International Journal of Production Research

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tprs20

Leveraging knowledge graphs in pharmaceutical supply chains: insights into key drivers of drug shortages

Katharina Eberhardt, Amelie Schwärzel, Florian Klaus Kaiser, Sonja Rosenberg & Frank Schultmann

To cite this article: Katharina Eberhardt, Amelie Schwärzel, Florian Klaus Kaiser, Sonja Rosenberg & Frank Schultmann (08 May 2025): Leveraging knowledge graphs in pharmaceutical supply chains: insights into key drivers of drug shortages, International Journal of Production Research, DOI: 10.1080/00207543.2025.2496671

To link to this article: https://doi.org/10.1080/00207543.2025.2496671

Published online: 08 May 2025.

INTERNATIONAL JOURNAL OF PRODUCTION RESEARCH
https://doi.org/10.1080/00207543.2025.2496671

Leveraging knowledge graphs in pharmaceutical supply chains: insights into key drivers of drug shortages

Katharina Eberhardt, Amelie Schwärzel, Florian Klaus Kaiser, Sonja Rosenberg and Frank Schultmann

Institute for Industrial Production (IIP), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

ABSTRACT
Supply chain issues pose a serious challenge to the pharmaceutical industry. The COVID-19 pandemic, raw material and labour bottlenecks, and rising demand from an ageing population have exacerbated disruptions, sparking discussions on the resilience of pharmaceutical supply chains. In Germany, drug shortages have drastically increased, leading to several supply deficits for specific medicines. The complex environment and extensive information necessitate a unified view that is both scalable and flexible. Therefore, we systematically collect and fuse disparate and heterogeneous medical data sources to build a comprehensive knowledge graph. By integrating interconnections between entities, the graph facilitates the identification of critical suppliers and drug patterns. The results reveal a significant increase in shortages in Germany. The drivers of these shortages are multifaceted and primarily relate to production, market, and process issues. Furthermore, we quantify a shortage's severity and economic impact to support targeted measures. Based on our research, we provide decision support and mitigation strategies, enabling the identification and prioritisation of shortages and strengthening resilience. For companies and governments, understanding the drivers of drug shortages is essential for shaping production and pricing strategies and ensuring supply security. Future studies can extend our dataset and methodology to apply artificial intelligence-based approaches.

ARTICLE HISTORY
Received 11 November 2024
Accepted 11 April 2025

KEYWORDS
Knowledge graphs; supply chain disruptions; network analysis; drug shortages; decision support

1. Introduction

In the last decade, supply chain disruptions have emerged as a growing concern in the pharmaceutical sector. The industry faces many challenging events, including natural disasters, pandemics, geopolitical conflicts, regulatory issues, and the emergence of novel medicinal modalities (Badejo and Ierapetritou 2024). Additionally, increasing cost pressures and fierce competition in the market are driving pharmaceutical companies to maintain lower inventory levels and source raw materials more efficiently from global partners (Huss et al. 2023; Shukar et al. 2021).

Moreover, the complex structure and unique demands of the highly regulated pharmaceutical industry render its supply chains susceptible to disruptions. Consequently, an increasing number of drug shortages occur, driven by insufficient production capacities, quality issues, increased demand, and logistical challenges. In 2023, multiple countries faced severe antibiotic shortages due to increased demand and long production lead times, while chemotherapy drugs like cisplatin also experienced significant disruptions (European Commission 2023; Pauwels et al. 2015). Furthermore, the diabetes drug semaglutide (EMA 2024) and essential paediatric drugs, like fever medicine and painkillers, have been in short supply (Huss et al. 2023). The consequences of such medical shortages are severe, leading to disease progression, therapy delays, extended hospitalisations, and exposure to falsified medications (Colin-Oesterlé and Mélin 2020; Phuong et al. 2019). In addition, it affects healthcare providers and the entire healthcare system by increasing efforts and costs to find alternative medication and manage the impacts of supply disruptions (Tucker and Daskin 2022).

Given the pharmaceutical supply chain's pivotal role, comprehensive risk assessment methods are essential to identify the structural causes of shortages and develop efficient management strategies. However, limited visibility often hinders effective strategies and results in delayed detection of supply chain problems, especially at the deeper levels, such as manufacturing issues, capacity bottlenecks, or quality problems (Ivanov 2024). Supply chains generally comprise complex, interconnected networks of stakeholders, where disruptions in one segment can cascade throughout the system (Dolgui and

CONTACT Katharina Eberhardt katharina.eberhardt@kit.edu Institute for Industrial Production (IIP), Karlsruhe Institute of Technology (KIT), Hertzstraße 16, Karlsruhe 76187, Germany

Ivanov 2021). Managing risks arising from these interconnections is challenging due to limited collaboration, hidden information, and the vast array of diverse data types and formats across various locations (Emrouznejad, Abbasi, and Sıcakyüz 2023). This complexity results in significant blind spots for decision-makers, as the disconnectedness makes it challenging to obtain a comprehensive overview of the situation. As a result, the valuable information embedded within the extensive and heterogeneous data sets must be filtered, analysed, and interpreted to detect and mitigate potential disruptions.

Artificial Intelligence (AI), business analytics, and data-driven solutions offer promising approaches to identify criticalities and support strategic decisions by leveraging visibility and transparency (Nguyen et al. 2022). Advances in AI and machine learning have introduced new methods for extracting information from supply chain data, modelling and integrating knowledge, and uncovering supply chain dependencies. Despite these new opportunities, managers often encounter the 'black-box' problem, where the lack of causal understanding in AI models results in a reluctance to integrate these technologies into decision processes (Kosasih et al. 2022). Consequently, Wyrembek, Baryannis, and Brintrup (2025) emphasise the growing need for methodologies that offer actionable insights to guide strategic and data-driven decision-making.

In this regard, Knowledge Graphs (KGs) have emerged as an efficient method for data mining and the structured representation of knowledge (Chaudhri et al. 2022). KGs represent data in a graph-based structure, where graph nodes depict concepts or entities of the natural world, and edges capture the interrelations between these entities (Cimiano and Paulheim 2017). This approach is beneficial in complex structures like supply chains, as it captures intricate interdependencies and allows additional contextual knowledge to arise through relationship structures between entities (Peng et al. 2023). Essentially, a KG is a digital network enriched with additional data, offering an efficient backbone for advanced AI applications (Buchgeher et al. 2021). KGs can enhance decision-making by providing a comprehensive and interconnected process view while enabling data extension without altering the underlying schematic structure. Their primary benefit is machine and human readability while supporting structured queries (Rolf et al. 2022).

Our paper leverages these strengths to address the mentioned issues in supply chain networks. By utilising the capabilities and advantages of knowledge graphs, we tackle data challenges across diverse supply chain tasks, enhancing visibility and decision support. Specifically, the following research questions frame our work:

RQ1: How can diverse entities and datasets be integrated within a KG, and what methods can be employed to identify critical nodes within the network?

RQ2: What insights can be derived from graph data, how can disruption patterns and root causes be identified, and which disruptions are most critical and economically impactful?

In addressing these questions, the main contributions of this study can be summarised as follows:

(1) We provide an ontology-based knowledge modelling method by developing and structuring a KG that systematically organises, links, and integrates data from various sources, facilitating in-depth analysis and enabling intricate queries.
(2) We employ a graph analysis method to identify critical entities within the network, focussing on centrality measures to derive information that is not immediately apparent. Additionally, we leverage query language and data analytics to address challenges in visibility, conducting comprehensive analyses of patterns, uncovering hidden connections, and identifying root causes of disruptions to obtain actionable and transparent findings.
(3) Furthermore, we demonstrate the application and effectiveness of KGs in analysing drug shortages within the pharmaceutical industry. We construct a substantial knowledge base through comprehensive data mining and integrating heterogeneous data sets. By examining the relations between various entities and key drivers, we provide valuable insights into the pharmaceutical network structure. Based on our findings, we offer decision support and prioritisation strategies to enhance informed risk management and strategic planning. In addition, we provide recommendations for potential extensions and future research.

The remainder of this paper is structured as follows. Section 2 reviews the state-of-the-art literature on the application of KGs and the management of disruptions in supply chains, with a particular focus on the pharmaceutical industry and drug shortages. Section 3 describes the methodology of the data sourcing process and framework of the underlying KG. Subsequently, Section 4 offers computational results and an in-depth analysis of the graph. Managerial insights, potential research avenues, and concluding remarks are presented in Sections 5 and 6.

2. Theoretical background and related work

2.1. Knowledge graphs in supply chain management

To effectively manage the complexity and interconnectivity of supply chains, network theory has become a prevalent foundational analytical approach. Stemming from social network analysis (SNA), supply chain systems can be represented as interconnected networks of nodes and edges, where nodes correspond to key entities such as suppliers or manufacturers, while edges define the relationships and interactions between them (Borgatti and Li 2009). Analysing structural properties such as network centrality, connectivity, and density helps identify critical actors, supply bottlenecks, and potential points of failure that help to improve the resilience of the supply chain networks (Bier, Lange, and Glock 2020; Kazemian et al. 2022). However, while traditional network analysis methods offer valuable insights into the structural characteristics of supply chain networks, they primarily focus on connectivity and relational patterns, lacking the ability to capture contextual semantics or deeper domain knowledge (Dörpinghaus et al. 2022). This has led to the growing adoption of knowledge graphs as an extension of network-based analysis.

The concept of a KG dates back to the 1960s when semantic networks were introduced as a method to represent knowledge (Ji et al. 2022). In a graph-based structure, knowledge is expressed through interconnected nodes and arcs, forming patterns to depict relationships between various pieces of information (Sowa 2006). This concept of interlinked knowledge is framed by the Semantic Web and the emergence of technologies aiming to represent knowledge in a structured, machine-readable format (Berners-Lee, Hendler, and Lassila 2001). This includes the development of technical standards and unified languages based on formal logic, such as the Resource Description Framework (RDF) and Web Ontology Language (OWL). These ontologies form a set of classes, objects, and data properties to articulate semantic rules that allow the formal representation of a domain's knowledge (Akroyd et al. 2021).

In 2012, Google announced its Knowledge Graph, enhancing Google's search functionality with semantic improvement (Singhal 2012). Since then, the term Knowledge Graph has become established, describing a directed, labelled, heterogeneous graph that includes diverse types of entities and their interrelations (Cimiano and Paulheim 2017). The architecture of a KG consists of different layers, including a data layer and a pattern layer, representing the hierarchical and logical structure of knowledge by defining knowledge classes such as entities, relations, and attributes (Tian et al. 2022). Hence, the network-based structure enables efficient analysis and processing of heterogeneous data within a machine-readable format, leading to substantial research on these solutions in recent years (Tian et al. 2022; G. Zhou et al. 2022). The research on KGs spans various areas, including, as indicated in Table 1, methods of constructing KGs from structured or unstructured data and exploring techniques for utilising KGs across specific domains (Zou 2020).

In the context of supply chain management, KGs have been employed to increase SC resilience, integrate knowledge, improve supplier selection, and enhance SC visibility. Karam et al. (2023) explore the usage of KGs to model activities and entities within a supply chain network. They integrate a Bayesian Network, which enables the early identification of bottlenecks and the timely prediction of disruption impacts. Zhang et al. (2020) construct a KG using multi-source heterogeneous data from daily company transactions, enabling graph mining for enhanced risk analysis. Similarly, Y. Yang et al. (2024) introduce a KG-based supply chain risk management framework, emphasising the role of graph-based models in identifying and managing risks in highly interconnected supply networks. In addition to KG construction, Liu et al. (2023) apply graph analytics, particularly centrality metrics, to support supply chain managers by automatically identifying critical suppliers, enabling closer monitoring and the development of mitigation strategies.

In the field of SC resilience, several studies emphasise the use of KG-based approaches for enhanced decision support. Wyrembek, Baryannis, and Brintrup (2025), for instance, highlight the importance of causal machine learning in supply chain intervention models. While most machine learning models identify correlations rather than causal relationships, the study leverages graph-based causal inference to enhance decision-making processes by identifying dependencies within the supply chain network. For improved decision support, S. Zhou, Liu, and Liu (2024) construct a KG, integrating multidimensional data layers to predict market convergence trends, enhancing supply chain resilience. Rolf et al. (2022) propose a concept utilising a KG and graph-based AI to recommend SC reconfigurations, incorporating a human-centric recommender system to support strategic decision-making. Düggelin and Laurenzi (2024) propose a KG-based decision support system (DSS) for designing resilient supply chain networks, focussing on risk-aware sourcing decisions. Accordingly, Lim et al. (2024) leverage a graph

Table 1. Overview of related research on KGs.

Reference	Objective	Method	Application	Data
AlMahri, Xu, and Brintrup (2024)	SC visibility	KG/LLM	Dependencies, sourcing	Web data
Brockmann, Elson Kosasih, and Brintrup (2022)	SC visibility	KG/GNN	DSCS, Link prediction	Web data
Düggelin and Laurenzi (2024)	SC resilience	KG	Sourcing	Fictitious
Hao et al. (2021)	Integrating knowledge	KG	Product design	Documents
Karam et al. (2023)	SC resilience	KG/BN	Disruptions	N/A
Kosasih et al. (2022)	SC visibility	KG/GNN	DSCS, Link prediction	SC data
L. Li et al. (2020)	Medical KG construction	KG	CDSS	EMR
Y. Li and Starly (2024)	Integrating knowledge	KG/LLM	Service discovery	Web data
Lim et al. (2024)	SC resilience	KG	Disruptions	SC data
Liu et al. (2023)	SC resilience	KG	Link prediction, GA	SC data
Lv et al. (2020)	Supplier selection	KG	Procurement	N/A
Rolf et al. (2022)	SC resilience	KG	Reconfigurable SC	N/A
Rotmensch et al. (2017)	Medical KG construction	KG	Diagnostic reasoning	EMR
Tu et al. (2024)	Supplier selection	KG/GNN	Disruptions	SC data
Wyrembek, Baryannis, and Brintrup (2025)	CML in SCRM	CML	Delay prediction	SC data
F. Yang et al. (2023)	Integrating knowledge	KG	Logistics industry	Web data
P. Yang et al. (2024)	Medical KG construction	KG	CDSS	EMR
Y. Yang et al. (2024)	SC resilience	KG	SCRM	N/A
Zhang et al. (2020)	SC resilience	KG	SCRM	SC data
G. Zhou et al. (2022)	Medical KG construction	KG	CDSS	EMR
S. Zhou, Liu, and Liu (2024)	SC resilience	KG	Market convergence	SC data

BN: Bayesian network, CDSS: Clinical decision support system, CML: Causal machine learning, DSCS: Digital supply chain surveillance, EMR: Electronic medical record, GA: Graph analytics, GNN: Graph neural network, LLM: Large language model, N/A: Not applicable, SC: Supply chain, SCM: Supply chain management, SCRM: Supply chain risk management

embedding-based mitigation DSS to facilitate disruption management by integrating supplier selection and material substitution strategies. This approach, based on a heterogeneous industrial KG, enhances supply chain resilience by mitigating propagated risks in production networks. Lv et al. (2020) and Tu et al. (2024) introduce supplier recommendation using KG-based representation of enterprise data.

While several works specifically address supply chain resilience, other studies focus on integrating and navigating complex decision-related knowledge. For instance, Hao et al. (2021) present a KG-based model that supports decision-making in product design. Similarly, F. Yang et al. (2023) build a KG of the logistics industry, facilitating knowledge sharing and intelligent upstream and downstream analyses of supply chains. In addition, Y. Li and Starly (2024) leverage a KG in combination with a Large Language Model (LLM) to improve manufacturing service discovery for enhanced reliability and interpretability.

Moreover, KG-based approaches have also been employed to increase overall supply chain visibility by addressing the challenge of integrating data from diverse supply chain actors. AlMahri, Xu, and Brintrup (2024) introduce an LLM-driven framework utilising KGs to reveal critical dependencies and alternative sourcing options, enhancing risk management and strategic planning. In contrast, Kosasih et al. (2022) and Brockmann, Elson Kosasih, and Brintrup (2022) target digital supply chain surveillance (DSCS) by using graph neural networks, KGs, and link prediction, allowing inferring and uncovering hidden risks in supply chains to increase the transparency between actors.

The ability of KGs to integrate heterogeneous data, uncover hidden dependencies, and facilitate sophisticated analytics has proven valuable beyond supply chain management, extending into specific domains such as healthcare. In the healthcare sector, the growing volumes of data from Electronic Medical Records (EMR) and biomedical datasets have made KGs essential for representing, storing, and organising knowledge to support clinical decision-making processes (Zou 2020). For instance, Rotmensch et al. (2017) and L. Li et al. (2020) present approaches for constructing medical KGs by mapping diseases to symptoms based on EMR data. The constructed KGs enhance clinical decision-making by information retrieval, diagnostic reasoning, and knowledge transfer. Similarly, P. Yang et al. (2024) investigate named entity recognition and relation extraction to build a large-scale, multi-source, and multi-lingual medical KG that can be used to assist various medical applications. G. Zhou et al. (2022) emphasise the representation and storage of clinical knowledge by applying KGs to tailored recommendations for hypertension medication.

2.2. Disruptions in the pharmaceutical supply chain

According to Snyder et al. (2016), disruptions in the supply chain are stochastic events that can lead to either a partial or complete interruption of operations for an indefinite period. In this regard, supply chain management aims to ensure operational efficiency and mitigate

Figure 1. A generalised pharmaceutical supply chain.

disruption risks arising from market-driven disruptions or external factors such as natural disasters, terrorist attacks, employee strikes, and technological and operational failures (Katsaliaki, Galetsi, and Kumar 2022; Sabouhi, Pishvaee, and Jabalameli 2018).

In the pharmaceutical industry, these risks are particularly critical, where a well-managed supply chain is essential for the reliable delivery of medications to patients worldwide. The pharmaceutical supply chain, outlined by Shah (2004), involves a complex network from procurement to delivering pharmaceutical products, as depicted in Figure 1. Its initial stages include raw material supply and primary manufacturing, which converts raw materials into Active Pharmaceutical Ingredients (API). APIs are defined as the medicinally active components of a pharmaceutical drug. Combined with other ingredients, they diagnose, cure, mitigate, and treat diseases (Kumar et al. 2022). Thereby, the Anatomical Therapeutic Chemical (ATC) classification system, recommended by the World Health Organisation, categorises drugs into different classes based on their APIs and their therapeutic and chemical characteristics (WHO 2024).

The next stage in the pharmaceutical supply chain is secondary manufacturing, where APIs are enhanced with additives to produce final dosage forms. After quality control and packaging, pharmaceutical products undergo initial storage and handling in global warehouses. From there, distribution occurs as the medications are transported to smaller, local distribution centres near the point of sale. Finally, they are delivered to various customers, including pharmacies, hospitals, and other healthcare providers, ensuring patient care.

The generic pharmaceutical supply chain is rather rigid due to the stringent regulatory framework and the specialised nature of manufacturing processes (Tucker and Daskin 2022). Consequently, the system is susceptible to disruptions such as drug shortages. The literature recognises this vulnerability by examining various strategies to enhance resilience within pharmaceutical supply chains. Quantitative studies in this domain encompass optimisation of the supply chain and logistics planning, as indicated by Diaz, Kolachana, and Falcão Gomes (2023) and Haial, Berrado, and Benabbou (2020). In addition, the review by Franco and Alfonso-Lizarazo (2017) underscores a predominant focus on inventory models in the existing body of research. For example, Sabouhi, Pishvaee, and Jabalameli (2018) present a two-stage possibilistic-stochastic programming model for the integrated selection of suppliers and the formulation of supply chain structures amidst disruptions and operational risks. Additionally, they analyse proactive measures, including prepositioning emergency inventory and adopting multiple sourcing strategies to strengthen the resilience of pharmaceutical supply chain companies. Similarly, other researchers consider the roles of inventory policies and reserve capacity in mitigating supply chain disruption risk, such as Azghandi, Griffin, and Jalali (2018), Lücker, Seifert, and Biçer (2019) and Lücker, Chopra, and Seifert (2021), Pathy and Rahimian (2023), and Saedi, Kundakcioglu, and Henry (2016). However, Schmitt and Tomlin (2012) emphasise that strategies like supplier diversification and backup sourcing provide viable options compared to stockpiling inventories.

Besides the presented optimisation approaches, the empirical operations management literature examines the causes and impacts associated with the rising frequency of shortages. For instance, Abu Zwaida, Elaroudi, and Beauregard (2022) identify knowledge gaps through a systematic literature search on drug shortages in the Canadian hospital pharmacy supply chain. Additionally, Ventola (2011) analyses various reports on the drug shortage crisis in the United States, scrutinising different causes, impacts, and management strategies. Likewise, Bogaert et al. (2015) adopt a qualitative approach, including semi-structured interviews, to comprehend the issues that underlie drug shortages in Belgium, France, and the European Union. Employing a mixed-methods approach, Vann Yaroson et al. (2023) investigates the impact of resilience strategies in mitigating medicine shortages in the United Kingdom. Bade et al. (2023) develop a comprehensive framework identifying supply-side causes of drug shortages in Germany from the perspectives of marketing authorisation holders. Based on a structured literature review, data analysis, and semi-structured interviews, this framework provides valuable insights and implications for mitigating such shortages. In addition to qualitative methods, Francas, Mohr, and

Hoberg (2023) employ logistic and negative binomial regression models to analyse how market characteristics, drug substance features, and regulatory factors influence the likelihood of shortages in the German market. Their analysis is based on the same dataset provided by the Federal Institute for Drugs and Medical Devices (BfArM).

To manage and mitigate disruptions in the pharmaceutical industry, Tucker et al. (2020) emphasise the necessity of quantifying shortages, encompassing aspects such as reporting counts of substitutions, affected doses, and impacted treatment groups. In addition, de Vries et al. (2021) highlight the importance of developing an evidence-based systemic view of drug shortages. This approach involves a comprehensive understanding of the causes of shortages, their relative significance, and their interrelationships, complementing publicly available datasets. Thus, there is a lack of advanced, data-driven methodologies capable of handling diverse data sources and unstructured datasets, for instance, retrieved from publicly accessible databases (Nguyen et al. 2022).

Unlike conventional databases, KGs offer enhanced context and logic, enabling tasks like extracting entities and attributes, linking relationships, and modelling ontologies (Shen et al. 2023). This approach allows for the straightforward identification and visualisation of the root causes of disruptions in a detailed and intuitive manner.

Although research on KG-based approaches for managing disruptions remains limited, existing studies highlight their strong potential for early bottleneck identification, timely prediction of network disruption consequences, and practical impact mitigation (Karam et al. 2023; Lim et al. 2024; Tu et al. 2024). Despite these promising applications, using KGs to address pharmaceutical supply chain disruptions remains relatively unexplored.

Therefore, we address the identified research gaps by introducing a KG-based approach to analyse and interconnect the characteristics, causes, and impacts of drug shortages within the pharmaceutical market in Germany. Our study goes beyond identification, offering insightful analysis and practical recommendations to facilitate informed decision-making in addressing and mitigating the challenges posed by drug shortages. Furthermore, we present the application of KGs in an adaptable manner, ensuring its relevance for various use cases.

3. Methodology

3.1. Framework design of the knowledge graph

A KG consists of a heterogeneous graph and its underlying structure. The graph holds the data through a network of nodes and edges, representing the various entities and their relationships. The schema, known as an ontology, provides a formalised structure that ensures consistency and semantic meaning in the data, governing how the data should be interpreted and connected. Before building a KG, we collect data and design an ontology. Once we establish the ontology, we create the KG from the data entries based on this schema. We assign each data item to entities (nodes) and represent relationships between data as edges, following the ontology's definitions.

Figure 2 provides our schematic methodology framework for fusing the datasets and building and visualising the KG. Several steps are performed to construct the graph, including (1) data collection and processing, (2) knowledge extraction and linking, (3) knowledge storage and visualisation, and (4) knowledge analysis and application.

Based on this framework, we ensure a systematic approach to integrating and analysing data for informed decision-making. The steps outlined in our framework are described in more detail in the following sections, providing a comprehensive guide to constructing and leveraging the KG for decision support.

3.2. Data collection and processing

The initial step in constructing the KG is to collect and process data, ensuring the integrity and reliability of the underlying information. Various data sources are identified and gathered for this purpose. Once collected, the data undergoes cleaning to eliminate inconsistencies or errors. Figure 3 illustrates the specific dataset processing procedure for this study. The primary dataset on drug shortages was obtained from BfArM on June 01, 2024 (BfArM 2024b). Additional data sources, including market data, product information, and API details, were collected. The data was then cleaned and combined to create the final dataset, as described in the following.

3.2.1. Reports of drug supply shortages

The primary information on drug shortages is obtained from BfArM, an independent authority within the German Federal Ministry of Health portfolio. BfArM's main tasks include authorising finished medicinal products under the Medicinal Products Act and ensuring their efficacy, safety, and pharmaceutical quality.

In this context, the institute also maintains a database for supply shortages of human-use medicines, excluding vaccines, in Germany (BfArM 2024b). Based on a reporting obligation introduced as part of the Medicinal Products Act (AMG) in 2017, BfArM specifies requirements for pharmaceutical companies to report drug shortages

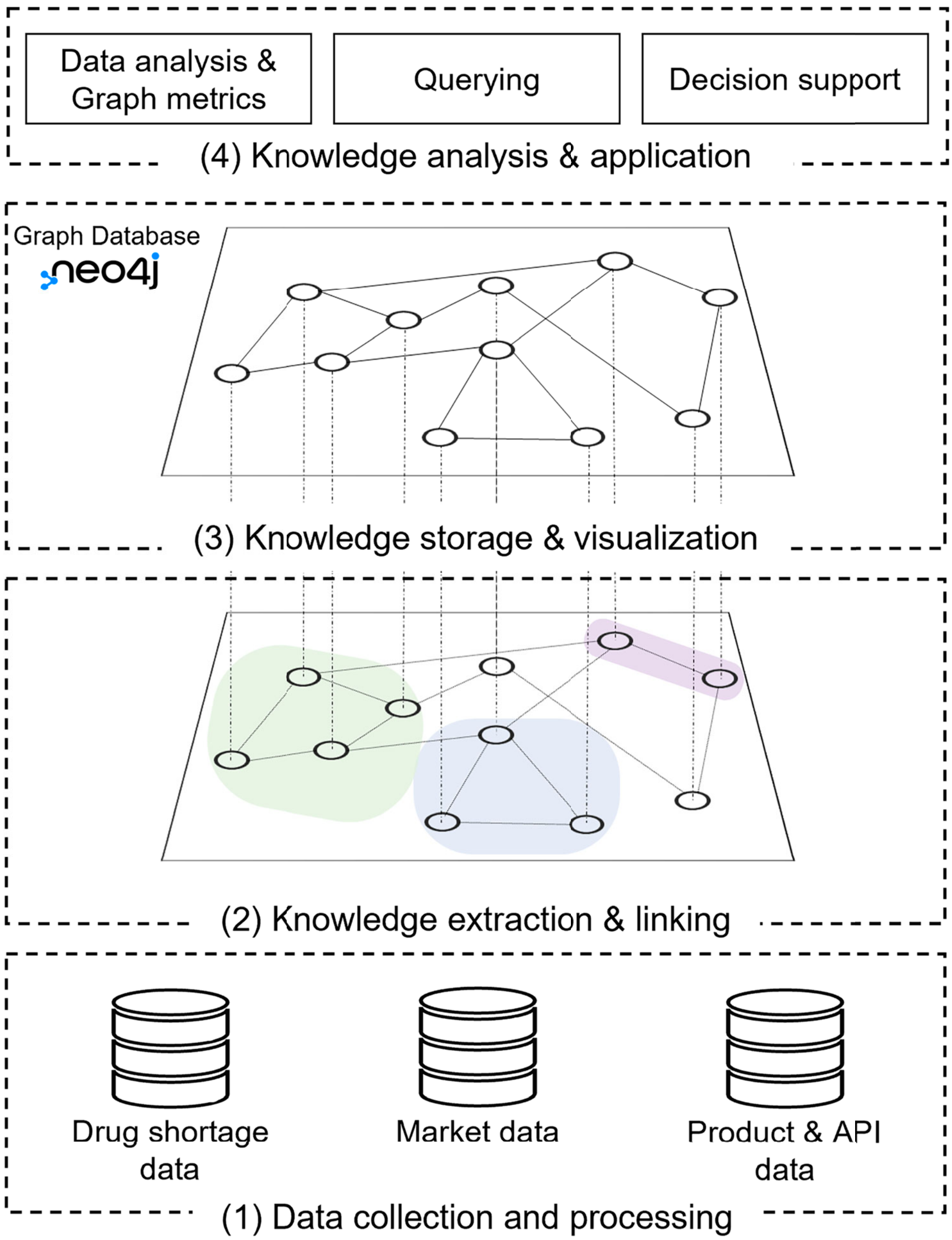


Figure 2. Knowledge Graph framework.

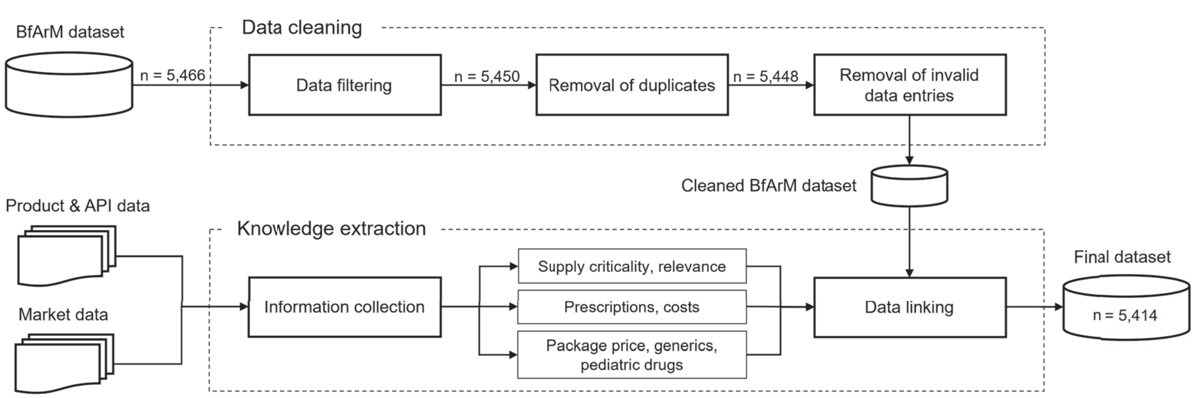


Figure 3. Overview of the data processing procedure with *n* as the number of shortage reports.

(BfArM 2024f). According to BfArM, a drug shortage is defined as either a disruption in the usual supply volume lasting more than two weeks or a significant increase in demand that cannot be adequately met (BfArM 2024a). In a supply bottleneck, the authorisation holder submits an initial notification of the affected product to BfArM. The report contains details on the shortage, such as the start date, the expected end date, and the reason for the occurrence. In particular, information about the affected medicinal product is transmitted, including the pharmaceutical identification number (PZN), submission number (ENR), name, and dosage form. The PZN, as an 8-digit numerical code, represents a standardised identification for medicinal products in Germany. Additionally, the database stores information on the drug's API, the hospital relevance of the drug, and suggestions for alternative medication. Each drug is assigned an 8-digit unique ATC code composed of letters and numbers, classifying the drugs according to their therapeutic use.

3.2.2. Market data

We utilise market data to enrich the drug shortages dataset as an additional source. A health insurance company (DKV AG) provides price data for drugs sold in the German market, including prices for each package size, which we use to supplement the listed medicinal products (DKV 2024). Moreover, a scientific institute of a medical insurance company in Germany (WIdO) collects and evaluates prescription data from statutory health insurance (WIdO 2024). The information provides insights into the quantities and costs of approximately 95% of all drug prescriptions in a given year for individuals insured by statutory health insurance. In Germany, about 90% of the population is insured by statutory health insurance (Destatis 2024). Thus, our study leverages this information to assess volumes and associated costs, enabling a comprehensive evaluation of market-relevant pharmaceutical prescriptions.

3.2.3. Product and API data

We incorporate additional information on medicinal products and APIs to supplement the drug shortages dataset. First, we integrate the ATC classification system provided by the WHO to systematically categorise the APIs into pharmacological groups (PHG) (WHO 2024). In this regard, we enrich the dataset by classifying the APIs according to the first ATC level, containing fourteen main anatomical groups, and the second ATC level of therapeutic subgroups. As part of its reporting obligation, BfArM operates a register of APIs classified according to their supply relevance and criticality in the medical domain, ensuring close surveillance of medications at high risk of supply disruptions. We utilise two lists of 530 supply-relevant and 278 supply-critical APIs published by BfArM according to legal obligations specified within the AMG in Germany AMG(Sec. 52b, para. 3c AMG) (BfArM 2024d, 2024e). Last, we include two further lists indicating if the drugs are for paediatric use (BfArM 2024c) and whether they are generics or biosimilars (ApoKonzept24 2023). This enrichment process of additional product and API data allows a detailed analysis of the APIs effectiveness against diseases and enables a nuanced understanding of the drug supply data.

3.2.4. Data cleaning

The data cleaning process is depicted at the top of the diagram in Figure 3. The initial dataset contained 5, 466 shortage reports of medicinal products with individual package sizes. These reports were thoroughly cleansed before further analysis. First, the data was filtered to remove all records with a notified start date before 2017, aligning the dataset with the reporting obligation introduced in that year. Duplicate records were then identified and removed. To ensure the unified identification of all medicinal products in individual package sizes, entries with missing PZN numbers were supplemented for several medicines through additional research. In cases where no PZN could be found, the corresponding entry was removed from the dataset. Similarly, missing ATC codes were added to ensure the completeness of the dataset. We standardised the data by removing non-machine-readable symbols to maintain uniformity across different datasets, facilitating easier integration and comparison. The final cleaned BfArM dataset used for further analysis included 5, 414 records of drug shortages reported between January 2017 and June 2024.

3.3. Knowledge extraction and linking

As presented in the methodology framework in Figure 2, the next step involves knowledge extraction and linking. Knowledge extraction entails identifying and extracting relevant information from the underlying dataset. In this context, the cleaned data is integrated to form a comprehensive final dataset from which meaningful insights and robust conclusions can be drawn. This process includes identifying entities such as organisations, products, or items and classifying them into predefined categories. Additionally, specific attributes or properties of the entities, such as name, price, or other characteristics, are identified. Knowledge linking involves integrating the extracted information into the KG structure by resolving entities, mapping relationships, and aligning attributes based on an ontology. This process results in entities represented as nodes and relationships as edges in the graph,

Table 2. Nodes and relevant properties of the MedSupplyKG.

'Authz. Holder'	'Report'	'Shortage'	'Drug'	'API'	'PHG'
Name	ID Notification type Notification date Start Expected end	Reason type Reason characteristic Reason comment	PZN, ENR Drug name Generic drug Paediatric drug Hospital relevance Dosage form Alternative drug Package price	ATC Code API name Supply relevance Supply criticality	PHG name Prescriptions DDD Costs

with attributes assigned as properties to the respective nodes.

In the context of our study, we proceed as follows. After collecting and processing the drug shortage dataset, we extract and integrate relevant information from product, API, and market data sources to form a consolidated dataset as depicted in the bottom part of Figure 3. Based on this dataset, we develop a multi-level knowledge graph called MedSupplyKG (Medicinal Supply Knowledge Graph), designed to capture and represent underlying relationships.

Definition: The $MedSupplyKG = \langle \mathbb{V}, \mathbb{E} \rangle$ consists of nodes $\mathbb{V}$ and edges $\mathbb{E}$ that illustrate the connections between entities v_1 and v_2. Each entity pair $v_1, v_2 \in \mathbb{V}$ and their linking relationship $\epsilon \in \mathbb{E}$ can be expressed as a triplet (v_1, ϵ, v_2). An ontology $\mathbb{O}$ is formally defined as $x \in \mathbb{V} \rightarrow \tau, (e_1, e_2) \in \tau \times \tau \rightarrow R_e \in R$, where τ represents the set of entity types and R denotes the set of relationship types.

In this framework, the ontology maps each entity $x \in \mathbb{V}$ to a specific entity type τ, and each pair of entity types $(e_1, e_2) \in \tau \times \tau$ to potential relationship types R_e. Table 2 provides a comprehensive overview of the nodes and their assigned properties.

In addition, Table 3 offers further details on the relationships between entities, providing a comprehensive understanding of the dependencies within the MedSupplyKG. Initially, 'Authorisation holders' transmit disruptions in the availability of medicinal products, which then issue a formal report linked to their authorised drugs. The node 'Report' captures essential data such as ID, notification type, and dates of shortage notifications. Moreover, the 'Report' includes information on the 'Shortage' node, specifying the type and reason for the shortage. This information, in turn, influences specific 'Drug' nodes. The 'Drug' node includes comprehensive details such as identifiers (e.g. PZN, ENR), generic status, paediatric use, hospital relevance, dosage forms, available alternatives, and pricing per package. If available, drugs are linked to an alternative substance, providing a substitute. Each drug contains its 'API', described by its ATC code, name, supply relevance, and criticality level. Additionally, APIs are classified by a specific 'PHG' according to the ATC classification system, which outlines their pharmacological properties and therapeutic uses. 'PHG' nodes are categorised by ATC levels and names, providing additional details such as prescriptions, defined daily dosages (DDD), and associated costs.

Figure 4(a) illustrates the connections between nodes within the graph schema, highlighting the relationships and interactions that define the structure.

3.4. Knowledge storage and visualisation

In the third step, the developed KG structure and its underlying data are stored in graph databases. These databases facilitate the storage of data as nodes, relationships, and properties rather than in traditional tables or documents. Graph databases are highly suited for dealing with relational network data by providing the advantage of visually analysing and storing the complex interconnections between entities, ensuring high transparency and traceability (Pokorný 2015). Additionally, they enable the transformation of the complex relationships in the underlying network data into a structured and scalable form (Neo4j 2024c).

We store the constructed MedSupplyKG in a Neo4j graph database using Python to efficiently manage and analyse it as depicted in Figure 4. Neo4j, an open-source graph database implemented in Java, is widely used and optimised for traversing and querying graph data (Neo4j 2024a).

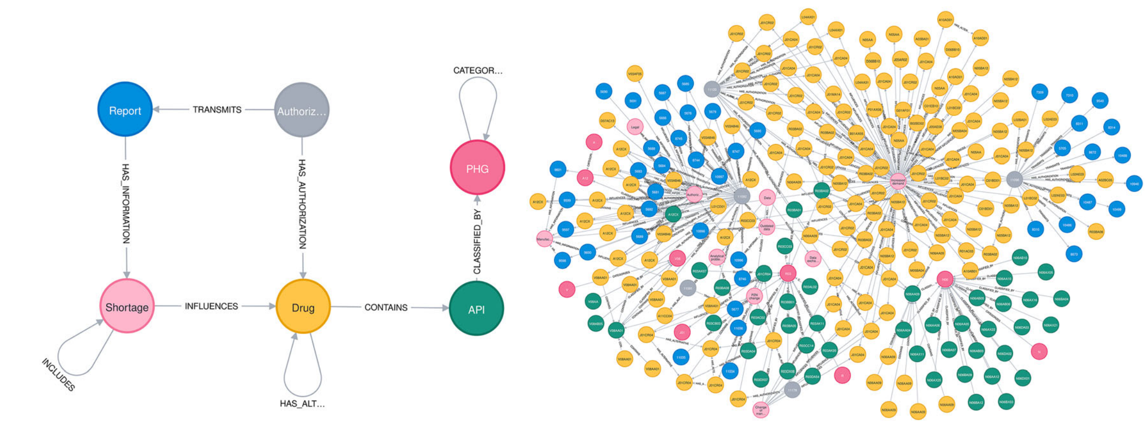
3.5. Knowledge analysis and application

The final step involves analyzing the constructed KG and applying the knowledge gained from the retrieved data. This process leverages the interconnected data within the KG to extract relevant information. By employing graph metrics and querying the graph, decision-makers can derive valuable insights that enhance understanding and inform strategic actions.

In our study, we employ diverse methods to extract valuable information from the graph to uncover patterns,

Table 3. Relationships between entities in the MedSupplyKG.

Relationship	Connecting nodes	Description
: TRANSMITS	'Authorisation holder' ⟶ 'Report'	Notification of an occurring drug shortage
: HAS_AUTHORIZATION	'Authorisation holder' ⟶ 'Drug'	Indicating the responsible license holder of the drug
: HAS_INFORMATION	'Report' ⟶ 'Shortage'	Detailing information on the shortage
: INCLUDES	'Shortage' ⟶ 'Shortage'	Specifying the shortage type
: INFLUENCES	'Shortage' ⟶ 'Drug'	Factors impacting drug availability
: HAS_ALTERNATIVE	'Drug' ⟶ 'Drug'	Referring to alternative medication
: CONTAINS	'Drug' ⟶ 'API'	Specifying the API in each drug
: CLASSIFIED_BY	'API' ⟶ 'PHG'	Indicating which APIs target which medical issues
: CATEGORIZES	'PHG' ⟶ 'PHG'	Specifying the relation between main and sub-PHG groups



(a) (b)

Figure 4. Graph schema (a) and excerpt (b) of the resulting MedSupplyKG using the graph database Neo4j.

identify critical nodes, analyse relationships, and derive insights that are not immediately apparent. These methods include descriptive approaches, graph analysis, and queries. Specifically, graph analysis encompasses a set of techniques used to understand the entities within a network and to analyse the structures that emerge from the relationships among these entities (Chiesi 2001). Centrality measures are commonly applied to explore a network locally and evaluate vertices' importance. The degree centrality $C_D(i)$ of a node i represents the number of nodes to which the focal node j is connected, indicating the node's level of involvement in the network (Opsahl, Agneessens, and Skvoretz 2010). It is quantified as the sum of the edges connected to the node and is calculated using Equation (1). In this equation, N is the total number of nodes, and a is the adjacency matrix in which the cell a_{ij} is defined as 1 if node i is connected to node j, and 0 otherwise.

$$C_D(i) = \sum_{j=1}^{N} a_{ij} \tag{1}$$

In addressing drug shortages, we utilise the MedSupplyKG to identify bottlenecks, analyze critical relationships, and derive insights. Specifically, we apply the centrality measure to pinpoint influential entities in the supply chain, as detailed in Section 4.1. Identifying these nodes is crucial because nodes with high centrality are often critical points in the network, and issues such as production delays or quality problems can cause significant disruptions. For example, a lack of raw materials at a central authorisation holder can lead to shortages of multiple drugs. The centrality measure also helps assess supply chain vulnerabilities by highlighting potential points of failure. Understanding the most central nodes enables decision-makers to develop targeted contingency plans.

Furthermore, we utilise the query language *Cypher*. Cypher is a language designed explicitly for querying graph data in Neo4j. It enables data retrieval and helps to identify specific sub-graphs, complex relationships, and structures, making it a crucial tool for exploring and analysing graph-based information (Neo4j 2024b). In the context of this research, we employ Cypher to identify and analyze sub-graphs within the larger MedSupplyKG. For instance, we can query the drivers of drug shortages by examining the relationships between shortage nodes and drug nodes. Additionally, we can trace which shortages affect specific drugs, identify the corresponding authorisation holders, and classify these drugs by their pharmacological groups, ultimately impacting different patient groups. By querying the graph, we can isolate the relevant nodes and edges, enabling a detailed

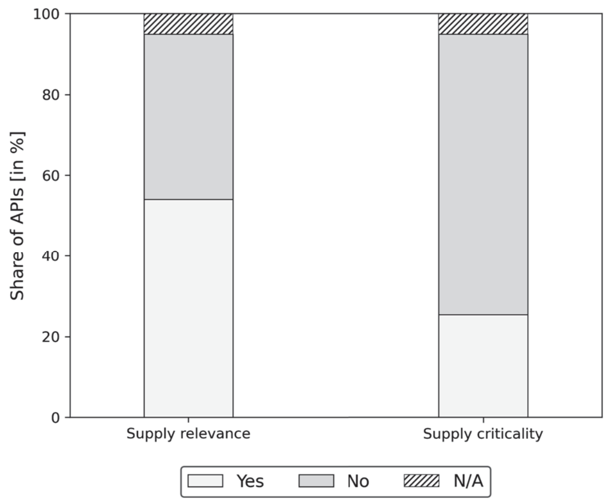


Figure 5. Distribution of APIs according to supply relevance and criticality. (Total number of APIs: 687)

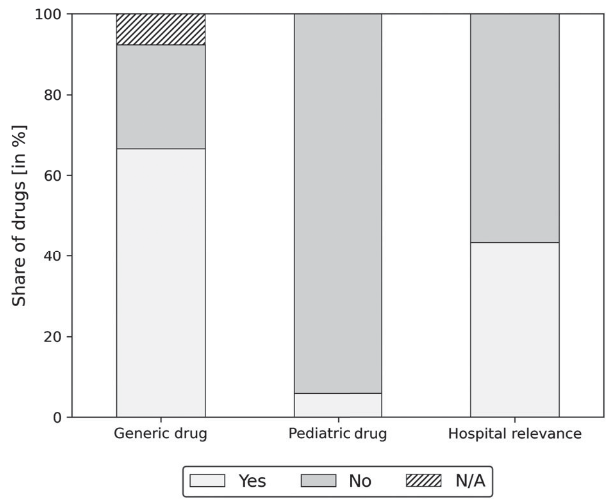


Figure 6. Distribution of drugs according to drug characteristics. (Total number of drugs: 3869)

examination of that supply chain segment. This approach allows decision-makers to trace dependencies, analyze the distribution of drug shortages, identify historical developments, determine the key drivers of these shortages, and assess the impact of disruptions. The results of these analyses are described in more detail in Section 4.

4. Results

4.1. Characteristics of the MedSupplyKG

The implemented KG consists of 10,351 nodes representing the entities: reports, drugs, authorisation holders, APIs, shortages, and PHGs connected by 24,820 edges denoting various relationships. The graph comprises 5414 report nodes detailing information on 3869 drugs licensed by 251 authorisation holders. These drugs comprise 687 distinct APIs, categorised into 90 PHGs, including 14 main groups and 76 therapeutic subgroups. The shortage nodes provide information on 32 shortage reasons categorised into 8 distinct reason types.

4.1.1. Distribution of drug and API data

Figures 5 and 6 present data on the distribution of APIs and drugs focussing on the relevance and criticality of APIs and the characteristics of drugs, being generic, paediatric, or hospital-related.

Figure 5 illustrates that 54.15% ($n = 372$) of the APIs are classified as supply-relevant, indicating that the WHO lists them as essential medicines and their unavailability would negatively impact patient outcomes (BfArM 2024f). Additionally, 25.62% ($n = 176$) of the APIs are deemed supply-critical. Supply-critical APIs are considered supply-relevant, have three or fewer authorisation holders or manufacturers, have previously experienced supply shortages, and are included in the substitution exclusion list (BfArM 2024f). Given their importance, the fact that 25.62% of the APIs fall into this category represents a significant portion.

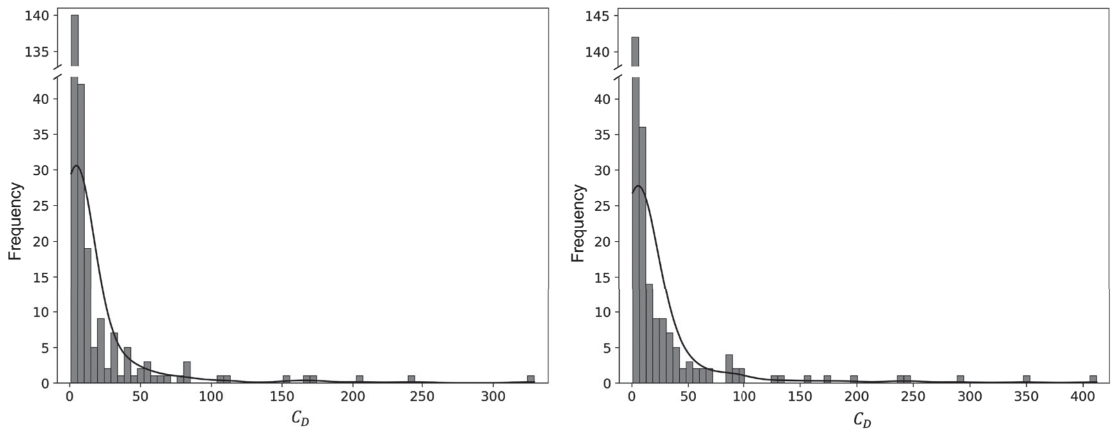
Furthermore, Figure 6 provides information on specific characteristics of the medicinal products. The majority of these products, with 66.63% ($n = 2578$), are generics, which are non-patented drugs that are typically more affordable. A small proportion of the drugs, namely 5.92% ($n = 229$), are paediatric medications, highlighting the therapeutic needs of children. Notably, 43.40% ($n = 1679$) of the medications are classified as hospital-relevant, underscoring their essential role in hospital treatments and the potential impact on patient care if shortages occur.

4.1.2. Degree distribution of authorisation holders

Within the graph scheme, authorisation holders play an essential role by licensing certain drugs and reporting shortages. To assess their engagement and impact on the pharmaceutical supply chain, we analyse the degree distribution of authorisation holders in the MedSupplyKG. The degree of a node in this context represents the number of direct connections an authorisation holder has with other entities in the graph.

Specifically, we examine connections between 'Drug' and 'Report' nodes associated with each 'Authorisation Holder'.

Figure 7(a) illustrates the degree distribution among authorisation holders regarding drug associations, whereas Figure 7(b) displays the distribution of submitted reports. The right-skewed distribution of both metrics implies that most authorisation holders have relatively low degree values, thus a limited number of



(a) Drug Degree. (b) Report Degree.

Figure 7. Degree distribution of authorisation holders with C_D as centrality degree. (a) Drug Degree and (b) Report Degree.

Table 4. Top 10 authorisation holders and their degrees.

Authorisation Holder	Drug Degree	Report Degree	Total Degree
Hexal Aktiengesellschaft	329	412	741
1 A Pharma GmbH	242	290	532
Glenmark Arzneimittel GmbH	168	348	516
Ratiopharm GmbH	204	247	451
ALIUD PHARMA GmbH	172	237	409
Aristo Pharma GmbH	152	158	310
Mibe GmbH Arzneimittel	108	195	303
Mundipharma GmbH	111	177	288
Sanofi-Aventis Deutschland GmbH	83	130	213
Aspen Pharma Trading Limited	60	135	195

connections to drugs and reports. However, moving to higher degree values, the gradual decline of the kernel density estimation (KDE) curve highlights several authorisation holders with high degree values compared to the majority, indicating the existence of a few dominant players in this field.

Table 4 complements the degree distribution plots by ranking the top ten authorisation holders by their total degree values. *Hexal Aktiengesellschaft* emerges as the most connected authorisation holder, with a total degree of 741. The company holds licenses for 329 drugs in the dataset, of which they reported 412 shortages, indicating substantial involvement in the pharmaceutical domain. This is followed by *1 A Pharma GmbH* and *Glenmark Arzneimittel GmbH* with total degree values of 532 and 516, respectively. If the number of reported shortages exceeds the number of licensed medicinal products, it is reasonable to deduce that a license holder reports multiple bottlenecks for a single product over time.

4.2. Distribution of drug shortages by pharmacological groups

Figure 8 shows the distribution of drug shortages across the PHGs. The results indicate that drugs within the nervous system (PHG N) are the most affected, with a total of 1, 614 reported shortages.

Within this category, the subgroups N02 (Analgesics) and N03 (Antiepileptics) are particularly impacted. The subgroup N02, covering general analgesics and antipyretics medications, contains 490 reported shortages. Among these, two supply critical opioids, hydromorphone (N02AA03) with 107 reports and oxycodone (N02AA05) with 97 reports, hold the highest share of notified shortages. Further 307 shortages are reported within the group N03, impacting medication availability for treating epilepsy. Particularly affected are medications that contain the supply-relevant APIs lamotrigine (N03AX09) with 50 reports and levetiracetam (N03AX14) with 49 reports.

Drugs treated as anti-infective for systemic use (PHG J) and the cardiovascular system (PHG C) also exhibit high numbers of total shortages, with 1001 cases reported in group J and 700 reported shortages in group C. In group J, the most significant number of shortages (721 reports) is attributed to APIs classified as antibacterials for systemic use (subgroup J01). This includes 70 reports of the API amoxicillin (J01CA04), a penicillin derivative used to treat infections caused by gram-positive bacteria (DrugBank 2024). Within medications

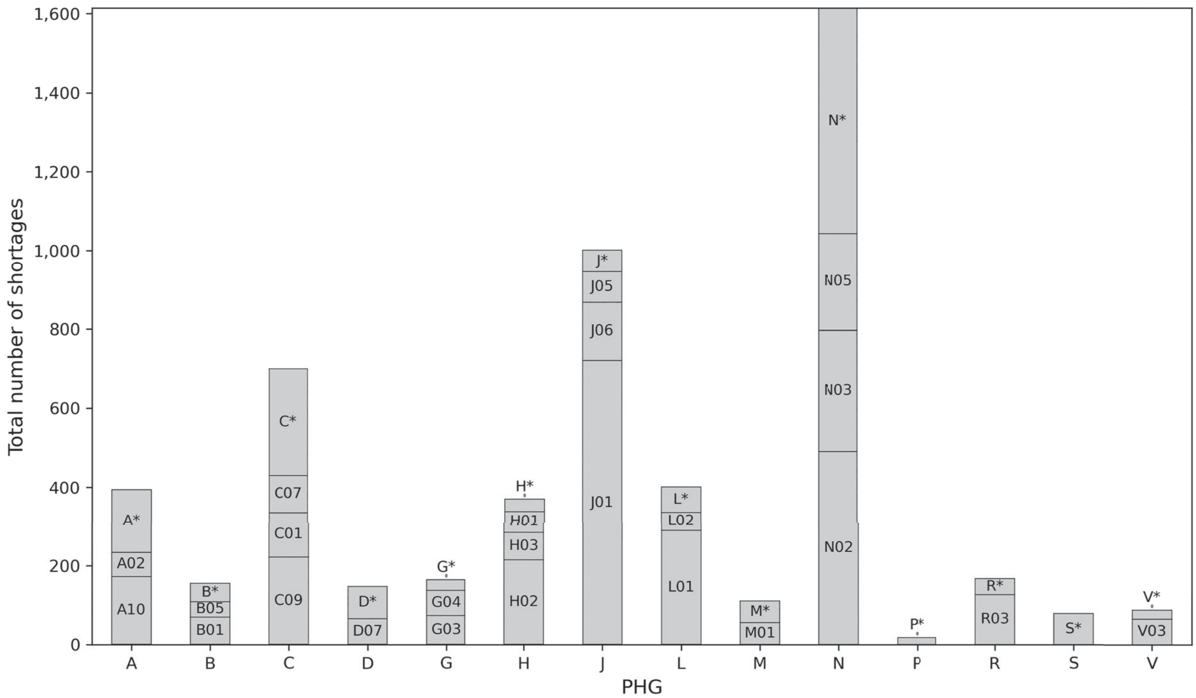


Figure 8. Distribution of drug shortages according to PHG.

used for the cardiovascular system in group C, APIs used as agents acting on the renin-angiotensin system in subgroup C09 are most affected. These drugs primarily treat hypertension and heart failure. Among the 222 reported shortages, 72 reports pertain to the API valsartan (C09CA03) and 68 to the API valsartan and diuretics (C09DA03), medication to control and treat high blood pressure.

Further notable findings include supply shortages in oncological medications (PHG L), alimentary tract and metabolism drugs (PHG A), and systemic hormonal preparations (PHG H). In particular, the subgroup L01 (290 shortages) referring to antineoplastic agents applied for cancer treatment highlights a gap in available medication in oncology care. This includes shortages of medications comprising the APIs methotrexate (L01BA01) and mitomycin (L01DC03), both supply-relevant APIs. Drugs in PHG A, particularly subgroups A10 (diabetes medications) and A02 (stomach acid treatments), received 172 and 61 shortage reports, respectively. PHG H recorded a total of 370 shortage reports, with the most significant shares in H02 (corticosteroids for inflammatory and autoimmune conditions) and H03 (levothyroxine for hypothyroidism).

4.3. Annual development of drug shortages

Figure 9 depicts the annual distribution of reported drug shortages, both in total and categorised by PHG, according to the transmitted start dates. The dataset tracks each PZN separately, ensuring that medications are uniquely identified. The bar chart shows the total number of quarterly shortage reports, while the line plots indicate the reports for each PHG over the same time frame. The figure reveals a noticeable increase in shortage reports from 2017 through 2024. This upward trend indicates an intensifying impact of disruptions within pharmaceutical supply chains, with significant spikes in certain quarters, particularly in 2018, 2020, and 2023. Certain PHGs, such as ‘N’, ‘J’, ‘C’, ‘L’, ‘A’, ‘H’, and ‘R’ show higher variability and peaks. In contrast, other groups like ‘B’, ‘D’, ‘G’, ‘M’, ‘P’, ‘S’, and ‘V’ have consistently low or zero reports.

In 2018, a spike in reports occurred, reaching 136, primarily related to shortages of drugs containing valsartan (C09CA03), commonly used for blood pressure. BfArM (2024g) attributes this shortage to a change in synthesis methods by a Chinese manufacturer, leading to nitrosamine contamination. This contamination resulted in a production halt, approval suspension, and a recall of affected batches, causing significant supply disruptions. Following this, the COVID-19 pandemic played a crucial role in the surge observed in early 2020, as it strained healthcare systems worldwide (EMA 2023). According to Piatek, Ning, and Touchette (2020), the demand for antibiotics and antiviral medications (PHG J) used to treat bronchitis and pneumonia, as well as medications for pain and sedation (PHG N), increased in 2020 due to COVID-19 infections and export bans

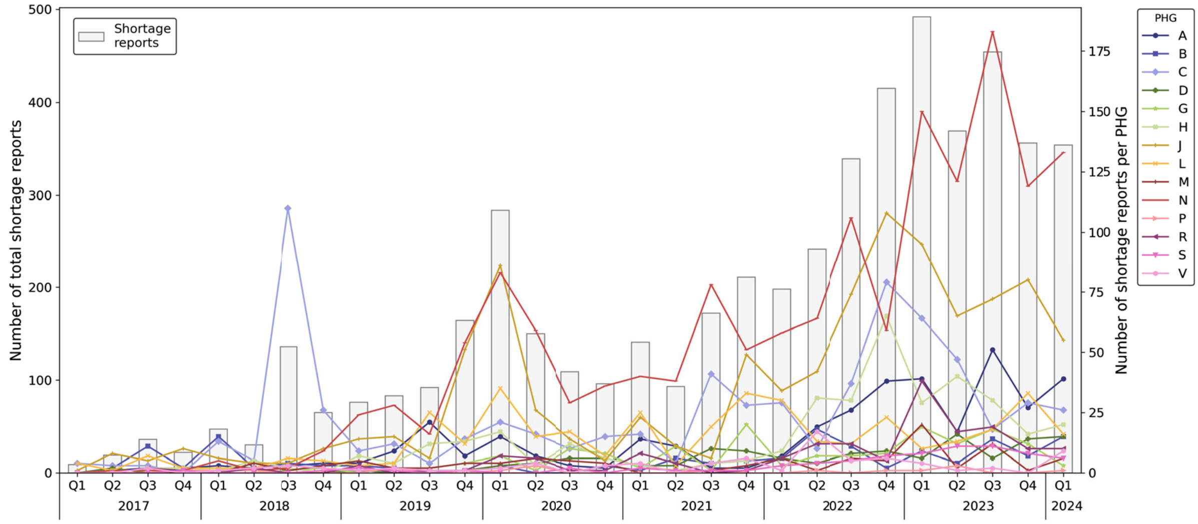


Figure 9. Distribution of the reported drug shortages per year and PHG.

of APIs from India. These findings of Piatek, Ning, and Touchette (2020) align with our analysis. According to our dataset, medication shortages during 2020 included, for example, clindamycin (J01FF01), trimethoprim (J01EA01), buprenorphine (N02AE01), morphine (N02AA01), and midazolam (N05CD08). Furthermore, medications of PHG L, such as bleomycin (L01DC01) used for cancer treatment, were affected.

Despite these short-term spikes, Germany continues to experience persistent drug shortages. In 2023, there were 1671 shortage reports, indicating a further increase. Antibiotics, antiviral medications (PHG J), and medications for the nervous system (PHG N), such as pain relievers, were once again in short supply. Additionally, drugs for the cardiovascular system (PHG C), alimentary tract and metabolism (PHG A), and systemic hormonal preparations (PHG H) used to treat diabetes also faced shortages. Multiple factors contribute to this development, such as increased demand and production problems, which are discussed in the following section.

4.4. Drivers of drug shortages

Figure 10 illustrates the reason types and subtypes extracted and defined based on the final dataset. This visualisation offers insights into the underlying drivers of drug shortages in Germany by highlighting the main reasons based on the degree centrality of 'Shortage' nodes.

In this context, the degree centrality of a 'Shortage' node (C_D) measures the importance of the nodes by counting the number of connections it has to 'Report' nodes in the network. Thus, it reflects how frequently a particular shortage reason appears within the reporting system, which can provide valuable insights into the severity and scope of the impacts.

Production, market, process, quality, legal, logistics, and data-related issues were the main reasons contributing to the drug shortages. *Production reasons* are the most commonly reported issues (1847 reports), particularly capacity bottlenecks ($C_D = 1,024$) and manufacturing issues ($C_D = 90$). Furthermore, bottlenecks in ingredients or packaging materials hinder or restrict the production and packaging of certain products, often leading to delays along the production line or supply chain. Compounding these issues are problems associated with Contract Manufacturing Organisations (CMOs), especially where manufacturer issues lead to temporary disruptions.

The second-largest category is related to *Market reasons* (1425 reports), with certain market conditions, such as increased demand ($C_D = 1299$) causing drug shortages. This surge is often driven by competitor supply issues and bottlenecks, stock-outs, market exits of competitors, or the inability of alternative suppliers to meet demands. These factors lead to significant overselling, creating further capacity bottlenecks across the industry. Increased sales also strain the supply chain when demand spikes unexpectedly. In addition, some drugs are being withdrawn from the market altogether, exacerbating the challenges of maintaining a steady supply of medicines.

Process reasons (632 reports) comprise changes and adaptations in manufacturing or packaging. For instance, transitions between packaging types (e.g. containers to blisters) may delay product availability. These reasons also encompass final release issues and testing issues with delays or damages affecting the release of produced batches, further contributing to the availability gaps. Moreover, changes in suppliers and manufacturers often lead to production delays, increasing the likelihood and occurrence of shortages of certain products.

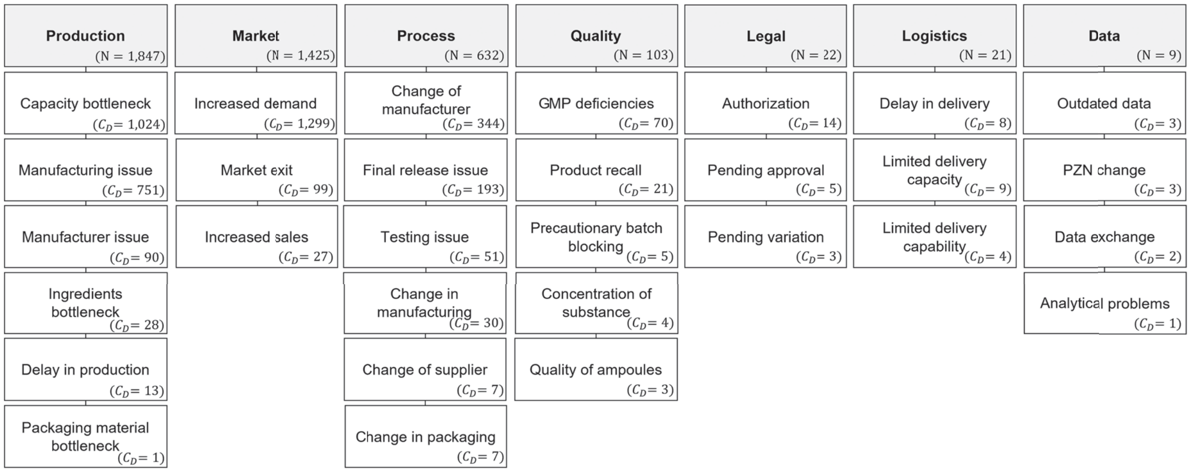


Figure 10. Drivers of drug shortages and their classification.

Quality reasons also contribute to these shortages (103 reports), mainly referring to failures or shortcomings in adhering to regulatory standards (GMP deficiencies, $C_D = 70$) that frequently lead to product recalls ($C_D = 21$). Additionally, unusually high concentrations of substances, such as excessive nitrosamine levels, can halt production. Precautionary batch blocking due to quality issues or defects in supplied ampoules further contributes to the interruptions in drug supply.

Moreover, several issues are related to *Legal reasons* (22 reports) that involve expired or withdrawn authorisation ($C_D = 14$) and pending approvals ($C_D = 5$), reflecting how regulatory processes can contribute to these shortages. Delays in obtaining or renewing licenses can lead to internal production capacity bottlenecks and release delays. Pending variations often result from contract manufacturer transfer, which may further slow down the scheduled production.

Logistics related issues (21 reports) encompass delivery delays, distribution changes, and limited capability or capacity. These challenges frequently arise from high utilisation rates at production facilities or distribution network changes, which result in decreased drug availability.

Lastly, *Data* related problems (9 reports), such as analytical problems, PZN changes, outdated data, and data exchange delays, contribute to inefficiencies in managing pharmaceutical information.

4.5. Drivers of drug shortages by pharmacological groups

The drivers of drug shortages can also be examined within individual PHGs, as depicted in Figure 11. The distribution indicates that the shortages are primarily related to production, market, and process issues. Additionally, specific reasons are particularly relevant for certain groups. Furthermore, the previous analysis of Figures 8 and 9 shows that PHGs N, J, and C are particularly affected by shortages.

Upon examining the causes of shortages, it is evident that the supply of PHG N is primarily hindered by production and market factors, with additional impacts from process and quality issues. According to Figure 10, the predominant subtypes within these categories are capacity bottlenecks, change of manufacturer, increased demand, and GMP deficiencies. In contrast, PHG J faces especially market challenges such as increased demand and disruptions in production, process, quality, and logistics. Similarly, PHG C encounters these same issues but is more affected by quality problems. Based on these findings, implementing targeted measures for these groups is beneficial in efficiently mitigating or preventing drug shortages.

4.6. Impact evaluation of drug shortages

We further analyse the impact of drug shortages using specific parameters derived from the MedSupplyKG. Our analysis quantifies a drug shortage's severity and economic impact, thereby supporting prioritisation measures for decision-makers. The shortage severity factor S_f is determined by a severity weight ω, which is derived from the substance's supply relevance and criticality, multiplied by the reported number of shortages n_s for the substance:

$$S_f = \omega \cdot n_s \tag{2}$$

The severity weight ω is calculated by assigning binary scores to the attributes: one if the substance is supply-relevant or critical and zero otherwise. The severity weight is then computed as a default value of one plus the

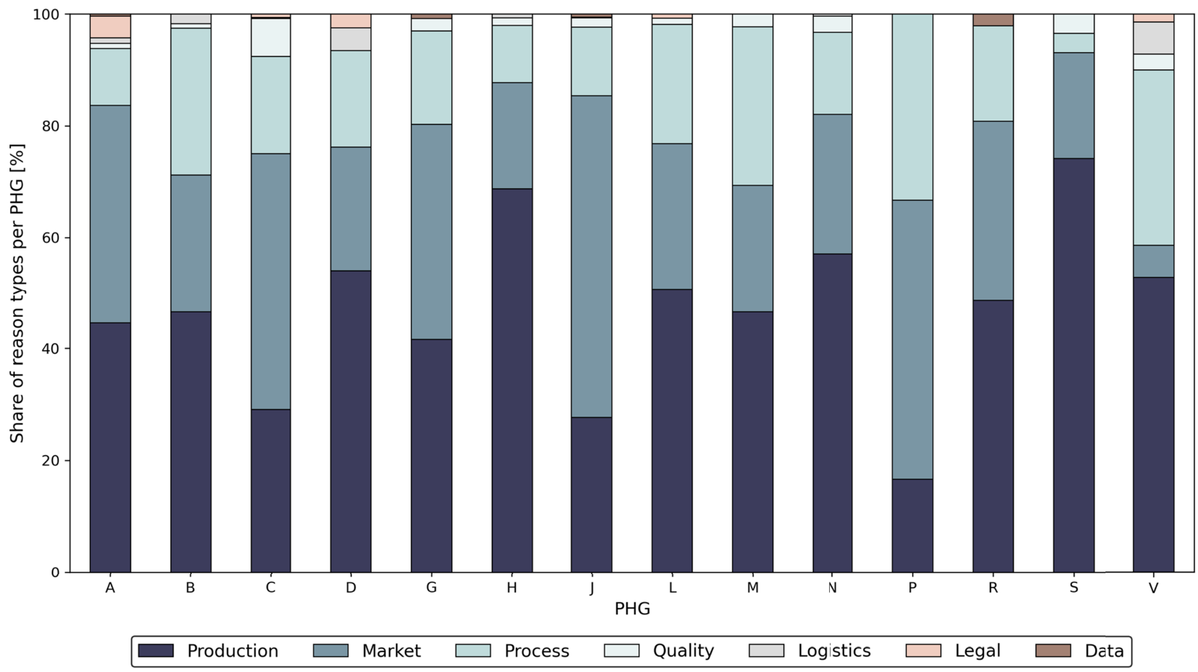


Figure 11. Distribution of shortage reasons according to PHG.

sum of these binary scores, resulting in a value ranging from one to three, depending on the values of the assessed attributes. After that, the shortage severity is normalised between zero and one using min-max normalisation.

The economic impact factor is determined by the ratio of C_{net}, which denotes the net cost of prescriptions, and N_{pres}, the total number of prescriptions for the relevant substance:

$$E_f = \left(\frac{C_{\text{net}}}{N_{\text{pres}}}\right) \tag{3}$$

This ratio is also normalised to enhance comparability. The combination of the factors S_f and E_f results in an impact matrix with four quadrants, as depicted in Figure 12.

The lower left quadrant indicates low shortage severity and minor economic impact, while the lower right quadrant comprises shortages with moderate to severe economic impact but low severity. Similarly, the upper left quadrant indicates high severity but low economic impact. The upper right quadrant is the most severe, indicating critical shortages with high severity and economic impact. The substances are integrated into the matrix based on the calculated scores and colour-coded to represent the impact levels of the combined factors. The most significant substances in this regard are annotated with their associated ATC Code and marked by dotted circles.

Multiple shortages exhibit low severity and minimal economic impact, primarily because they involve widely used and well-established medications. Examples include eye drops for allergic diseases (R06AX17), medicines for treating high blood pressure (C09AA04), or analgesics commonly used for pain relief and fever reduction, such as paracetamol (N02BE01). However, specific shortages are characterised by a higher economic impact, greater shortage severity, or both, making them particularly significant for companies and public decision-makers.

For instance, the substances B01AF01, N02BB02, A02BC02, and C09CA06 are especially relevant regarding the economic impact. They belong to the PHGs B, N, A, and C, which include medications for blood and blood-forming organs, the nervous system, the alimentary tract and metabolism, and the cardiovascular system. The significant economic impact arises from a high volume of prescriptions, their associated net costs, or both. Contributing factors to these elevated costs include complex manufacturing processes, prolonged use in chronic disease management, and the availability of patented versions since B01AF01, N02BB02, and A02BC02 are non-generic drugs.

Looking at the shortage severity, specific substances of the PHGs H, J, and N exhibit a high relevance and criticality comprising systemic hormonal preparations, anti-infectives for systemic use, and the nervous system. These medications are vital for treating common yet significant health conditions. For instance, morphine (N02AA01) is critical for pain management, while antibiotics like amoxicillin (J01CR02) and phenoxymethylpenicillin (J01CE02) are essential for treating bacterial infections. Prednisolon (H02AB06) is vital

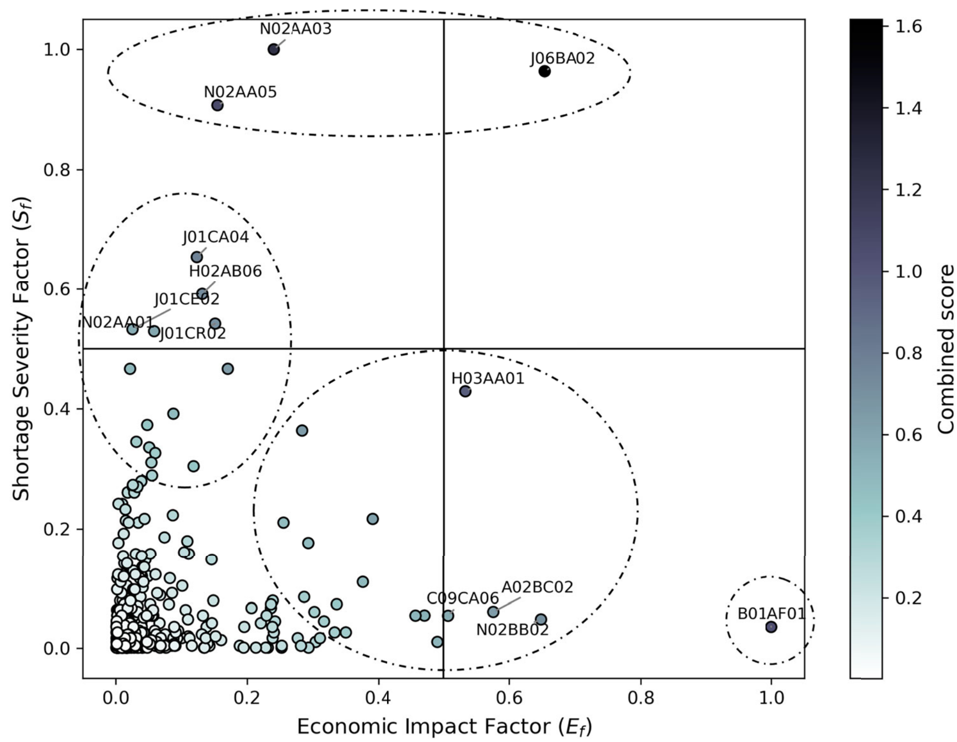


Figure 12. Evaluation of economic impact and shortage severity.

in managing inflammatory and autoimmune conditions. While they are relevant and critical for treatment, many of these medications, particularly antibiotics and pain relievers, are available as generics, significantly lowering their market price.

Substances with significant impact in both areas of the MedSupplyKG include H03AA01 (Levothyroxine sodium), used for treating thyroid deficiency states, and J06BA02 (Human Immunoglobulin), which supports the immune system against certain diseases. These medications are essential for a large number of patients, especially H03AA01. Additionally, J06BA02 is quite costly in Germany, contributing to its high impact in both dimensions.

5. Discussion

5.1. Findings

In this study, we address the need of the pharmaceutical industry to develop robust strategies to manage the growing challenges of drug shortages. These shortages, driven by supply chain disruptions, regulatory issues, or demand spikes, require innovative solutions to ensure a reliable supply of essential medicines (Shukar et al. 2021). Therefore, we apply a KG-based approach to investigate the underlying drivers of drug shortages, their characteristics, and their impacts on the pharmaceutical market in Germany. To construct the graph, we follow a systematic, multi-stage approach including data collection and processing of disparate and heterogeneous data sources. This integration facilitates the mapping of dependencies and interactions relevant to the identification and causal understanding of the factors contributing to drug shortages.

The results reveal several key findings underscoring the potential of KGs in improved management and decision-making. In this regard, the visualisation and query functionalities facilitate determining the root causes behind shortages. These insights show that drug shortages in Germany have steadily risen, with a 161.44% increase between 2020 and 2023. The reasons for this development are multifaceted, mainly concerning production bottlenecks, increased demand, quality issues, and supply delays. These challenges are also linked to the dependence on foreign markets as the production of generics, which account for the largest share, is predominantly concentrated in Asia (Palit and Bhogal 2022). Drawing on the implemented interconnections between entities, the KG enables the uncovering of critical supplier patterns, particularly those frequently reporting shortages. By analysing the degree centrality of authorisation holders, we identify the most influential suppliers in the network and highlight key players whose disruptions could significantly impact the whole drug supply chain.

Furthermore, the analysis allows us to find shortage variations between different PHGs, highlighting which groups are most susceptible to supply issues. Notably, many shortages are constantly reported for medications for treating nervous system disorders, anti-infectives for systemic use, and cardiovascular medications. These shortages affect supply-critical and supply-relevant APIs that are relevant for the production of various drugs. Understanding these disparities is crucial for securing the long-term supply of drugs that are critical to the public health system. Additionally, by quantifying the economic impact and shortage severity factor, the KG enables the evaluation of the drugs' relevance to the market. The developed matrix serves as a foundation for informed decision-making through priority assessment. For companies, insights into disruptions of drug supply are pivotal in shaping production strategies, pricing models, and supply chain planning, ultimately influencing profitability. This information is also crucial for public decision-makers ensuring access to essential medications, managing budgets, and formulating effective policies.

In summary, our KG reveals the intricate interconnections between multiple factors, the duration of shortages, and specific product characteristics, such as supply relevance and generic status. It also highlights the role of stakeholders and the economic impact these shortages have on the market and various patient groups. From a broader perspective, our results align with those of Bade et al. (2023) and Francas, Mohr, and Hoberg (2023), based on the BfArM dataset like our study, making them particularly interesting for comparison.

Bade et al. (2023) identify supply issues, manufacturing problems, logistics disruptions, product recalls, and discontinuations as the primary drivers of drug shortages. In comparison, our analysis highlights slightly different factors, including production, market, process, quality, legal, logistics, and data-related issues, with production and market-related causes being the most frequently reported. Interestingly, our analysis of the BfArM dataset revealed greater detail regarding the causes of shortages, suggesting some improvements in the reporting system. However, gaps remain in the depth of reporting. While the BfArM data identify the causes of drug shortages, they do not fully illustrate their interconnections, highlighting a lack of transparency and communication across the supply chain and among stakeholders. To address this gap, we enriched the dataset within our KG, enabling a more comprehensive understanding of the factors driving shortages.

In contrast, Bade et al. (2023) derive additional drivers through extensive literature research and interviews with authorisation holders, while our approach establishes links between the underlying causes of shortages, authorisation holders, product characteristics, and market data.

Francas, Mohr, and Hoberg (2023), using logistic and negative binomial regression models, analyse the impact of market, drug substance, and regulatory characteristics on the likelihood of a drug shortage. Similar to our approach, they enrich the BfArM data with additional information such as market data, drug characteristics, and classifications to identify relationships between variables. For example, they find that drug characteristics, such as still-patented and biopharmaceutical drugs, are less likely to be associated with shortages and state that substances classified as essential medicines have a higher probability of shortage. We observe similar trends concerning generic drugs, their criticality, and supply relevance. However, these findings should be interpreted cautiously due to methodological differences and the specific reporting obligations in Germany.

In comparison to Bade et al. (2023) and Francas, Mohr, and Hoberg (2023) as well as other works, our approach aims to establish a backbone for connecting heterogeneous datasets, enabling a more dynamic and comprehensive analysis. By visually mapping the complex relationships between various entities, our method reveals interdependencies and patterns that might remain hidden in traditional data analysis or qualitative research methods.

5.2. Managerial implications

Managerial implications are derived for companies and public authorities, such as BfArM, which oversees monitoring supply shortages for human medicines in Germany. Given the vulnerability of pharmaceutical supply chains and the severe impact disruptions can have on this critical sector, designing resilient supply chains is a significant challenge for companies and public stakeholders. In this context, developing a scalable methodology for constructing a comprehensive KG from diverse, publicly available data sources enables decision-makers to analyse large and heterogeneous data sets. Therefore, we integrated the primary dataset of drug shortages with market data, product information, and API details to provide managers with a novel perspective on the network, offering a unified view for knowledge generation, retrieval, visualisation, analysis, and risk identification.

For example, the KG links shortage reports to authorisation holders and their medications, integrating BfArM classifications to indicate whether a drug is considered supply-relevant or supply-critical. This approach allows authorities to assess the overall supply situation and identify companies essential for maintaining supply stability. In this context, graph metrics highlight relationships between authorisation holders and other entities, supporting the pinpointing of key players within the network. As a result, 'Hexal Aktiengesellschaft' emerges as the most connected authorisation holder, reflecting its significant role in the pharmaceutical sector.

Additionally, integrating supplementary datasets that classify drugs based on attributes such as whether they are generic, paediatric, or intended for hospital use provides deeper insights into shortages. For instance, this approach helps uncover the factors contributing to paediatric drug shortages and supports the development of targeted strategies to address shortages in hospital medications. Managerial insights can also be gained by analysing historical shortage trends and associating them with the respective drugs, their authorisation holders, and the relevant PHG. This analysis facilitates the identification of past peak occurrences and their underlying causes. For instance, the study reveals that certain groups, such as drugs for the nervous system, anti-infectives for systemic use, or medications for the cardiovascular system, show more significant variability and frequent peaks. These groups are also linked to each drug's primary causes of shortages.

Furthermore, the graph connects drug shortages with their corresponding pharmacological groups, prescription volumes, and cost per prescription to highlight the populations affected by each shortage. These established connections enable companies to evaluate the economic impact of a shortage and, through a more profound analysis of the graph, identify the root causes of financial losses.

Addressing the identified key drivers of shortages, such as production, market, process, quality, legal, logistics, and data issues, involves measures like identifying and improving the monitoring of critical drugs with limited alternatives and a higher risk of shortages. This approach assists authorities in focussing their efforts and taking quick action to ensure a stable supply of essential medications for the population. Mandatory stockpiling is another strategy to reduce the impact of production or logistics disruptions. The reserve can be established centrally for Germany and Europe or decentralised at medical facilities. For example, at the end of 2019, the Netherlands created a strategic reserve of medicines to prevent supply shortages, including stockpiles at companies and wholesalers (Medicines For Europe 2024). In this context, implementing targeted preparedness contracts with relevant companies can ensure the availability of critical medicines during shortage situations (Schwarz 2021).

Further measures aim to strengthen individual supply chains and secure the procuring of critical raw materials through increased diversification (Spieske et al. 2022). Kalaitzi and Tsolakis (2023) discuss strategies such as dual sourcing, engaging multiple suppliers, and reshoring the production of critical supplies to Europe. Consequently, these strategies mitigate dependency on single sources, shorten supply chains, and avoid risks due to export stops or shutdowns of plants in distant markets. Therefore, public authorities must focus on enhancing Europe's attractiveness as a production location by accelerating technological advancements, digitalisation, and streamlining approval processes (European Commission 2020). In addition, political regulations are necessary to support mitigation measures, including specific benefits such as tax reductions or subventions to offset the higher production costs in Europe (Francas, Fritsch, and Kirchhoff 2022).

However, these approaches require transparency and robust data analysis systems between various stakeholders to detect shortages and ensure a continuous supply of essential pharmaceuticals. KGs encourage smoother collaboration and more efficient information sharing within organisations and among various stakeholders. As a result, investments in digital processes, big data, and AI are necessary, as announced by several companies and federal authorities, including the BfArM (Francas, Fritsch, and Kirchhoff 2022). Therefore, the BfArM reporting system and its usage would benefit from restructuring to enhance reporting detail and information exchange between stakeholders. In this regard, our approach provides a robust foundation for decision-makers by enabling the integration and interconnection of diverse datasets. Furthermore, it serves as a structural framework for AI-driven applications.

5.3. Limitations and future research

Despite the contributions of our study, we acknowledge several limitations that merit further investigation and create opportunities for future research. The developed KG relies on limited public data sources, which lack interesting supply chain information and detailed market data. This limitation prevents in-depth analysis of specific disruptions and their impacts, including detailed effects on the population and patients. In addition, the static structure of the graph fails to capture the dynamic and evolving nature of supply chain networks. As a result, its reliability hinges on the credibility of the underlying sources, and despite our efforts to refine the data, outdated or inaccurate information may still lead to errors. Furthermore, the construction process of the KG requires manual effort, making it labour-intensive and costly.

In this regard, future research should explore opportunities to integrate additional or more granular data, such as production locations, suppliers, transportation routes, inventory levels, sales data, alternative medication options, and behaviour relating context into the provided KG structure. Incorporating such data, for example, through targeted studies, investments in data sets, surveys, or interviews, will significantly enhance the graph's capacity, revealing deeper connections and providing more nuanced insights. Therefore, new reporting standards and collaboration forms might be valuable for collecting data from different stakeholders and sources to build a resilient and efficient network.

Given that many knowledge graphs suffer from incompleteness, synthetic data generated through simulations or machine learning techniques such as link prediction and graph completion could serve as valuable supplements to real-world datasets. Moreover, dynamic knowledge graphs combined with learning algorithms that capture temporal dynamics could overcome the static nature and update the graph in real time as new information becomes available. In addition, further structural analysis techniques, such as those employed in SNA, provide a more nuanced understanding of the relationships and dynamics within knowledge graphs. These techniques and metrics can offer deeper insights into the interconnectedness and influence of various entities within the graph.

Moreover, the challenges associated with manual construction can be addressed by integrating the structural advantages of KGs with natural language processing techniques, particularly LLMs. These methods automate data mapping from unstructured sources into a pre-existing graph structure. LLMs further enhance decision-support capabilities by explaining the connections between entities and the significance of their relationships, making insights more accessible. This functionality enables decision-makers to interact with knowledge graphs through intuitive queries.

While we have demonstrated using KGs in the pharmaceutical industry, their adaptability extends to various sectors and supply chain tasks, such as product management and recommendations, predictive maintenance, digital twin modelling, SNA, or threat intelligence for mapping attack patterns. Future research can explore industry-specific applications, leveraging the structural framework of KGs to map and analyse complex relationships, identify vulnerabilities and sentiment trends, and uncover opportunities contributing to more resilient and efficient networks.

6. Conclusion

This study underscores the urgent need for the pharmaceutical industry to develop robust strategies to address the growing challenges of drug shortages and the broader demand for greater supply chain transparency. By leveraging a KG-based approach, we explore the underlying drivers, characteristics, and consequences of drug shortages in the pharmaceutical market, particularly in Germany. As a result, companies and public authorities should enhance data integration by adopting scalable methodologies for constructing KGs to improve supply chain management and address shortages. Measures such as prioritising critical drugs, implementing stockpiling strategies, diversifying suppliers, establishing targeted contracts, fostering greater collaboration, and investing in AI and digital processes can mitigate shortage risks and ensure a stable supply.

Our findings demonstrate that a KG-based approach is valuable for identifying systemic issues within supply chain networks, promoting transparency, and facilitating comprehensive data connections and visualisations. Moreover, the method provides a solid foundation for AI-driven applications that have the potential to significantly improve supply chain management and strategic decision-making across various industries.

Disclosure statement

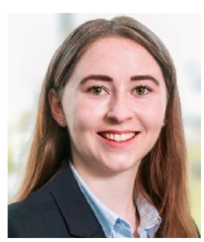
No potential conflict of interest was reported by the author(s).

Data availability

Data supporting the findings of this study are available on a reasonable request from the authors.

Notes on contributors

Katharina Eberhardt is a researcher at the Institute for Industrial Production (IIP) at the Karlsruhe Institute of Technology (KIT), Germany. She holds a M.Sc. in Industrial Engineering and Management from the KIT. Her research interests focus on enhancing supply chain resilience, risk management strategies, optimising logistics network design, and leveraging data analytics to support informed decision-making.



Amelie Schwärzel is a researcher at the Institute for Industrial Production (IIP) at the Karlsruhe Institute of Technology (KIT), Germany. She holds a M.Sc. in Industrial Engineering and Management from the KIT. Her research interests comprise strategic supply chain management, risk management strategies towards disruptions, and logistics network design.

Florian Klaus Kaiser holds B.Sc. and M.Sc. degrees in technically oriented business administration, focussing on information management from the University of Stuttgart, Germany, and a D.Eng. from Karlsruhe Institute of Technology (KIT), Germany. He was a senior researcher at the Institute of Information Security and Dependability, head of the Institute for Industrial Production research group, and lecturer at KIT. He is a senior security consultant and independent researcher. His main research interests are in data analytics and the engineering of secure systems.

Sonja Rosenberg is a postdoctoral researcher at the Institute for Industrial Production (IIP) at the Karlsruhe Institute of Technology (KIT), Germany, and is head of the Resilient Systems and Risk Management research group. She holds a M.Sc. in Management Science from the Catholic University Eichstätt-Ingolstadt and a doctorate in engineering (Dr.-Ing) from the KIT. Her research interests are developing methods and tools that allow decision support in uncertain and complex systems to provide resilient solutions.

Frank Schultmann is a Chair Professor of Business Administration, Production, and Operations Management at the Karlsruhe Institute of Technology (KIT), Germany, as well as a Professor of Complex Project Management at the University of Adelaide, Australia, and the Director of the Institute for Industrial Production, KIT, and the KIT's French-German Institute for Environmental Research (DFIU). He received the diploma, doctorate, and habilitation degrees in industrial engineering and management science from KIT (formerly the University of Karlsruhe) in 1993, 1998, and 2003, respectively. His research interests include sustainable manufacturing and logistics, decision support, supply chain management and optimisation, project management, technology assessment, construction management, and information and communication technologies.

References

Abu Zwaida, T., K. Elaroudi, and Y. Beauregard. 2022. "The Challenges of Drug Shortages in the Canadian Hospital Pharmacy Supply Chain—A Systematic Literature Review." *Journal of Public Health* 30 (11): 2593–2604. https://doi.org/10.1007/s10389-021-01485-2.

Akroyd, J., S. Mosbach, A. Bhave, and M. Kraft. 2021. "Universal Digital Twin—A Dynamic Knowledge Graph." *Data-Centric Engineering* 2:e14. https://doi.org/10.1017/dce.2021.10.

AlMahri, S., L. Xu, and A. Brintrup. 2024. "Enhancing Supply Chain Visibility with Knowledge Graphs and Large Language Models." https://arxiv.org/abs/2408.07705.

ApoKonzept24. 2023. "Deutsche Apothekengruppe – ApoKonzept24." Retrieved June 19, 2024, from https://www.apokonzept24.de/.

Azghandi, R., J. Griffin, and M. S. Jalali. 2018. "Minimization of Drug Shortages in Pharmaceutical Supply Chains: A Simulation-Based Analysis of Drug Recall Patterns and Inventory Policies." *Complexity* 2018:1–14. https://doi.org/10.1155/cplx.v2018.1.

Bade, C., A. Olsacher, P. Boehme, H. Truebel, and L. Fehring. 2023. "Reasons for Supply Side Driven Drug Shortages—A Mixed-Methods Study on First-Level, Higher-Level, and Root Causes from the Perspective of Marketing Authorization Holders." *Research in Social and Administrative Pharmacy* 19 (8): 1218–1227. https://doi.org/10.1016/j.sapharm.2023.05.007.

Badejo, O., and M. Ierapetritou. 2024. "Enhancing Pharmaceutical Supply Chain Resilience: A Multi-Objective Study with Disruption Management." *Computers & Chemical Engineering* 188:108769. https://doi.org/10.1016/j.compchemeng.2024.108769.

Berners-Lee, T., J. Hendler, and O. Lassila. 2001. "The Semantic Web." *Scientific American* 284 (5): 34–43. https://doi.org/10.1038/scientificamerican0501-34.

BfArM. 2024a. "BfArM – Lieferengpässe." Retrieved July 22, 2024, from https://www.bfarm.de/DE/Arzneimittel/Arzneimittelinformationen/Lieferengpaesse/_artikel.html.

BfArM. 2024b. "Lieferengpassmeldungen." Retrieved June 19, 2024, from https://anwendungen.pharmnet-bund.de/lieferengpassmeldungen/faces/public/meldungen.xhtml?jfwid=AA097A4A702660FE88E069CBCBE9186B%3A0.

BfArM. 2024c. "Liste der Kinderarzneimittel gem. § 35 Abs. 5a SGB V." Retrieved June 19, 2024, from https://www.bfarm.de/SharedDocs/Downloads/DE/Arzneimittel/Zulassung/amInformationen/Lieferengpaesse/Liste_kinderarzneimittel_35absatz5a_sgb_v_pdf.pdf?__blob=publicationFile.

BfArM. 2024d. "Liste der versorgungskritischen wirkstoffe nach § 52b abs. 3c amg (BfArM, Ed.)." Retrieved October 15, 2024, from https://www.bfarm.de/SharedDocs/Downloads/DE/Arzneimittel/Zulassung/amInformationen/Lieferengpaesse/Liste_versorgungskritischer_wirkstoffe.html.

BfArM. 2024e. "Liste der versorgungsrelevanten wirkstoffe nach § 52b abs. 3c amg (BfArM, Ed.)." Retrieved October 15, 2024, from https://www.bfarm.de/SharedDocs/Downloads/DE/Arzneimittel/Zulassung/amInformationen/Lieferengpaesse/ListeVersorgungsrelevanteWirkstoffe.html.

BfArM. 2024f. "Reporting Obligations." Retrieved June 12, 2024, from https://www.bfarm.de/EN/Medicinal-products/Information-on-medicinal-products/Supply-shortages/Reportingobligations/_node.html.

BfArM. 2024g. "Valsartan (BfArM, Ed.)." Retrieved July 25, 2024, from https://www.bfarm.de/DE/Arzneimittel/Arzneimittelinformationen/Rapid-Alert-System/Valsartan/_node.html.

Bier, T., A. Lange, and C. H. Glock. 2020. "Methods for Mitigating Disruptions in Complex Supply Chain Structures: A Systematic Literature Review." *International Journal of Production Research* 58 (6): 1835–1856. https://doi.org/10.1080/00207543.2019.1687954.

Bogaert, P., T. Bochenek, A. Prokop, and A. Pilc. 2015. "A Qualitative Approach to a Better Understanding of the Problems Underlying Drug Shortages, as Viewed from Belgian, French and the European Union's Perspectives." *PloS One* 10 (5): e0125691. https://doi.org/10.1371/journal.pone.0125691.

Borgatti, S., and X Li. 2009. "On Social Network Analysis in a Supply Chain Context." *Journal of Supply Chain Management* 45:5–22. https://doi.org/10.1111/jscm.2009.45.issue-2.

Brockmann, N., E. Elson Kosasih, and A. Brintrup. 2022. "Supply Chain Link Prediction on Uncertain Knowledge Graph." *ACM SIGKDD Explorations Newsletter* 24 (2): 124–130. https://doi.org/10.1145/3575637.3575655.

Buchgeher, G., D. Gabauer, J. Martinez-Gil, and L. Ehrlinger. 2021. "Knowledge Graphs in Manufacturing and Production: A Systematic Literature Review." *IEEE Access* 9:55537–55554. https://doi.org/10.1109/ACCESS.2021.3070395.

Chaudhri, V. K., C. Baru, N. Chittar, X. L. Dong, M. Genesereth, J. Hendler, A. Kalyanpur, et al. 2022. "Knowledge Graphs: Introduction, History, and Perspectives." *AI Magazine* 43 (1): 17–29.

Chiesi, A. 2001. "Network Analysis." In *International Encyclopedia of the Social & Behavioral Sciences*, edited by Neil J. Smelser and Paul B. Baltes, 10499–10502. Amsterdam, New York: Pergamon. https://doi.org/10.1016/B0-08-043076-7/04211-X.

Cimiano, P., and H. Paulheim. 2017. "Knowledge Graph Refinement: A Survey of Approaches and Evaluation Methods." *Semantic Web* 8 (3): 489–508.

Colin-Oesterlé, N., and J Mélin. 2020. "Report on the Shortage of Medicines – How to Address an Emerging Problem (2020/2071(ini)) (Committee on the Environment, Public Health and Food Safety, Ed.)." Retrieved July 31, 2024, from https://www.europarl.europa.eu/doceo/document/A-9-2020-0142_EN.pdf.

Destatis. 2024. "Krankenversicherungsschutz." Retrieved October 15, 2024, from https://www.destatis.de/DE/Themen/Arbeit/Arbeitsmarkt/Qualitaet-Arbeit/Dimension-2/krankenversicherungsschutz.html.

de Vries, H., M. Jahre, K. Selviaridis, K. E. van Oorschot, and L. N. van Wassenhove. 2021. "Short of Drugs? Call upon Operations and Supply Chain Management." *International Journal of Operations & Production Management* 41 (10): 1569–1578. https://doi.org/10.1108/IJOPM-03-2021-0175.

Diaz, R., S. Kolachana, and R. Falcão Gomes. 2023. "A Simulation-Based Logistics Assessment Framework in Global Pharmaceutical Supply Chain Networks." *Journal of the Operational Research Society* 74 (5): 1242–1260. https://doi.org/10.1080/01605682.2022.2077661.

DKV. 2024. "Medikamente Preisvergleich." Retrieved June 19, 2024, from https://www.dkv.com/gesundheit-arzneimittel-preis-vergleich.html.

Dolgui, A., and D. Ivanov. 2021. "Ripple Effect and Supply Chain Disruption Management: New Trends and Research Directions." *International Journal of Production Research* 59 (1): 102–109. https://doi.org/10.1080/00207543.2021.1840148.

Dörpinghaus, J., S. Klante, M. Christian, C. Meigen, and C. Düing. 2022. "From Social Networks to Knowledge Graphs: A Plea for Interdisciplinary Approaches." *Social Sciences & Humanities Open* 6 (1): 100337. https://doi.org/10.1016/j.ssaho.2022.100337.

DrugBank. 2024. "Antibacterials for Systemic Use." Retrieved September 10, 2024, from https://go.drugbank.com/categories/DBCAT002130.

Düggelin, W., and E. Laurenzi. 2024. "A Knowledge Graph-Based Decision Support System for Resilient Supply Chain Networks." In *Research Challenges in Information Science* edited by J. Araújo, J. L. de la Vara, M. Y. Santos, and S. Assar, 66–81. Cham: Springer Nature Switzerland.

EMA. 2023. "Availability of Medicines During COVID-19 (The European Medicines Agency, Ed.)." Retrieved October 11, 2024, from https://www.ema.europa.eu/en/human-regulatory-overview/public-health-threats/coronavirus-disease-covid-19/covid-19-public-health-emergencyinternational-concern-2020-23/availability-medicines-during-covid-19.

EMA. 2024. "Eu Actions to Tackle Shortages of GLP-1 Receptor Agonists (European Medicines Agency, Ed.)." Retrieved October 11, 2024, from https://www.ema.europa.eu/en/news/eu-actions-tackle-shortages-glp-1-receptor-agonists.

Emrouznejad, A., S. Abbasi, and Ç. Sıcakyüz. 2023. "Supply Chain Risk Management: A Content Analysis-Based Review of Existing and Emerging Topics." *Supply Chain Analytics* 3:100031. https://doi.org/10.1016/j.sca.2023.100031.

European Commission. 2020. "Pharmaceutical Strategy for Europe (European Commission, Ed.)." https://health.ec.europa.eu/system/files/2021-02/pharma-strategy_report_en_0.pdf.

European Commission. 2023. "Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Addressing Medicine Shortages in the EU (European Commission, Ed.)." Retrieved July 31, 2024, from https://commission.europa.eu/system/files/2023-10/Communication_medicines_shortages_EN_0.pdf.

Francas, D., M. Fritsch, and J. Kirchhoff. 2022. "Resilienz pharmazeutischer lieferketten: Studie für den verband forschender arzneimittelhersteller e. v. (vfa) (Verband Forschender Arzneimittelhersteller e. V., Ed.)." https://www.iwkoeln.de/studien/manuel-fritsch-jasmina-kirchhoff-resilienz-pharmazeutischer-lieferketten.html.

Francas, D., S. Mohr, and K. Hoberg. 2023. "On the Drivers of Drug Shortages: Empirical Evidence from Germany." *International Journal of Operations & Production Management* 43 (10): 1520–1538. https://doi.org/10.1108/IJOPM-09-2022-0581.

Franco, C., and E. Alfonso-Lizarazo. 2017. "A Structured Review of Quantitative Models of the Pharmaceutical Supply Chain." *Complexity* 2017:1–13. https://doi.org/10.1155/2017/5297406.

Haial, A., A. Berrado, and L. Benabbou. 2020. "Redesigning a Transportation Network: The Case of a Pharmaceutical Supply Chain." *International Journal of Logistics Systems and Management* 35 (1): 90. https://doi.org/10.1504/IJLSM.2020.103866.

Hao, J., L. Zhao, J. Milisavljevic-Syed, and Z. Ming. 2021. "Integrating and Navigating Engineering Design Decision-Related Knowledge Using Decision Knowledge Graph." *Advanced Engineering Informatics* 50:101366. https://doi.org/10.1016/j.aei.2021.101366.

Huss, G., S. Barak, L. Reali, C. Magendie, A. Carrasco-Sanz, E. Somekh, R. Cohen, et al. 2023. "Drug Shortages in Pediatrics in Europe: The Position of the European Pediatric Societies." *The Journal of Pediatrics* 261:113472. https://doi.org/10.1016/j.jpeds.2023.113472.

Ivanov, D. 2024. "Digital Supply Chain Management and Technology to Enhance Resilience by Building and Using End-to-end Visibility during the COVID-19 Pandemic." *IEEE Transactions on Engineering Management* 71:10485–10495. https://doi.org/10.1109/TEM.2021.3095193.

Ji, S., S. Pan, E. Cambria, P. Marttinen, and P. S. Yu. 2022. "A Survey on Knowledge Graphs: Representation, Acquisition, and Applications." *IEEE Transactions on Neural Networks and Learning Systems* 33 (2): 494–514. https://doi.org/10.1109/TNNLS.2021.3070843.

Kalaitzi, D., and N. Tsolakis. 2023. "Facing Shortages: Practical Strategies to Improve Supply Chain Resilience." In *Handbook for Management of Threats: Security and Defense, Resilience and Optimal Strategies*, edited by K. P. Balomenos, A. Fytopoulos, and P. M. Pardalos, 59–80. Cham: Springer

International Publishing. https://doi.org/10.1007/978-3-031-39542-0_4.

Karam, N., S. Matini, R. Laas, and T. Hoppe. 2023. "A Hybrid Knowledge Graph and Bayesian Network Approach for Analyzing Supply Chain Resilience." In *The Semantic Web: ESWC 2023 Satellite Events*, edited by C. Pesquita, H. Skaf-Molli, V. Efthymiou, S. Kirrane, A. Ngonga, D. Collarana, R. Cerqueira, M. Alam, C. Trojahn, and S. Hertling, 27–31. Cham: Springer Nature Switzerland.

Katsaliaki, K., P. Galetsi, and S. Kumar. 2022. "Supply Chain Disruptions and Resilience: A Major Review and Future Research Agenda." *Annals of Operations Research* 319 (1): 965–1002. https://doi.org/10.1007/s10479-020-03912-1.

Kazemian, I., S. A. Torabi, C. W. Zobel, Y. Li, and M. Baghersad. 2022. "A Multi-Attribute Supply Chain Network Resilience Assessment Framework Based on SNA-inspired Indicators." *Operational Research*22 (3): 1853–1883. https://doi.org/10.1007/s12351-021-00644-3.

Kosasih, E. E., F. Margaroli, S. Gelli, A. Aziz, N. Wildgoose, and A. Brintrup. 2022. "Towards Knowledge Graph Reasoning for Supply Chain Risk Management Using Graph Neural Networks." *International Journal of Production Research* 62 (15): 1–17.

Kumar, V., V. Bansal, A. Madhavan, M. Kumar, R. Sindhu, M. K. Awasthi, P. Binod, and S. Saran. 2022. "Active Pharmaceutical Ingredient (API) Chemicals: A Critical Review of Current Biotechnological Approaches." *Bioengineered* 13 (2): 4309–4327. https://doi.org/10.1080/21655979.2022.2031412.

Li, Y., and B. Starly. 2024. "Building a Knowledge Graph to Enrich ChatGPT Responses in Manufacturing Service Discovery." *Journal of Industrial Information Integration* 40:100612. https://doi.org/10.1016/j.jii.2024.100612.

Li, L., P. Wang, J. Yan, Y. Wang, S. Li, J. Jiang, Z. Sun, et al. 2020. "Real-World Data Medical Knowledge Graph: Construction and Applications." *Artificial Intelligence in Medicine* 103:101817. https://doi.org/10.1016/j.artmed.2020.101817.

Lim, K. Y. H., L. Liu, C. -H. Chen, and X. Gu. 2024. "Manufacturing Resilience through Disruption Mitigation Using Attention-Based Consistently-Attributed Graph Embedded Decision Support System." *Computers & Industrial Engineering* 197:110494. https://doi.org/10.1016/j.cie.2024.110494.

Liu, Y., B. He, M. Hildebrandt, M. Buchner, D. Inzko, R. Wernert, E. Weigel, D. Beyer, M. Berbalk, and V. Tresp. 2023. "A Knowledge Graph Perspective on Supply Chain Resilience." https://arxiv.org/abs/2305.08506.

Lücker, F., S. Chopra, and R. W. Seifert. 2021. "Mitigating Product Shortage Due to Disruptions in Multi–stage Supply Chains." *Production and Operations Management* 30 (4): 941–964. https://doi.org/10.1111/poms.13286.

Lücker, F., R. W. Seifert, and I. Biçer. 2019. "Roles of Inventory and Reserve Capacity in Mitigating Supply Chain Disruption Risk." *International Journal of Production Research* 57 (4): 1238–1249. https://doi.org/10.1080/00207543.2018.1504173.

Lv, C., Y. Lu, X. Yan, W. Lu, and H. Tan. 2020. "Supplier Recommendation Based on Knowledge Graph Embedding." In *2020 Management Science Informatisation and Economic Innovation Development Conference (MSIEID)*, 514–518. https://doi.org/10.1109/MSIEID52046.2020.00105.

Medicines For Europe. 2024. "Medicine Shortages and National Stockpiling Requirements in the EU (Medicines For Europe, Ed.)." Retrieved October 22, 2024, from https://www.medicinesforeurope.com/wp-content/uploads/2024/04/240327_Position-paper-stockpiling_FINAL3.pdf.

Neo4j. 2024a. "Genai Apps, Grounded in Your Data (Neo4j, Inc., Ed.)." https://neo4j.com/.

Neo4j. 2024b. "Query a Neo4j Database Using Cypher (Neo4j, Inc., Ed.)." Retrieved July 19, 2024, from https://neo4j.com/docs/getting-started/cypher-intro/.

Neo4j. 2024c. "Why Graph Databases?" Retrieved July 22, 2024, from https://neo4j.com/why-graph-databases/.

Nguyen, A., S. Lamouri, R. Pellerin, S. Tamayo, and B. Lekens. 2022. "Data Analytics in Pharmaceutical Supply Chains: State of the Art, Opportunities, and Challenges." *International Journal of Production Research* 60 (22): 6888–6907. https://doi.org/10.1080/00207543.2021.1950937.

Opsahl, T., F. Agneessens, and J. Skvoretz. 2010. "Node Centrality in Weighted Networks: Generalizing Degree and Shortest Paths." *Social Networks* 32 (3): 245–251. https://doi.org/10.1016/j.socnet.2010.03.006.

Palit, A., and P. Bhogal. 2022. "COVID19, Supply Chain Resilience, and India: Prospects of the Pharmaceutical Industry." In *Globalisation Impacts: Countries, Institutions and COVID19*, edited by A. Palit, 159–181. Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-16-7185-2_9.

Pathy, S. R., and H. Rahimian. 2023. "A Resilient Inventory Management of Pharmaceutical Supply Chains under Demand Disruption." *Computers & Industrial Engineering* 180:109243. https://doi.org/10.1016/j.cie.2023.109243.

Pauwels, K., S. Simoens, M. Casteels, and I. Huys. 2015. "Insights into European Drug Shortages: A Survey of Hospital Pharmacists." *PloS One* 10 (3): e0119322. https://doi.org/10.1371/journal.pone.0119322.

Peng, C., F. Xia, M. Naseriparsa, and F. Osborne. 2023. "Knowledge Graphs: Opportunities and Challenges." http://arxiv.org/pdf/2303.13948v1.

Phuong, J. M., J. Penm, B. Chaar, L. D. Oldfield, and R. Moles. 2019. "The Impacts of Medication Shortages on Patient Outcomes: A Scoping Review." *PloS One* 14 (5): 1–17. https://doi.org/10.1371/journal.pone.0215837.

Piatek, O. I., J. C. -M. Ning, and D. R. Touchette. 2020. "National Drug Shortages Worsen during COVID-19 Crisis: Proposal for a Comprehensive Model to Monitor and Address Critical Drug Shortages." *American Journal of Health-System Pharmacy* 77 (21): 1778–1785. https://doi.org/10.1093/ajhp/zxaa228.

Pokorný, J. 2015. "Graph Databases: Their Power and Limitations." In *Computer Information Systems and Industrial Management*, edited by K. Saeed and W. Homenda, 58–69. Cham: Springer International Publishing.

Rolf, B., N. Mebarki, S. Lang, T. Reggelin, O. Cardin, H. Mouchère, and A. Dolgui. 2022. "Using Knowledge Graphs and Human-Centric Artificial Intelligence for Reconfigurable Supply Chains: A Research Framework." *IFAC-PapersOnLine*55 (10): 1693–1698. https://doi.org/10.1016/j.ifacol.2022.09.641.

Rotmensch, M., Y. Halpern, A. Tlimat, S. Horng, and D. Sontag. 2017. "Learning a Health Knowledge Graph from Electronic Medical Records." *Scientific Reports* 7 (1): 5994. https://doi.org/10.1038/s41598-017-05778-z.

Sabouhi, F., M. S. Pishvaee, and M. S. Jabalameli. 2018. "Resilient Supply Chain Design under Operational and Disruption Risks considering Quantity Discount: A Case Study of Pharmaceutical Supply Chain." *Computers & Industrial Engineering* 126:657–672. https://doi.org/10.1016/j.cie.2018.10.001.

Saedi, S., O. E. Kundakcioglu, and A. C. Henry. 2016. "Mitigating the Impact of Drug Shortages for a Healthcare Facility: An Inventory Management Approach." *European Journal of Operational Research*251 (1): 107–123. https://doi.org/10.1016/j.ejor.2015.11.017.

Schmitt, A. J., and B. Tomlin. 2012. "Sourcing Strategies to Manage Supply Disruptions." In *Supply Chain Disruptions*, edited by H. Gurnani, A. Mehrotra, and S. Ray, 51–72. Dordrecht: Springer London. https://doi.org/10.1007/978-0-85729-778-5_3.

Schwarz, M. 2021. "Warum wir stärkere lieferketten bei generischen arzneimitteln brauchen: Konzept für mehr resilienz und für mehr versorgungssicherheit (Sarticon GmbH, progenerika, Ed.)." https://www.progenerika.de/app/uploads/2022/04/Studie_Warum-wir-staerkere-Lieferketten-brauchen_2021.pdf.

Shah, N. 2004. "Pharmaceutical Supply Chains: Key Issues and Strategies for Optimisation." *Computers & Chemical Engineering* 28 (6–7): 929–941. https://doi.org/10.1016/j.compchemeng.2003.09.022.

Shen, X., X. Li, B. Zhou, Y. Jiang, and J. Bao. 2023. "Dynamic Knowledge Modeling and Fusion Method for Custom Apparel Production Process Based on Knowledge Graph." *Advanced Engineering Informatics* 55:101880. https://doi.org/10.1016/j.aei.2023.101880.

Shukar, S., F. Zahoor, K. Hayat, A. Saeed, A. H. Gillani, S. Omer, S. Hu, Z.-U.-D. Babar, Y. Fang, and C. Yang. 2021. "Drug Shortage: Causes, Impact, and Mitigation Strategies." *Frontiers in Pharmacology*12:693426. https://doi.org/10.3389/fphar.2021.693426.

Singhal, A. 2012, May 16. "Introducing the Knowledge Graph: Things, Not Strings [Google]." Retrieved January 23, 2024, from https://blog.google/products/search/introducing-knowledge-graphthings-not/.

Snyder, L. V., Z. Atan, P. Peng, Y. Rong, A. J. Schmitt, and B. Sinsoysal. 2016. "Or/ms Models for Supply Chain Disruptions: A Review." *IIE Transactions* 48 (2): 89–109. https://doi.org/10.1080/0740817X.2015.1067735.

Sowa, J. F. 2006. "Semantic Networks." In *Encyclopedia of Cognitive Science*, edited byNadel Lynn. John Wiley & Sons, Ltd. https://doi.org/10.1002/0470018860.s00065.

Spieske, A., M. Gebhardt, M. Kopyto, and H. Birkel. 2022. "Improving Resilience of the Healthcare Supply Chain in a Pandemic: Evidence from Europe during the COVID-19 Crisis." *Journal of Purchasing and Supply Management* 28 (5): 100748. https://doi.org/10.1016/j.pursup.2022.100748.

Tian, L., X. Zhou, Y.-P. Wu, W.-T. Zhou, J.-H. Zhang, and T.-S. Zhang. 2022. "Knowledge Graph and Knowledge Reasoning: A Systematic Review." *Journal of Electronic Science and Technology* 20 (2): 100159. https://doi.org/10.1016/j.jnlest.2022.100159.

Tu, Y., W. Li, X. Song, K. Gong, L. Liu, Y. Qin, S. Liu, and M. Liu. 2024. "Using Graph Neural Network to Conduct Supplier Recommendation Based on Large-Scale Supply Chain." *International Journal of Production Research* 62 (24): 8595–8608. https://doi.org/10.1080/00207543.2024.2344661.

Tucker, E. L., Y. Cao, E. R. Fox, and B. V. Sweet. 2020. "The Drug Shortage Era: A Scoping Review of the Literature 2001–2019." *Clinical Pharmacology and Therapeutics* 108 (6): 1150–1155. https://doi.org/10.1002/cpt.1934.

Tucker, E. L., and M. S. Daskin. 2022. "Pharmaceutical Supply Chain Reliability and Effects on Drug Shortages." *Computers & Industrial Engineering* 169:108258. https://doi.org/10.1016/j.cie.2022.108258.

Vann Yaroson, E., L. Breen, J. Hou, and J. Sowter. 2023. "Examining the Impact of Resilience Strategies in Mitigating Medicine Shortages in the United Kingdom's (UK) Pharmaceutical Supply Chain (PSC)." *Benchmarking: An International Journal* 31:683–706. https://doi.org/10.1108/BIJ-07-2022-0460.

Ventola, C. 2011. "The Drug Shortage Crisis in the United States: Causes, Impact, and Management Strategies." *Pharmacy and Therapeutics* 36:740–757.

WHO. 2024. "Anatomical Therapeutic Chemical (ATC) Classification." Retrieved June 19, 2024, from https://www.who.int/tools/atc-ddd-toolkit/atc-classification.

WIdO. 2024. "PharMaAnalyst." Retrieved June 19, 2024, from https://arzneimittel.wido.de/PharMaAnalyst/?0.

Wyrembek, M., G. Baryannis, and A. Brintrup. 2025. "Causal Machine Learning for Supply Chain Risk Prediction and Intervention Planning." *International Journal of Production Research* 1–20. https://doi.org/10.1080/00207543.2025.2458121.

Yang, F., J. Li, Y. Shen, and M. Liu. 2023. "Research on the Construction of Logistics Industry Knowledge Graph Based on Big Data." In *2023 7th Asian Conference on Artificial Intelligence Technology (ACAIT)*, 1221–1228. https://doi.org/10.1109/ACAIT60137.2023.10528528.

Yang, Y., C. Peng, E.-Z. Cao, and W. Zou. 2024. "Building Resilience in Supply Chains: A Knowledge Graph-Based Risk Management Framework." *IEEE Transactions on Computational Social Systems* 11 (8): 1–9.

Yang, P., H. Wang, Y. Huang, S. Yang, Y. Zhang, L. Huang, Y. Zhang, et al. 2024. "Lmkg: A Large-Scale and Multi-Source Medical Knowledge Graph for Intelligent Medicine Applications." *Knowledge-Based Systems* 284:111323. https://doi.org/10.1016/j.knosys.2023.111323.

Zhang, W., Y. Liu, L. Jiang, N. Shah, X. Fei, and H. Cai. 2020. "The Construction of a Domain Knowledge Graph and Its Application in Supply Chain Risk Analysis." In *Advances in E-Business Engineering for Ubiquitous Computing*, edited by K.-M. Chao, L. Jiang, O. K. Hussain, S.-P. Ma, and X. Fei, 464–478, Springer International Publishing.

Zhou, S., Y. Liu, and Y. Liu. 2024. "A Market Convergence Prediction Framework Based on a Supply Chain Knowledge Graph." *Sustainability* 16 (4): 1696.

Zhou, G., L. Tan, X. Xie, J. Li, and H. Luo. 2022. "Clinical Decision Support System for Hypertension Medication Based on Knowledge Graph." *Computer Methods and Programs in Biomedicine* 227:107220. https://doi.org/10.1016/j.cmpb.2022.107220.

Zou, X. 2020. "A Survey on Application of Knowledge Graph." *Journal of Physics: Conference Series* 1487 (1): 012016.