al. Predicting frailty domain impairments and mortality with the Hospital Frailty Risk Score among older adults with cancer: the ELCAPA-EDS cohort study. Age Ageing. 2024 Oct 1;53(10):afae222.

88. Thietart S, Rozes A, Tubach F et al. In-hospital mortality of older patients with COVID-19 throughout the epidemic waves in the great Paris area: a multicenter cohort study. BMC Geriatr. 2023 Sep 18;23(1):573.

89. Aparicio T, Layese R, Hemery F et al. The 10-month mortality rate among older patients treated for digestive system cancer during the first wave of the COVID-19 pandemic: The CADIGCOVAGE multicentre cohort study. J Geriatr Oncol. 2023 Mar;14(2):101443.

90. Assi R, Schwab C, El Abd A, Fernandez C, Hindlet P. Which Potentially Inappropriate Medications List Can Detect Patients At Risk of Readmissions in the Older Adult Population Admitted for Falls? An Observational Multicentre Study Using a Clinical Data Warehouse. Drugs Aging. 2022 Feb;39(2):175-182.

91. Fartoukh M, Voiriot G, Guérin L et al. Seasonal burden of severe influenza virus infection in the critically ill patients, using the Assistance Publique-Hôpitaux de Paris clinical data warehouse: a pilot study. Ann Intensive Care. 2021 Jul 29;11(1):117.

92. Windsor C, Hua C, De Roux Q et al. Healthcare trajectory of critically ill patients with necrotizing soft tissue infections: a multicenter retrospective cohort study using the clinical data warehouse of Greater Paris University Hospitals. Ann Intensive Care. 2022 Dec 20;12(1):115.

93. Urbina T, Lavillegrand JR, Garnier M et al. Delayed inflammation decrease is associated with mortality in Tocilizumab-treated critically ill SARS-CoV-2 patients: A retrospective matched-cohort analysis. Innate Immun. 2022 Jan;28(1):3-10.

94. Lombardi Y, Azoyan L, Szychowiak P et al. External validation of prognostic scores for COVID-19: a multicenter cohort study of patients hospitalized in Greater Paris University Hospitals. Intensive Care Med. 2021 Dec;47(12):1426-1439.

95. Celante H, Oubaya N, Fourati S et al. Prognosis of hospitalised adult patients with respiratory syncytial virus infection: a multicentre retrospective cohort study. Clin Microbiol Infect. 2023 Jul;29(7):943.e1-943.e8.

96. Marois C, Nedelec T, Pelle J et al. Comparison of Clinical Profiles and Mortality Outcomes Between Influenza and COVID-19 Patients Invasively Ventilated in the ICU: A Retrospective

Study From All Paris Public Hospitals From 2016 to 2021. Crit Care Explor. 2022 Jul 25;4(7):e0737.

97. Destival JB, Urbina T, Bougouin W et al. Severe autoimmune hemolytic anemia in ICU: a place for emergency plasma exchange? A French multicenter retrospective study. Crit Care. 2025 Sep 30;29(1):412.

98. Plomp L, Chassepot H, Psimaras D et al. Features of myositis and myasthenia gravis in patients treated with immune checkpoint inhibitors: a multicentric, retrospective cohort study. Lancet Reg Health Eur. 2025 Jan 14:50:101192.

99. Molto A, Pinson P, Beeker N, Roux C. Evaluation of the prevalence of new-onset musculoskeletal symptoms in patients hospitalized for severe SARS-CoV-2 infection during the first two COVID waves in France: A descriptive analysis of the clinical data warehouse of 39 hospitals in France. Joint Bone Spine. 2022 Nov;89(6):105450.

100. Chevalier K, Genin M, Petit Jean T et al. CovAID: Identification of factors associated with severe COVID-19 in patients with inflammatory rheumatism or autoimmune diseases. Front Med (Lausanne). 2023 Mar 22:10:1152587.

101. Mageau A, Géradin C, Sallah K, Papo T, Sacre K, Timsit JF. Risk of systemic lupus erythematosus flare after COVID-19 hospitalization: A matched cohort study. PLoS One. 2024 Oct 10;19(10):e0309316.

102. Gérard AL, Vieira M, Cohen A, Hassanaly O, Lambert J, Saadoun D. Combotherapies in immune-mediated inflammatory diseases: A study using the Clinical Data Warehouse from Paris Hospitals' Public Assistance. Semin Arthritis Rheum. 2025 Apr:71:152660.

103. Jouffroy J, Feldman SF, Lerner I, Rance B, Burgun A, Neuraz A. Hybrid Deep Learning for Medication-Related Information Extraction From Clinical Texts in French: MedExt Algorithm Development Study. JMIR Med Inform. 2021 Mar 16;9(3):e17934.

104. Gault N, Sallah K, Régent A, Le Jeunne C, Lechat P. Analysis of off-label prescriptions of medicines in hospital in adult patients and feasibility study of their detection using CIM-10 coding. Therapie. 2022 May-Jun;77(3):329-338.

105. Sallah K, N. Gault N, Van Gysel D, Luquiens G, Lechat P. Élaboration d'un référentiel mettant en lien des champs textuels d'indications thérapeutiques avec des libellés diagnostiques, en vue de la détection de prescriptions hors AMM dans les dossiers médicaux. Revue d'Épidémiologie et de Santé Publique. Volume 67, Supplement 3, May 2019, Pages S168-S169

106. Blangis F, Bergerat M, Allali S et al. Adequacy of the Diagnostic Workup of Suspected Physical Abuse in Young Children: An Audit in the Paris Metropolitan Area. Acad Pediatr. 2025 Sep-Oct;25(7):102866.

107. Hennocq Q, Adjed C, Chappuy H et al. Injuries and child abuse increase during the pandemic over 12942 emergency admissions. Injury. 2022 Oct;53(10):3293-3296.

108. Semeraro M, Pinson P, Populaire M et al. Analysis of the impact of the SARS-CoV-2 infection on the pediatric population hospitalized during the pandemic in the Greater Paris University Hospitals. Front Pediatr. 2023 Feb 27:11:1044352.

109. Rech JS, Yao K, Bachmeyer C et al. Prognostic Value of Hyponatremia During Acute Painful Episodes in Sickle Cell Disease. Am J Med. 2020 Sep;133(9):e465-e482.

110. Fayand A, Dzierzynski N, Georgin-Lavialle S et al. The association of psychiatric comorbidities with emergency visits and hospitalisations in adult sickle-cell patients: a cohort study. Br J Haematol. 2020 Apr;189(1):e21-e23.

111. Rech JS, Cohen A, Bartolucci P et al. Shift in emergency department utilization by frequent attendees with sickle cell disease during the COVID-19 pandemic: A multicentre cohort study. Br J Haematol. 2024 Aug;205(2):463-472.

112. Mahé I, Skaff Y, Helfer H et al. Care pathway for patients hospitalized with venous thromboembolism. Eur J Clin Invest. 2025 Apr;55(4):e14383.

113. Coscas R, Dubosq M, Charton J et al. Thrombosis of Medium-Sized and Large Arteries During Covid-19 Infection: Results of the COVIVASC Study. Ann Vasc Surg. 2022 Oct:86:35-42.

114. Gasnier M, Pinson P, Beeker N et al. Acute COVID-19 severity markers predict post-COVID new-onset psychiatric disorders: A 2-year cohort study of 34,489 patients. Mol Psychiatry. 2025 Apr;30(4):1329-1337.

115. Lampros A, Montardi C, Journeau L et al. Association between psychiatric comorbidities and hospital length of stay in post-emergency internal medicine patients. Rev Med Interne. 2020 Jun;41(6):360-367.

116. Sanchis-Borja M, Fallet V, Fabre E et al. Characterization of lung cancers in patients with BRCA germline variants: A multicenter series. Lung Cancer. 2022 Nov:173:67-70.

117. Dupont A, Couffignal C, Arias C et al. Outcomes and risk factors with COVID-19 or influenza in hospitalized asthma patients. Respir Res. 2022 Dec 13;23(1):342.

118. Barret M, Collier M, Tran VT, Dellagi M, Beeker N. Disease burden of Barrett's esophagus across the Paris metropolitan area. Clin Res Hepatol Gastroenterol. 2024 Jun;48(6):102371.

119. Thiebaud PC, Wassermann E, de Caluwe M et al. Assessment of Prognostic Scores for Emergency Department Patients With Upper Gastrointestinal Bleeding. Ann Emerg Med. 2025 Jan;85(1):31-42.

120. Julla JB, Jolivet T, Estellat C et al. Incidence of death and amputation in patients with a first diabetic foot ulcer: results from the CODIA cohort. Diabetes Metab. 2025 Sep 5;51(6):101700.

121. Diedisheim M, Dancoisne E, Gautier JF et al. Diabetes Increases Severe COVID-19 Outcomes Primarily in Younger Adults. J Clin Endocrinol Metab. 2021 Aug 18;106(9):e3364-e3368.

122. Bataille P, Layese R, Claudepierre P et al. Paradoxical reactions and biologic agents: a French cohort study of 9303 patients. Br J Dermatol. 2022 Nov;187(5):676-683.

123. Chaigne B, Tieu A, Beeker N et al. Cluster analysis reveals eosinophilia and fibrosis as poor prognostic markers in 128 patients with eosinophilic fasciitis. J Am Acad Dermatol. 2022 Nov;87(5):997-1005.

124. Le TT, Gutiérrez-Sacristán A, Son J et al. Multinational characterization of neurological phenotypes in patients hospitalized with COVID-19. Sci Rep. 2021 Oct 12;11(1):20238. doi: 10.1038/s41598-021-99481-9.

125. Senot N, Gibier JB, Rabant M et al. Epidemiological and clinicopathological characteristics of vascular-limited renal AL amyloidosis. Nephrol Dial Transplant. 2025 Jun 30;40(7):1396-1407.

126. Detroit M, Collier M, Beeker N et al. Predictive Factors of Response to Immunotherapy in Lymphomas: A Multicentre Clinical Data Warehouse Study (PRONOSTIM). Cancers (Basel). 2023 Aug 9;15(16):4028.

127. Khonsari RH, Bernaux M, Vie JJ et al. Risks of early mortality and pulmonary complications following surgery in patients with COVID-19. Br J Surg. 2021 Apr 30;108(4):e158-e159.

128. Chouchana L, Beeker N, Garcelon N et al. Association of Antihypertensive Agents with the Risk of In-Hospital Death in Patients with Covid-19. Cardiovasc Drugs Ther. 2022 Jun;36(3):483-488. doi: 10.1007/s10557-021-07155-5. Epub 2021 Feb 17.