

OTDK PAPER

2025

**UNIVERSITY OF DEBRECEN
FACULTY OF ECONOMICS AND BUSINESS**

**EGY VÁLLALAT PÉNZÜGYI HELYZETÉNEK ÉRTÉKELÉSE
PÉNZÜGYI MUTATÓK ÉS SZÖVEGES ELEMZÉSEK
ALKALMAZÁSÁVAL AZ ÉVES JELENTÉSEKBEN: EGY
ALKALMAZOTT ESETTANULMÁNY**

**ASSESS FINANCIAL HEALTH OF A COMPANY USING FINANCIAL
RATIOS AND TEXTUAL ANALYSIS IN ANNUAL REPORTS: AN
APPLIED CASE STUDY**

Author: Do Quang Hieu

Consultant: Dr. Vilmos Lakatos

Closure date of the manuscript: 29.10.2024

TABLE OF CONTENTS

INTRODUCTION	1
1. RAISING OF TOPIC	2
2. REVIEW OF RELEVANT LITERATURE	3
2.1. Overview of financial health assessment and financial ratios.....	3
2.2. Textual analysis in financial reporting	7
2.3. Pharmaceutical industry and its characteristics.....	12
2.4. Pharmaceutical industry of Hungary	16
3. MATERIAL AND METHOD	20
3.1. Overview of company Gedeon Richter	20
3.2. Methodology	21
3.2.1. Financial ratio analysis.....	21
3.2.2. Textual analysis of annual reports	22
3.2.3. Data collection.....	24
3.2.4. Analytical framework.....	25
4. FINDINGS AND THEIR ANALYSIS	26
4.1. Financial health analysis using ratios.....	26
4.2. Textual analysis results.....	30
4.3. Integration of financial ratios and textual findings	42
5. CONCLUSIONS AND RECOMMENDATIONS	43
5.1. Methodological contributions	43
5.2. Limitations of the study.....	43
5.3. Suggestions for future research	44
SUMMARY	45

INTRODUCTION

Evaluating the financial health of a company is an important task for the management, creditors, investors and policy makers. Accurate financial assessment is not only useful for internal decision making but also for external stakeholders to know a company's stability and performance. Traditionally, information is limited to financial ratios analysis derived from financial statements which are informative but fail to capture the complexity of financial situations.

Even though significant developments have been made inside the financial analysis field like advance statistical techniques, machine learning approaches in financial modellings, they all depend on quantitative data in financial statements and neglect qualitative data from management.

Moreover, in the situation of modern times, when the market is constantly evolving, relying on financial information from financial statements is not sufficient. To understand better the financial health of a company, a new type of data could be potentially explored even though it is often overlooked in traditional analyses which is textual data.

The deployment of textual analysis methods in accounting have been a popular tool and developed from traditional manual content analysis to text mining with the help of computer. It is performed on various types of textual data from companies and social media to capture the trend and increase predictability of many aspects like bankruptcy, stock price, return on equity. The technique is commonly used in large textual data, while the usage of it in small text samples is hardly seen.

The study wants to address the gap in financial analysis by make use of textual analysis for small, focused text data which is annual report of 5 years in the company case study (Gedeon Richter) and demonstrate that even the small text data could provide valuable information on company financial health which usually be overlooked by conventional analyses.

1. RAISING OF TOPIC

The study is exploratory in nature. applying a new methodology: textual analysis in comparison with financial ratios analysis on a case study – Gedeon Richter in the period of 2019 to 2023. The primary objective is to present a holistic view of the financial health of Gedeon Richter, from there contribute fresh perspectives to the traditional financial analysis methods.

Firstly, the reason case study research is conducted instead of using a sample of companies is because of the substantial difference in the nature of each company. While sample-based studies provide generalizability, it would not capture the distinctive interactions between company-specific financial situation, strategies, products and operations with the external forces. Using sample-based study would risk oversimplifying the complexity of company-specific settings.

In addition, company-specific analysis is a well-established practice in financial analysis, especially investment analysis. The study wants to specifically contribute to this area by using a case study as an exploratory approach to illustrate the contribution of integrating two methods.

Secondly, the reason choosing this case study (Gedeon Richter) is because the company is working in pharmaceutical market where the regulatory, operational and market risk is high especially in the period of 2019 to 2023 when crisis like Covid-19 and Russian-Ukrainian war happen.

The specific tasks to achieve the objectives of the research are:

1. Assess financial health using traditional financial ratios analysis method.
2. Conduct textual analysis on textual data of annual reports (in this case is CEO's letters).
3. Integrate 2 methods to see if they give a more comprehensive assessment with new insights.

The expected outcome of this study is a well-rounded evaluation of Gedeon Richter's financial condition with both numerical measurement and management justification. Ultimately, the study aims to demonstrate the benefit of using both methods: financial ratios and textual analysis to accurately reflect the picture of company financial health.

2. REVIEW OF RELEVANT LITERATURE

2.1. Overview of financial health assessment and financial ratios

Financial health is a complex concept which does not have concrete definitions (Guan et al., 2021). It could refer to people or a variety of businesses, for example small businesses, non-profit businesses and for-profit businesses. For people, financial health reflects ability to manage expenses, being ready and recovering from financial shocks, have minimal debt and building wealth (Weida et al., 2020). For businesses, the concept is not that different, in case of small businesses, it is the ability to withstand adverse conditions and expand due to market opportunities (Koryński, n.d.). With non-profit businesses, it has two elements which are financial capacities and financial sustainability (Guan et al., 2021). In general for any organization, it is ability to use resources most effectively for generating profit with the lowest degree of risk (Alhasanko, 2024). Some factors that affect financial health are revenue, expense, debt management and cash flow (*Figure 1*).

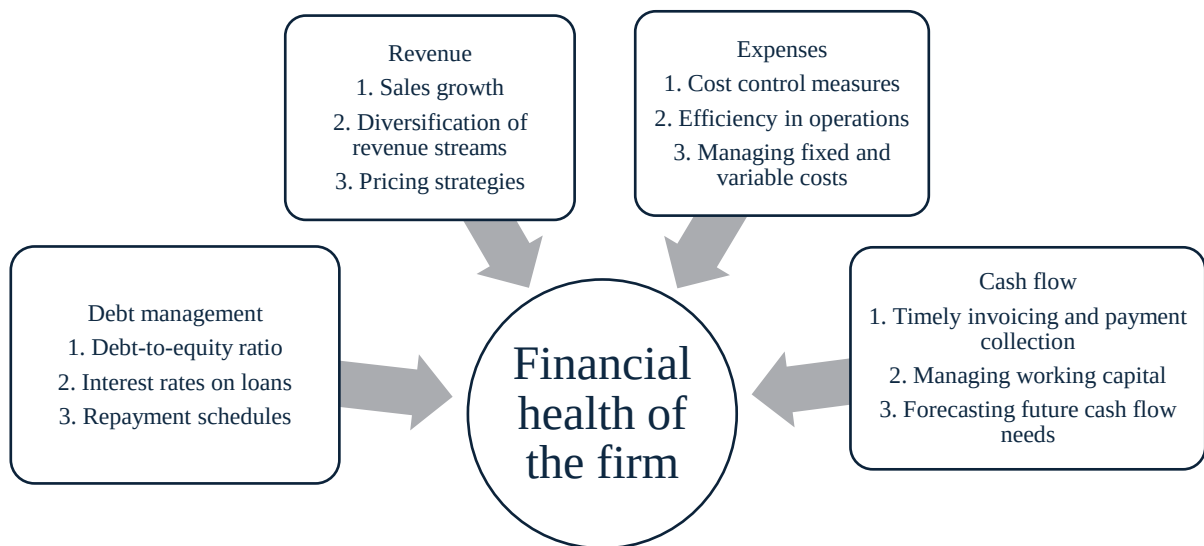


Figure 1: **Factors affecting of the financial health of the firm**

Source: Alhasanko (2024)

Compared to financial health, financial performance focuses on profitability of the company and strongly related to 4 financial ratios: return on assets (ROA), return on equity (ROE), return on sales (ROS), and earnings per share (EPS) (Martin – Siehl, 1990). While on the other hand, financial position emphasizes elements in the balance sheet which are associated with assets, equities and liabilities. Other concepts which are closely related to financial health could be named are financial distress, default and bankruptcy. A firm which are in trouble pay its debt is considered to be in financial distress, to the point cannot paid its debt or required

interest is deemed to be defaulted and if the debt holder decides to take the ownership, it becomes bankruptcy (Berk – DeMarzo, 2017).

Financial health is important for stakeholders even with different perspectives. Firstly, for business owners, financial health is a comprehensive picture of company and equip management decisions, secondly, for lenders, financial health is a tool for lending decisions, thirdly, for policy makers, the information will be used to develop better policies (Koryński, n.d.).

Alhasanko (2024) had created a roadmap to analyzing financial health of a firm in 5 steps (*Figure 2*): the first one is looking at goals, objective and operating characteristics of the firm, the second step is assessing future sales, market growth and competitive forces, the third step is estimating value of investment which support the product-market strategy, the fourth one is evaluate the profitability of the firm as it affect debt financing ability, willingness of common stock issue and valuation and show the sustainable growth of the firm, and final step is appraise future external financing needs from prior steps.

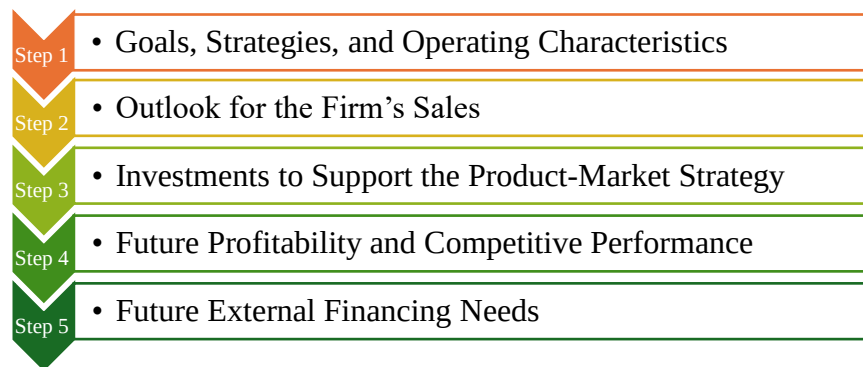


Figure 2: Steps to evaluate the financial health of the firm

Source: Alhasanko (2024)

Assessing financial health need financial information in financial statements from which financial ratios could be calculated as it is the primary tools of analysis, the financial ratios can also be benchmarked with competing companies to assess financial position and predict future bankruptcy (Alhasanko, 2024; Bertoneche – Knight, 2001). According to Bertoneche and Knight (2001), the fourth key dimensions of financial health and financial ratios are: profitability ratios, efficiency ratios, leverage ratios and liquidity ratios.

Financial ratios is defined as quantitative value created from two or more value in financial statements usually in in percentage, ratios or multiple form, the advantage of financial ratios is it is independent of firm's size and can allow meaningful comparison (Arkan, 2016). Financial ratios are used for all types of purpose but there are 2 principle uses: traditional, normative use of firm ratios compared with standard and positive use estimating relationships

mostly for predictive purposes (Barnes, 1987). According to Arkan (2016), there are 3 types of comparisons:

- Cross sectional comparison: this means the firm will be compared with other firms. Cross sectional comparison could be divided into intra-industry which other firms come from the same industry with ratio is the industry average and Inter-industry which other firms from different industries with ratios is the average of other industries. Between two comparisons, inter-industry is challenging because the difference among the industries makes it difficult to compare.
- Intertemporal comparison (horizontal analysis or trend analysis): the financial ratios of the firm will be compared over time to find the trend.
- Arbitrary standards comparison: the financial ratios of the firm are compared with the traditional benchmarks which are acceptable or traditional of its industry.

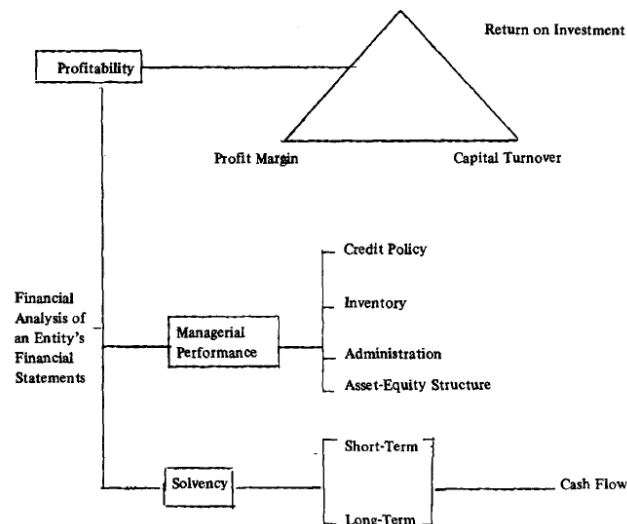


Figure 3: **Financial ratios categoric framework**

Source: Courtis (1978)

DENOMINATOR OF RATIO	NUMERATOR OF RATIO																		Totals			
	Annual Funds Flow	Cash	Cash Flow	Cost of Sales	Current Assets	(Each) Current Asset	Current Liabilities	Debt	Dividends	EBIT	Fixed Assets	Gross Profit	Inventory	Net Income	Net Worth	Op.Exp. & C.G.S.	Quick Assets	Retained Earnings		Sales	Total Debt	Working Capital
Accounts Receivable																			CP			1
Cash Flow									R												CF	2
Current Assets						AE							I									2
Current Liabilities	CF	L	CF		L						AE											5
Debt											AE											1
Fixed Assets															S				C			2
Fixed Charges										S												1
Interest Expense										S												1
Inventory				I		I													I			3
Net Income								R							AE			AE				3
Net Worth			CF			L					AE			R				C		S	AE	7
Number of Common Shares			CF					R						R								3
Sales		L	CF		A	L						P		R		A		L				8
Total Assets		L	CF				AE	AE	R	AE	R	I		R	AE		L	AE	C	S	L	16
Total Debt			CF				AE	AE					I	R								3
Working Capital							AE	AE						R				C				5
	1	3	6	2	3	1	4	3	3	3	3	2	3	6	3	1	3	2	6	2	3	63

*The abbreviations used in the body of the table are: R = return on investment; P = profit margin; C = capital turnover; CP = credit policy; I = inventory; A = administration; AE = asset equity; L = short term liquidity; S = long term solvency; CF = cash flow

Figure 4: A visual approximation of collinearity of 63 ratios

Source: Courtis (1978)

There are numerous combinations for numerator and denominator of financial ratios and to decide whether a financial ratio is useful and how to categorize them is crucial in financial statements analysis. In prior literatures, there are groups of ratios: Activity, efficiency, solvency, leverage, turnover, profitability, coverage, cash flow, Courtis (1978) had created a matrix to categorize 63 financial ratios in a more logical way of approach 3 questions which is present in the framework: is the firm profitable? Is management good? And is the firm stay in business? (Figure 3,4). The ratios are many and some was found to be useful and be advocated in different textbooks, predictive studies, so the analysts should carefully choose which ratios to calculate that suit their need instead of calculating all of them (Courtis, 1978).

There is development from comparison of financial ratios to model. Around 1919, the birth of DuPont model to appraise operating result of the firm by decomposing return on equity (ROE) into 3 financial ratios: turnover, margin and leverage is a famous model represent the normative use of financial ratios (Saus-Sala et al., 2021). The positive use of financial ratios has 2 types: by accountant and analysts to forecast financial variable like future profit by take the product of predicted sales with profit margin, and by researchers who use statistical methods to predict bankruptcy possibility, credit rating, risk assessment and testing hypothesizes with input of financial ratios (Barnes, 1987). Although the second use of financial ratios is varied but the focus is on predicting bankruptcy, Beaver (1966) use univariate analysis (profile analysis and dichotomous classification test) to assess the prediction ability of each ratios from 30 most common ratios to find out the best predictors and Altman (1968) improve Beaver work with

famous Altman Z-score model using multiple discriminant analysis (MDA) method which is a multivariate analysis, the model was revised in 1997 with ZETA model and many other papers addressing business default with the same approach. Wilcox (1976) used probabilistic model gambler's ruin which describe businesses as gamblers but the results is not good as it considered cash flow as independent from each other. Ohlson (1980) used logit and probit model as it has some advantages over MDA – no assumptions must be made on prior probabilities of bankruptcy and distribution of predictors. Theodossiou (1993) predicted business failure by predicting shifts in the mean of a multivariate time series process, from assessing the drastic change to the mean of time dependent financial indicator variables.

To predict bankruptcy, beside statistical approaches, the development of machine learning methods like Artificial Neural Networks (ANN), Support Vector Machines (SVM), Decisions Trees (DT) and other machine learning methods are used as it provide better accuracy for smaller data while statistical methods are used for large dataset (Dept. of CSE, GIT, GITAM University, Visakhapatnam, Andhra Pradesh, India et al., 2018).

Financial ratios have some limitations, firstly, as it depends on accounting principles, accounting classifications and accounting methods varied in time and among firms, so it reduces the comparability of the ratios (Faello, 2015). Secondly, the ratios are not as good in predicting the future, even though the usefulness of them in credit analysis is indisputable, usually the future of a firm would be affected by uncontrollable variables, lastly, financial ratios is not giving answers but it directs the attention of analyst to where the important question needs to be answers to make investment decisions (Arkan, 2016).

2.2. Textual analysis in financial reporting

The International Accounting Standards Board (2010) and Financial Accounting Standards Board (2018) defines financial reporting as process of providing financial information on reporting entity that is useful for current and future investors, lenders, and creditors to make decisions: (1) Should they buy, keep or sell the equity or debt instruments? (2) giving and settling loans or other type of credits, (3) affecting management's action by using right to votes. That process also brings about information on the reporting entity's economic resource, claim against it and effect of transactions, events, conditions change their resource and claim, some could include management's expectations, strategies and forward-looking information. However, financial reporting is not a perfect communication tool as the users are different from each other, the information is only used as estimation tools, not provide the value of entity.

To be suitable for the need of important stakeholders especially investors, financial statements are used to convey information of 4 areas: financial performance which is the return on resource it controls, financial position which is the economic resource, financial structure, liquidity and solvency and capacity to change, the generation and use of cash and lastly financial adaptability which is short term ability to raise capital, pay debt, obtain cash from disposable asset without disruption and rapid improvement in net cash flow (Elliott – Elliott, 2009).

The quality characteristics of financial information are important in financial reporting, with the constraint of cost, financial information needs to keep these characteristics to be useful (*Figure 5*). Firstly, the information must be relevant, which can influence the decision making by assisting to predict future or confirm current, past event like looking-forward statements, business opportunities and risks, use of fair value (van Beest et al., 2009). Secondly, it must be reliable, faithfully reflect the economic phenomenon in words and numbers, neutral and prudent in presenting, free from errors, and the information must be complete. Thirdly, comparability is essential to determine financial position, cash flow and performance of entity across time and among firms in the same period, to do this, the note needs to disclose all the change in accounting policies, the consistency in application of accounting principles is crucial (Herath, 2017). Fourthly, understandability quality could be improved through effective communications like graphs and Tables, giving the users the information they need in a clear and sufficient way. Finally, the concept of materiality is a threshold which means that any items which is material if it is wrongly or not reported will affect the decision of users or assessment of management's stewardship, the concept justify the reason not to report an item which could cause extra cost for reporters and hinder decision-making processes with excessive information (Elliott – Elliott, 2009).

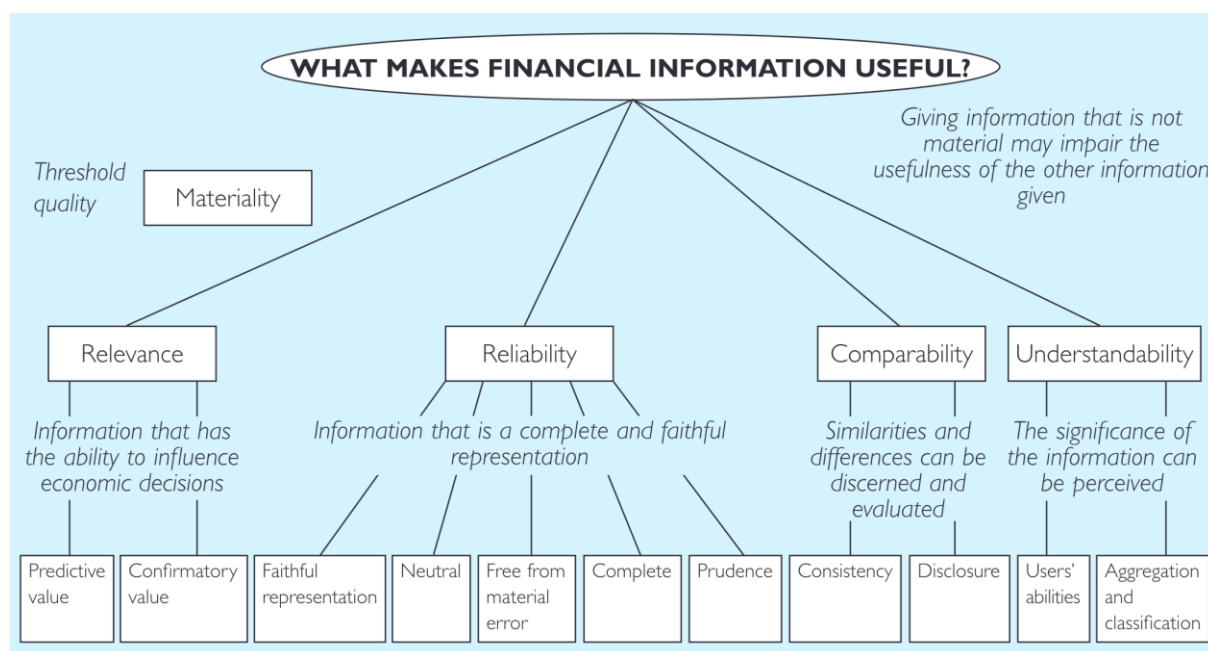


Figure 5: **What makes financial information useful?**

Source: Elliott – Elliott (2009)

The formal communication vehicle between corporation and its stakeholders is annual reports which not only contain quantitative information, graphs, Tables, photographs and narratives (Stanton et al., 2004). Annual reports usually have 2 parts: the rear sections with compulsory financial statements and the front sections with voluntary disclosures, this part have various names: management's discussion and analysis (MD&A), operating and financial review and strategic report (Stanton et al., 2004; International Financial Reporting Standards [IFRS], 2021). The MD&A and President's letter are the text used to provide overview of past performance and future plan, even though these part receive little recognition and people tend to focus more on financial part but it have proved to be an important research materials (Yuthas et al., 2002). As recognition of the impact of the management commentary, in 2010, IASB published IFRS Practice Statement *Management Commentary*, which aim to help management presenting management commentary in accordance with IFRS, even though the statement itself is not IFRS, meaning not compulsory (International Financial Reporting Standards [IFRS], 2010). The new exposure draft suggests 6 elements needed in the management commentary, which are also stated in the 2010 practice statement (*Figure 6*): (1) the business model or nature of the business which mean how it generates money, (2) management objectives and strategy to develop the business model, (3) resources and relationships which business model and strategy depend, even asset not on financial statements, (4) the risks which the business is exposed to, (5) external environment could affect the business and (6) management evaluation

and measurement of financial performance and position against stated objectives (International Financial Reporting Standards [IFRS], 2021).

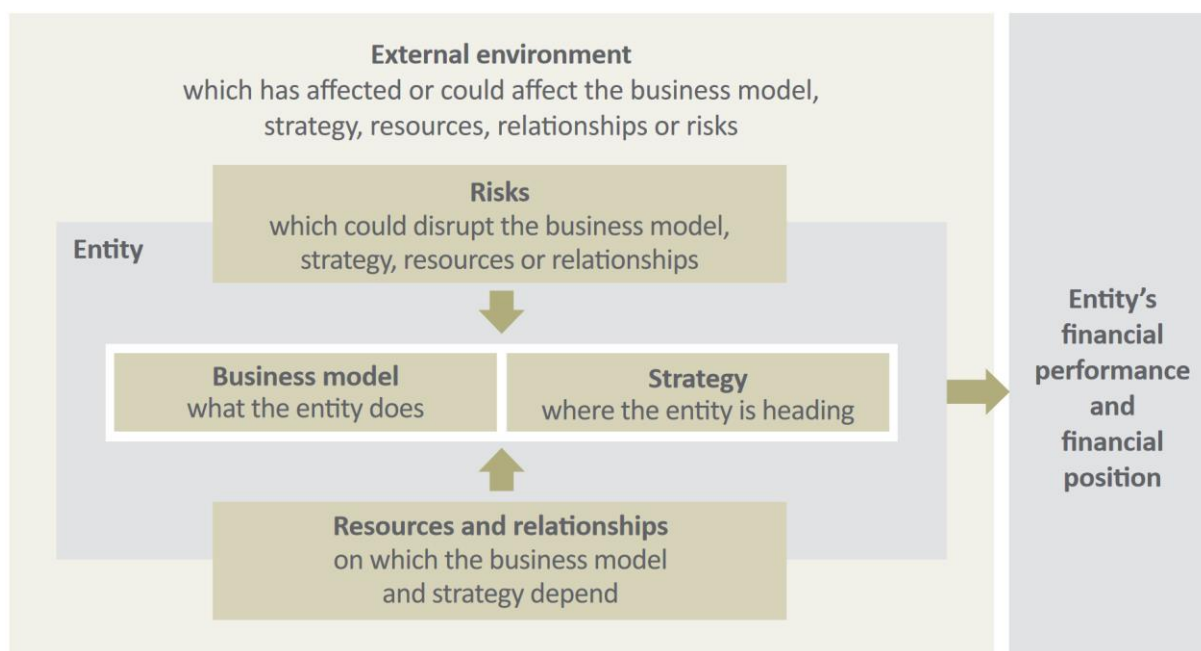


Figure 6: **Relationships between the six areas of content in management commentary**

Source: International Financial Reporting Standards [IFRS] (2021)

Although management commentary possesses a great amount of information, to analyze it could be difficult compared to financial statements analysis because the information here is unstructured text data. One effective methodology is textual analysis. Textual analysis is defined as a methodology in academic research to collect and analyze information which “make an educated guess at some of the most likely interpretations that might be made of that text.” (McKee, 2001) or “any systematic reduction of a flow text (or other symbols) to a standard set of statistically manipulable symbols representing the presence, the intensity, or the frequency of some characteristics relevant to social science” (Mehl, 2006). The data used for textual analysis is ‘text’ which is defined as “an utterance (speech) or set of utterances fixed by writing in which the writing itself is a constituent of the text” (Ricoeur, 1971).

Textual analysis could have many forms but divided into quantitative textual analysis and qualitative textual analysis. Qualitative textual analysis has popular approaches like classical hermeneutics, coding in Grounded Theory, content analysis and qualitative content analysis, research practice as described in textbook, ... In context of accounting fields, qualitative content analysis is mostly used. Quantitative textual analysis, which refers to using methods to make statistical reference from a text populations (Roberts, 2000), includes methods in accounting research are text mining and quantitative content analysis.

Content analysis is a method to analyze variety of data like visual or verbal data, which is used to find patterns, relationships and categories in data through iterative process of coding and interpretation of researchers, the object data could be variety include text, image, video,... and the output are Tables, graphs and text citations, qualitative content analysis will take into consideration of latent meaning while quantitative content analysis will only consider the objective meaning and text is coded automatically using a dictionary. On the other hand, text mining is process of looking for insight from text with the help of coding and analyzing using computer program and the role of subjectivity of research is limited, and for text mining its object of research is only text with output are Tables and graphs. (Bologna University, Italy – Aureli, 2017, Kwartler, 2017).

The using textual analysis in accounting is an emerging field with developmental technologies for automation over time (Chakraborty – Bhattacharjee, 2020). For content analysis approach, Smith and Taffler (2000) predict bankruptcy of a dataset of manufacturing and construction companies using textual variable to create a discriminant function like Altman's Z-score, they divided the methods into 2 approaches: keyword variables (form oriented) which using the appearance of keywords as predictors for failure of company as some words appear exceptionally in chairman's letter of bankrupted firms, theme variable (meaning oriented) which classified using investigator code. An important result here is that chairman's letter is able to classify failed and non-failed firms, and even though the documents is not audited, the distortion is also limited as it is looked at by outside parties (Smith – Taffler, 2000).

The quantitative content analysis approach continues to develop with the creation of dictionary, there are two types of dictionaries: general purpose dictionaries (Diction, GI/Harvard...) and finance specific dictionaries (Loughran and McDonald, Henry...). Diction is a computer-aided content analysis program developed by Roderick Hart who is specialist in politics and mass media and General Inquirer (GI) which is a system of computer programs developed by Philip Stone who is specialist in social psychology to do content analysis, they both commonly used for general content analysis purpose (Green et al., 1967; Henry – Leone, 2009; Short – Palmer, 2008). Henry (2006b, 2006a) compiled specific wordlist to measure tone of earning announcements, even though processed through Diction software but the author reminded that the software also allowed using customized wordlists and in the author suggest researchers want to measure qualitative information in financial disclosure should use domain specific dictionary to increase the test power (Henry – Leone, 2009). Loughran and Mcdonald (2011) created new dictionary with 7 categories: Negative, Positive, Uncertainty, Litigious, Strong Modal, Weak Modal, Constraining, the list has more power than general purpose one

like GI because as there are misclassification from industry effects. Loughran and McDonald (2014) have pointed some disadvantage of using Fog index as readability metric for accounting and finance text, the metrics is calculated by aggregate average words in the sentence and percentage of complex words which is proportion of words with 2 syllables, and for conclusion the use of file size is emphasize as a useful measure of textual readability. Loughran and McDonald (2024) developed a wordlist to measure the complexity of companies.

The methodologies above belong to BOW (Bag-of-words method) in which it analyzes the group of words in document without considering the grammar and order, the method is easy to use. The second method which syntactic parsing which is more complex, it tag POS (part of speech) to find out text block building the sentence in a grammatically context (Kwartler, 2017).

For machine learning approach, Antweiler and Frank (2001) used machine learning methods: Naive Bayes (NB) and Support Vector (SV) to classified messages into “buy”, “sell”, and “hold” then aggregate them to calculate the market sentiment, from that he could find out how it impacts stock trade volume. Das and Chen (2007) approach the same relationship of sentiment from stock message boards and stock market through various classifiers: Naive classifier, Vector distance classifier, Discriminant-Based Classifier, Adjective-Adverb Phrase Classifier., Bayesian Classifier. Balakrishnan et al (2010) summarized data of narrative disclosures from annual reports using TF-IDF (term frequency-inverse document frequency) then train a model to predict performance of next year using this year data with and without text characteristic variables, the results show an improved prediction. Brown and Tucker (2011) used Vector Space Model (VSM) to calculate the cosine similarity of 2 documents which is MD&A modification through years and examine which factor contribute to the score. Li (2010) utilized Naive Bayesian algorithm for categorizing tone and content of FLSs (forward-looking statements) in MD&A, the tone has predictability for future performance, besides the dictionary approach seems to not work well in the research but at this time, dictionaries used are too general and not specialized for accounting and finance. Hajek et al (2014) use multiple forecasting methods like logistic regression, AIRS2-p algorithms...to include the sentiment categories which is categorize from the Loughran and McDonald (2011) dictionary to see if it improves prediction ability of company’s performance classified by Altman-Z score.

2.3. Pharmaceutical industry and its characteristics

The pharmaceutical industry has grown significantly in size during the past few years. The entire pharmaceutical market was projected to be worth approximately 1.6 trillion US dollars in 2023. This represents a rise of more than \$100 billion from 2022. (IQVIA, 2024).

With the continuous evolution throughout history, it is more structured with scientific, technological, regulatory, and market changes and developments (Malerba – Orsenigo, 2015). The pharmaceutical industry is different from other industries with its special characteristics:

- Heavily regulated industry: The entry requirement (evaluation for safety, efficacy, and quality) for new drugs and pricing of (patented) drugs should be approved by high authority. Besides, regulatory entities are strict on marketing information about which one is permitted (Danzon, 2006; Tepperman, 2009). Patent protections are important as the industry significantly depends on it; absence of the barriers could result in free riding R&D effort. (Morse, 2003).
- Multiple agents in the market: The agents are diverse and heterogeneous in preferences: consumers, firms, government entities, research centers and universities,... (Malerba – Orsenigo, 2015). For prescription drugs which need doctors' consent to be able to purchase from licensed pharmacists, in this way doctors are the imperfect agents for consumers who make decision for other but based on consideration of professional responsibility and malpractice liability risks. For Over-the-counter (OTC) drugs which could be bought in retail outlets, the pharmacies will decide which version will be on stock. (Scherer, 2000; Tepperman, 2009). For consumers, as the price of drug will be covered some part from insurance, their demand would be less elasticity. Besides, even if the insurance not cover the expense, consumers who have enough money would be willing to buy (Scherer, 2000). *Figure 7* adapted from Tepperman (2009) to represent the relationship between the main agents.

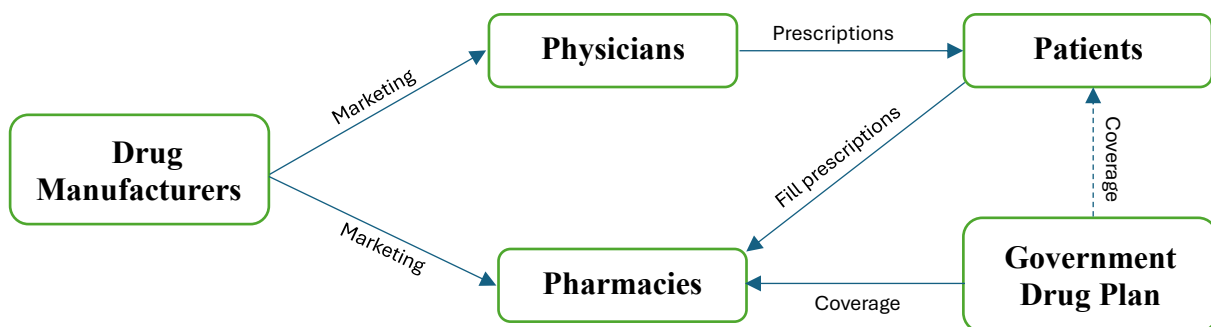


Figure 7: Agents in pharmaceutical market

Source: Adapted from Tepperman (2009)

- Long and costly R&D process: According to Joint Research Centre – European Commission, analyzing 2500 companies that invested the largest sums into R&D worldwide 2019, the pharmaceutical industry has the highest R&D intensity of 15.4%.

In 2022, the R&D spent globally is estimated to be 244 billion USD (Evaluate Pharma 2023). The development of a new medical product could take decades, from information of European Federation of Pharmaceutical Industries and Associations (EFPIA) , about 10 years will spend on R&D (*Figure 8*), after those 2-3 years of administrative procedures, and when the patent expired around year 20, maximum 5 years SPC (supplementary protection certificate) could be provided. Moreover, the estimate cost of developing a new medical product is varies widely, ranging from US\$ 43.4 million to US\$ 4.2 billion (Moon et al., 2022).

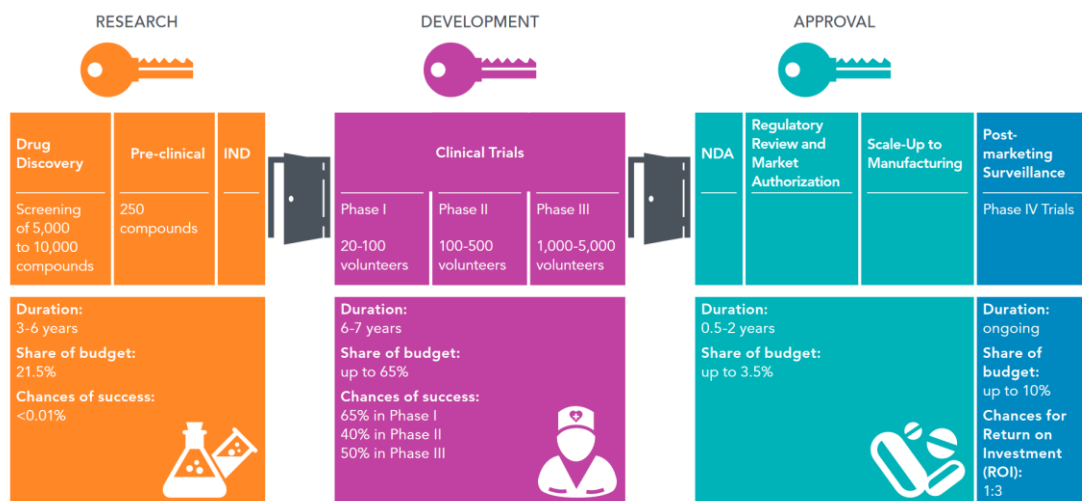


Figure 8: Research and development process

Source: Pugatch Consilium (2015)

- Product differentiation and marketing: Marketing cost of medical product (branded drugs) spend mostly on communication with physicians as their decisions on therapy are dominant, and they compete on attributes not price. For OTC drugs, consumers have more powers to decide so marketing effort could target other important consumers like pharmacists, hospitals, wholesalers...(Szalkai, 2004; Tepperman, 2009).
- Complex payment system: Adapt from Tepperman (2009), the *Figure 9* represent the payment system of the industry, when pharmacies buy from wholesalers or manufacturers, in case of generic drugs, manufacturers could be provided rebates to stock the sellers' drugs. On the other hand, patients pay all or part for medicine which includes the markup and dispensing fees, and insurance will pay the residual part according to policies.

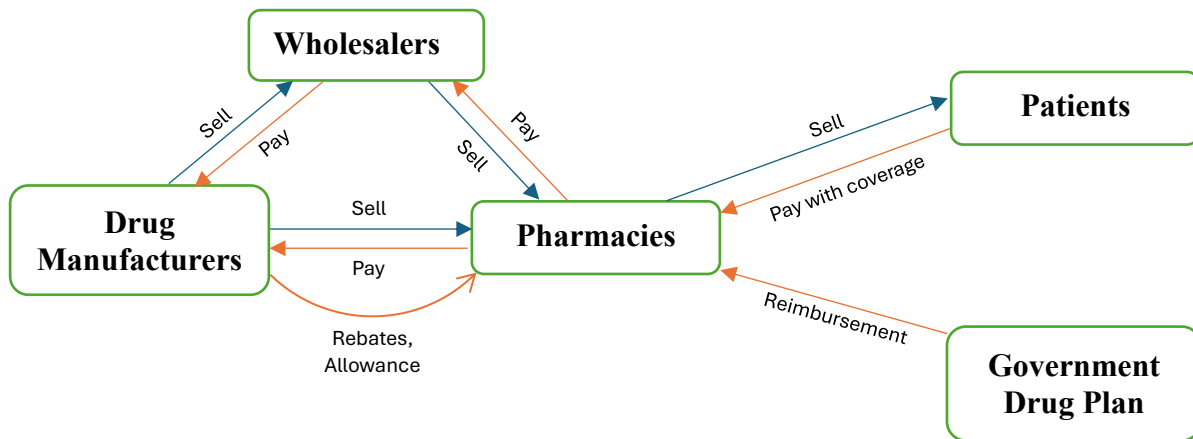


Figure 9: Payment system in pharmaceutical industry

Source: Adapted from Tepperman (2009)

Regarding business model in the pharmaceutical, companies could choose to perform the whole process from development to consumers which is vertically integrated pharmaceutical supply chain, or it could outsource part of production to Contract Manufacturing Organizations (CMOs) (which is more and more common) or call for specialized actors support at different stages like University at research stage, Contract Research Organizations at clinical trials stage...(Figure 10) (Capo et al., 2014).

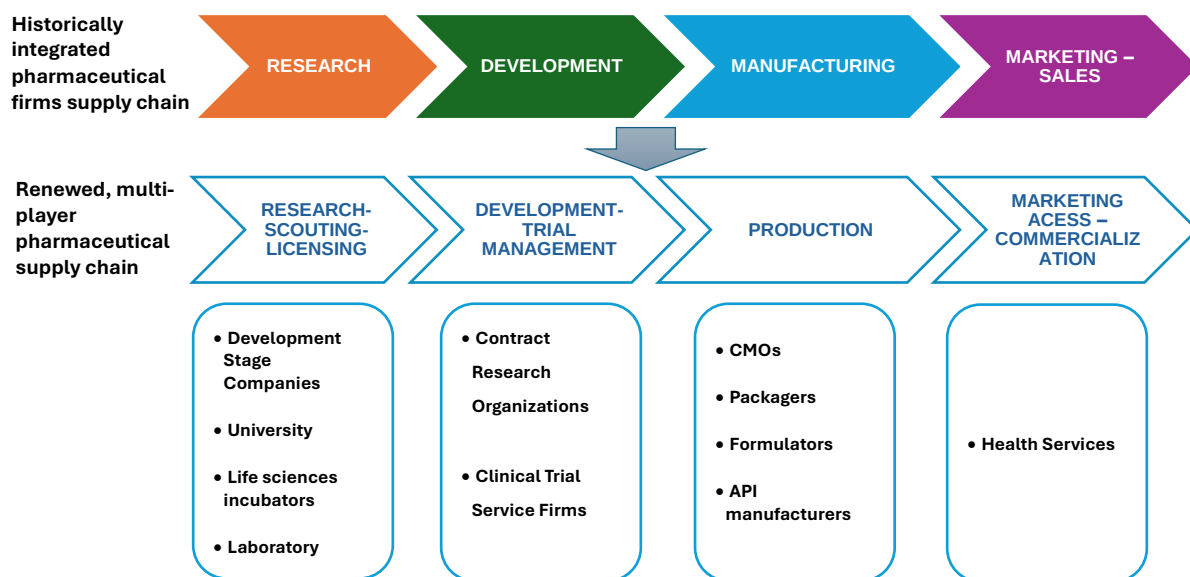


Figure 10: Pharmaceutical business model

Source: Capo et al. (2014)

In terms of portfolios of pharmaceutical industry, there are variety of products with different risk, cost, and return (Figure 11). The drug discoveries like New molecular entities (NMEs), Novel biologics (NBEs) carry high risk with potentially significant revenue and cost recovery. Conversely, the low-risk sectors like generics and super generics have potential as

they are rapidly growing but the competition in these is intense, and the margin is narrow too. Besides, life cycle management opportunities (drug combinations, new drug delivery options, ...) are crucial for optimizing return. Large companies also want to enlarge their portfolio with medium risk options like drug repositioning, orphan drugs and life extension opportunities (Khanna, 2012).

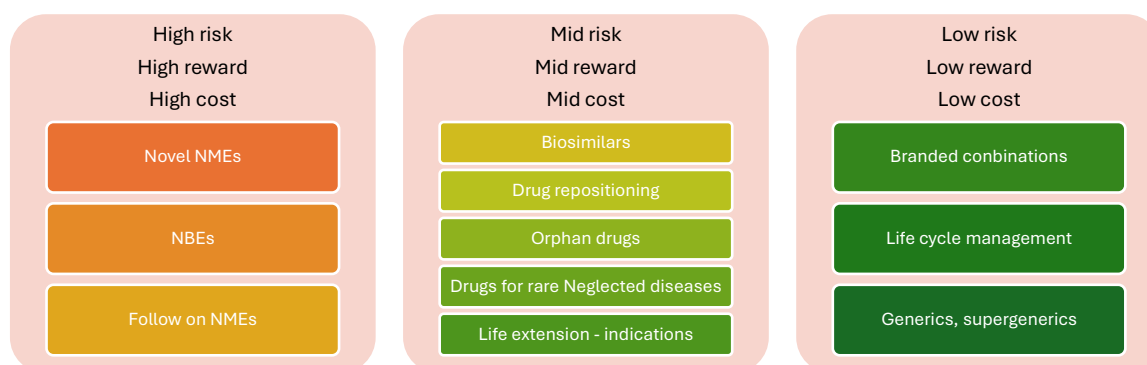


Figure 11: Risk balanced portfolio

Source: Khanna (2012)

2.4. Pharmaceutical industry of Hungary

The Hungarian pharmaceutical industry is very traditional with the first drug production company in mid-19th century but ceased due to competitors from Germany (Kerpel-Fronius, 2013). Significant manufacturers with different specialization in the sector in 20th century were Chinoin (now owned by Sanofi Avenis and specialized in in synthetic pharma production), Egis (specialized in nutrition products now owned by French Servier group), and Richter Gedeon (specialized in pharma production based on natural materials), these three are established between 1901-1913, later Biogal (now owned by Teva) is founded in 1950. After World War II, 4 mentioned companies plus Humán and Alkaloida become leading companies in the market (Antalóczy et al., 2019). As an important supplier to the former Soviet Union, Hungary's pharmaceutical industry played a significant role in the country's economy during the communist era. Becoming privatized in the 1990s, the local sector attracted substantial foreign investment and continued to play a key role in the economy (EIU, 2023; Antalóczy et al., 2019; Gulácsi et al., 2002, 2009; Kerpel-Fronius, 2013; Lőrinczy, 2013).

The current system of pricing and reimbursement for pharmaceutical products in Hungary is briefly represented in the *Figure 12*. According to the website of OGYÉI (former National Pharmaceutical and Food Health Institute before merged into National Center for Public Health and Pharmacy (NCPHP)) for pharmaceutical products to have marketing authorization, successfully going through assessment by professional authority is necessary. After classification, the National Health Insurance Fund (NEAK), advised by the Technology

Evaluation Committee (TÉB) will decide whether the products are reimbursable or not by a complex process and coordination of multiple departments. Finally, the product will go to the distribution process. The important thing in the pricing process is manufacturers have freedom to price their products while NEAK can only accept or negotiate the price (Österreichisches Bundesinstitut für Gesundheitswesen, 2006). On the other hand, on wholesale and pharmacy level, all pharmaceuticals' price is regulated by markup scheme (Österreichisches Bundesinstitut für Gesundheitswesen, 2006).

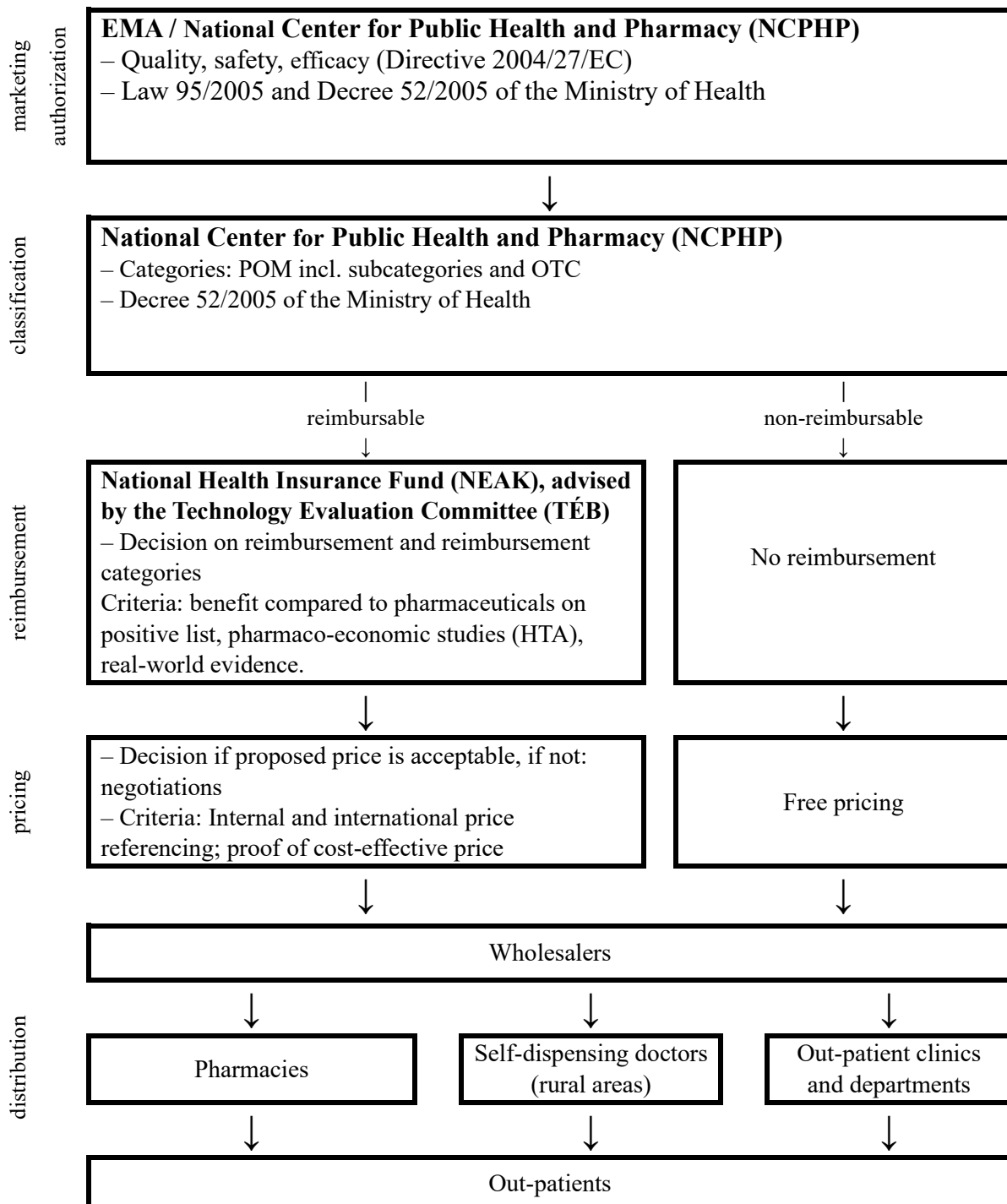


Figure 12: Hungary - pharmaceutical system

Source: Adapted from Österreichisches Bundesinstitut für Gesundheitswesen (2006)

Regarding current economic situations of industry from the data of Hungarian Central Statistical Office (2024b), the Share of pharmaceutical and medical goods in total retail sales value for internal trade in Hungary tend to decrease with 4.33% in 2023 and total value tend to increase with 783217 million HUF in 2023 which imply a shift in consumers spending to other products categories or the market is already saturated (*Figure 13*). The total value of production is 127130 million HUF in April 2024 and increase slightly overtime which show the growth of the industry, which could be driven aging population, health awareness or the pandemic Covid-19 might dramatically affect it (*Figure 14*) with total number of registered manufacturing pharmaceutical company of 124 in 2023, the number is stable over year which indicate a mature industry with decent barrier to entry (*Figure 15*). Additionally, there are also 17,559 people working in the industry with 2,343 pharmacies all over the country in 2023 (Hungarian Central Statistical Office, 2024e). In terms of external trade, there is a positive balance of payment in duration from Jan 2024 to April 2024 with import value of 818,032.7 million HUF and export value of 1,196,870.2 million HUF which is understandable as Hungary is in high position in pharmaceutical industry of local region (Hungarian Central Statistical Office, 2024c).

The effective strategies to continue develop in the industry in Hungary should be focusing on R&D as the market is saturated, competition is high with moderate barrier, or companies could extend their product portfolio to adapt to constant changing customer needs and for the government, they could keep the top position if they could attract foreign investment.

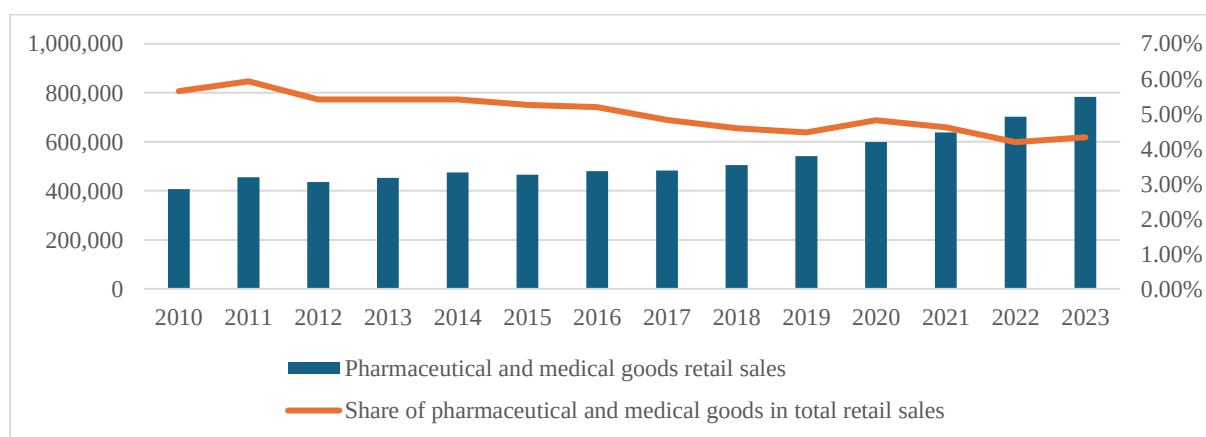


Figure 13: Value and share of pharmaceutical and medical goods in total retail sales value in Hungary from 2010 to 2023 in million HUF

Source: Adapted from Hungarian Central Statistical Office (2024b)

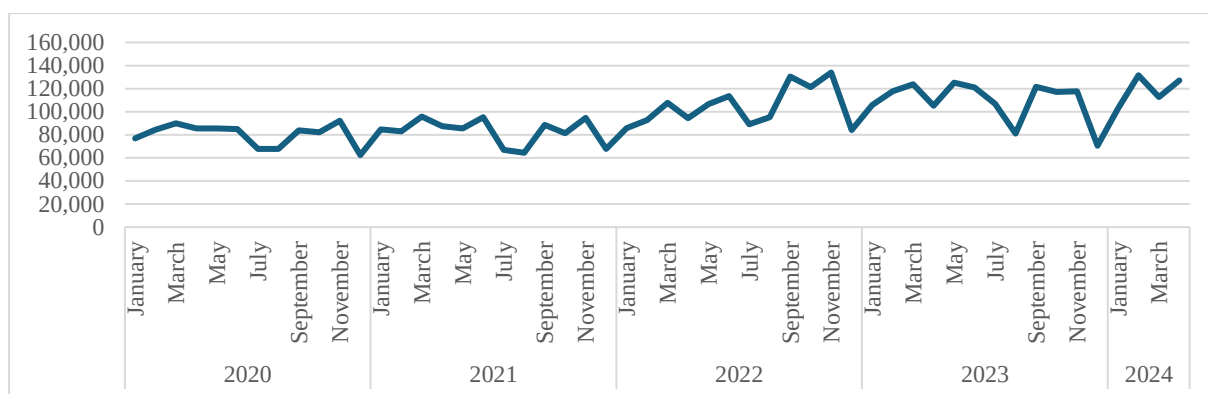


Figure 14: Value of manufacturing of pharmaceuticals, medicinal chemical and botanical products in million HUF

Source: Adapted from Hungarian Central Statistical Office (2024d)

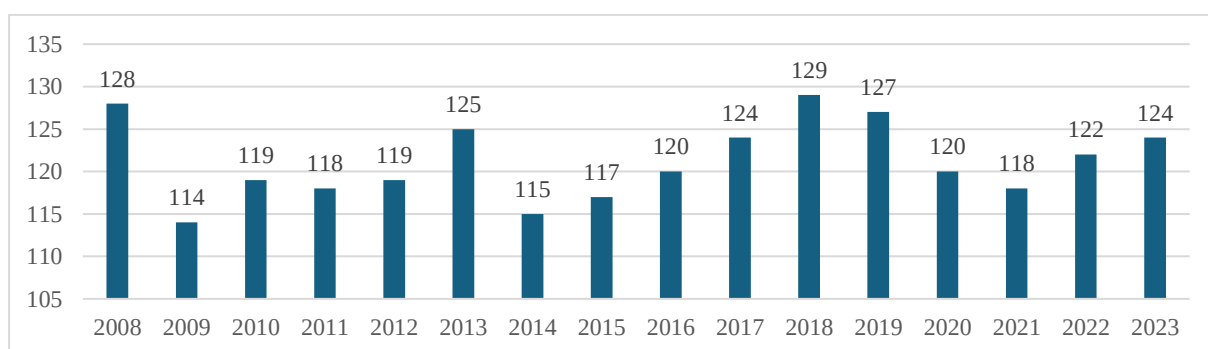


Figure 15: Number of registered pharmaceutical manufacturing companies in Hungary

Source: Adapted from Hungarian Central Statistical Office (2024a)

3. MATERIAL AND METHOD

3.1. Overview of company Gedeon Richter

Gedeon Richter is a public company with headquarter in Budapest Hungary, the company is established in 1901. There are two main business activities, the fundamental one is research and development, manufacturing and marketing of pharmaceutical products and the second one is wholesale and retail of these products. The other activities are providing support service for the other parts. The company has revenue of 805,158 million HUF in 2023 with EPS of 860 HUF and its total market capitalization is 1631 billion HUF. Around one third of the total share is domestically owned while the other is internationally owned. In the same year, the number of employees is 11603 people with around half is working in Hungary.

The Group has a network of subsidiaries, joint ventures and associate companies operating together to help the company to have a solid presence globally. There are 6 types of member companies in Gedeon Group: the headquarters in Hungary doing R&D, production, logistics and management of total sales activities, pharmaceutical subsidiaries and joint ventures producing pharmaceutical products, trading offices support sales and marketing in local regions, wholesalers and retailers which receive marketing support from parent company, service companies support headquarters activities and other units.

Gedeon Group has 4 business units: Neuropsychiatry (CNS), Women's Healthcare (WHC), Biotechnology (BIO) and General Medicines (GM).

1. Neuropsychiatry (CNS) is a unit hold one of the most valuable assets of company – Cariprazine which is a blockbuster developed by the company and sold globally, the aim of activities here is to maximize value of the drug through marketing and market access, besides company is actively researching and developing more drug to meet the medical need of the field.
2. Women's Healthcare (WHC) is the most important niche areas of operation of the group with a large product portfolio and strong foundation in Eastern and Central Europe, even in countries with no direct appearance, women there still be able to access the products, the strategy Gedeon Richter set for this business segment is to be leading provider and achieve significant profitability.
3. Biotechnology (BIO) is a unit independent from other as the source of knowledge is external and it comprises of management of biosimilar portfolio, revenue return from teriparatide biosimilar and biotechnology CDMO - Contract Development and Manufacturing services from Richter-Helm Biologics joint venture.

4. General Medicines (GM) is business unit manage the generic products of company, it is a new unit which created in 2023 focus on Eastern and Central Europe, the products are related to central nervous system and cardiometabolic treatments with aim to be a leading generic provider.

3.2. Methodology

3.2.1. Financial ratio analysis

To assess the financial health of Gedeon Richter over the period, intertemporal comparison is adopted. For intertemporal comparison, financial ratios will be compared overtime to find trend.

From the model of Alhasanko (2024), there are 4 areas that affect the financial health of companies which are: debt management, revenue, cash flow and expense. From those categories of financial ratios of Courtis (1978), 14 ratios is chosen which reflect the best of these categories and group them into 4 groups: Debt management, liquidity, profitability and Revenue – Cash flow. For Debt management, the debt to equity and interest coverage ratio is selected as they reflect the capital structure and how they manage the cost of debt:

$$\text{Debt to equity ratio} = \frac{\text{Total Liabilities}}{\text{Total Equity}} \quad \text{Interest coverage ratio} = \frac{\text{EBIT}}{\text{Interest expense}}$$

For profitability, 3 ratios: gross margin, net profit margin and operating margin are opted as they not only show the profitability, but they also show the expense for each level of profit difference.

$$\begin{aligned} \text{Gross margin} &= \frac{\text{Gross profit}}{\text{Revenue}} & \text{Operating margin} &= \frac{\text{EBIT}}{\text{Revenue}} \\ \text{Net margin} &= \frac{\text{Net profit}}{\text{Revenue}} \end{aligned}$$

For liquidity, this is an area added to the assessment of financial health as it shows how liquid the company is to manage the short-term debt, 3 ratios used here are cash ratio, current ratio and quick ratio.

$$\begin{aligned} \text{Cash ratio} &= \frac{\text{Cash} - \text{Cash equivalent}}{\text{Short-term liabilities}} & \text{Current ratio} &= \frac{\text{Current assets}}{\text{Short-term liabilities}} \\ \text{Quick ratio} &= \frac{\text{Current assets} - \text{inventory} - \text{Prepaid expense}}{\text{Short-term liabilities}} \end{aligned}$$

For revenue and cash flow, 5 indicators are selected which are sale growth, payable days, receivable days, inventory days, cash conversion cycle. These ratios reflect the credit policy of the company, and how it utilizes its working capital to generate revenue and cash flow.

$$\text{Sale growth} = \frac{\text{Current revenue}}{\text{Last year revenue}} \times 100\% \quad \text{Payable days} = \frac{\text{Account payable}}{\text{COGS}} \times 365$$

$$\text{Receivable days} = \frac{\text{Account receivable}}{\text{Sales}} \times 365 \quad \text{Inventory days} = \frac{\text{Inventory}}{\text{COGS}} \times 365$$

$$\text{Cash conversion cycle} = \text{Inventory days} + \text{Receivable days} - \text{Payable days}$$

Additionally, the aggregate indicator which will be used here is Altman Z-score. Even though the model has some limitations as the origin is from US companies' data so it will not reflect the country-specific differences, the model still works well in an international context. However, a country specific model is still more desirable if the bankruptcy prediction is the focus. But in case the bankruptcy prediction is not primary concern, the Altman Z-score model which has been tested across countries, also is very reliable and easy to interpret is recommended (Altman et al., 2014).

$$Z = 1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + X_5$$

Where:

$$X_1 = \frac{\text{Working Capital}}{\text{Total Assets}} \quad X_2 = \frac{\text{Retained Earnings}}{\text{Total Assets}} \quad X_3 = \frac{\text{EBIT}}{\text{Total Assets}}$$

$$X_4 = \frac{\text{Market Capitalization}}{\text{Total Liabilities}} \quad X_5 = \frac{\text{Sales}}{\text{Total Assets}}$$

In this research, that weakness would not significantly affect the analysis as the comparison to make is in the same company. The information to focus on is the change in the score which will reflect the shifts in financial health of Gedeon Richter, as the ratios used in the formula change when the financial health changes. The approach now focusses on internal financial health changing over time instead of the absolute thresholds for bankruptcy or financial health.

3.2.2. Textual analysis of annual reports

For textual analysis, the quantitative textual analysis approach is employed as it will minimize subjective. Both quantitative content analysis which utilizes the Loughran-McDonald Master Dictionary and text mining which use R programming languages to process text is applied into the research with the general flowchart in *Figure 16*.

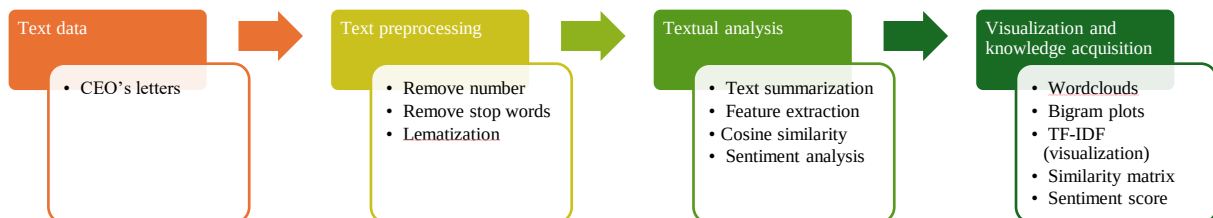


Figure 16: Flowchart of textual analysis process

Source: Adapted from Luo et al (2021)

The first step is choosing the text data. There are several text bodies that have been used in textual analysis research from different entities, particularly in firm level it can be letter of shareholders, annual or quarterly reports, 8-Ks, Press releases, Conference calls, firms' websites, twitters posts (Crowley, 2019, p. 7). In this research, the MD&A will be used as others type of text are not officially published and less stable year by year. This part is inside the Management reports of annual reports of Gedeon Richer.

As the aim is to measure the financial health of Gedeon Richter, not all the text in MD&A part of the financial report will be collected. However, they serve different content and purpose, so if analyzing the whole might make the results not good as not all the text will support financial health and for sentiment analysis, not all the text gives subjective information. As a result, only the relevant textual data and method will be chosen with reference to prior textual analysis in accounting research. In this case, the CEO's letters will be analyzed.

Then text preprocessing is performed by remove number, stop words which contribute little to the content like "do", "from", "get", and then lemmatize words in here is to bring words with inflected forms to their original forms so they could be analyze as one.

The third and fourth step are to perform textual analysis, create visualization and knowledge acquisition. summarization of text is used to see the main ideas of the text with the word clouds to show words with high frequency (number of appearance) and then bigram plots which treat words appear next to each other as a pair to show the text structure and relationship of words is applied. The most common co-occurring word pairs (does not mean they appear the most, but they appear together relative to appear separately) will appear in the graph, the phi coefficient ≥ 0.5 is set. The phi coefficient is used to measure the associate of binary variables and in this study, correlation among words is examined in sentence level (Yan, 2022).

$$\phi_{X,Y} = \frac{n_{11}n_{00} - n_{10}n_{01}}{\sqrt{n_{.1}n_{.0}n_{1.}n_{0.}}}$$

Where:

Table 1: Value in phi coefficient formula

	Has word Y	No word Y	Total
Has word X	n_{11}	n_{10}	$n_{1.}$
No word X	n_{01}	n_{00}	$n_{0.}$
Total	$n_{.1}$	$n_{.0}$	n

Source: Yan (2022)

The TF-IDF (Term Frequency-Inverse Document Frequency) is used to extract the distinctive features of the texts by words with high TF-IDF weight, it is calculated as the product

of TF (term frequency) and IDF (inverse document frequency) (Yan, 2022). Here it is used to show the specific information of the company in that year which might have strong effect or explain the change in the risks, performance or positions of the company that direct or indirectly affect the financial health.

$$tf - idf(term) = \frac{n_{term \text{ in document}}}{n_{total \text{ word in document}}} \times \log \frac{n_{total \text{ documents}}}{n_{documents \text{ contain term}}}$$

Cosine similarity is a measurement of the similarity of 2 non-zero vectors, and by converting 2 documents into vectors, the similarity of the documents could be calculated. Besides removing stop words, using tf-idf weight is to further reduce the effect of common-used words. The results of cosine similarity of documents each year will be cosine similarity matrix (Anand et al., 2020).

$$\cos(u, v) = \frac{u \cdot v}{|u||v|}$$

For sentiment analysis, the dictionary-based techniques with the help of Loughran-McDonald Master Dictionary will be utilized, the dictionary has lists of words categories: "negative", "positive", "litigious", "uncertainty", "constraining", "strong modal" and "weak modal". By counting words belong to which categories, the tone different between years could be revealed, and the sentiment score could be calculated to assess the positive of the tone (Anand et al., 2020).

$$\text{Tone}(\%) = 100\% \times \frac{n_{positive \text{ words}} - n_{negative \text{ words}}}{n_{total \text{ words}}}$$

For technology to process text data, R programming is used to analyze and visualize as it has a lot of libraries that support textual analysis and visualization like: tidytext, wordcloud, tm, ggraph,...

3.2.3. Data collection

As the need of data in financial ratios analysis and textual analysis all come from the financial reports so the financial reports of Gedeon Richter from 2019 to 2023 will be collected from the company websites. The reason to choose to analyze information from this period is because there are a lot of crises happen like Covid-19, Russian-Ukrainian war which will reflect in the trend analysis and give good insight of the company's financial health and how the methods show the financial health.

From the reports, data of financial data will be collected from financial statements in consolidated financial statements part of the annual reports: consolidated income statements, consolidated statements of comprehensive income, consolidated balance sheets, consolidated cash-flow statement and consolidated statement of change in equity. Text data will be collected

in the Management Report parts of financial reports where contain chief executive officer's reviews (CEO's letters).

3.2.4. Analytical framework

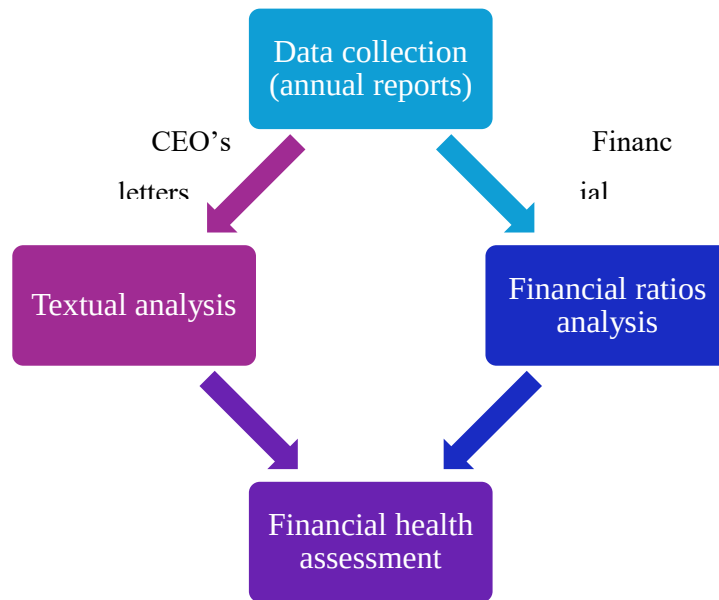


Figure 17: **Analytical framework**

Source: Author (2024)

The analytical framework (*Figure 17*) will be the general roadmap of analyzing financial health of Gedeon Richer, after using both methods: Textual analysis and financial ratios analysis, the comparison and integration of 2 methods will be made in the end to see if they give better assessment of company financial health.

4. FINDINGS AND THEIR ANALYSIS

4.1. Financial health analysis using ratios

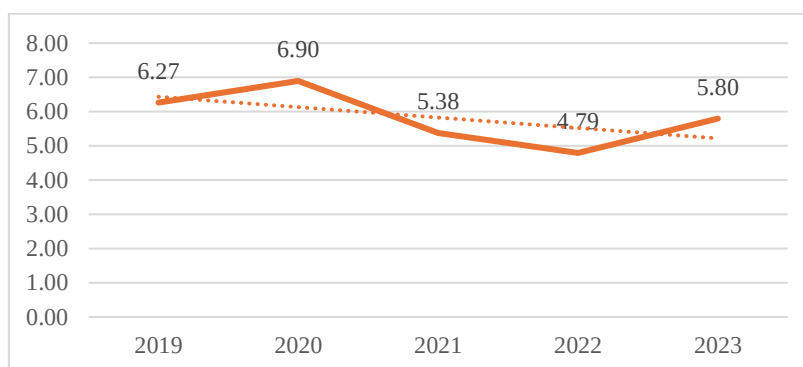


Figure 18: Altman Z-score of Gedeon Richter 2019-2023

Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)

The Altman Z-score of Gedeon Richter seems to decrease over time even though they are all on the safety threshold, which is more than 3.0, with the highest score in 2020 (6.9) and lowest in 2022 (4.79) and increase back in 2023 (5.8). It shows that there is some financial change which affect in 2021 as there is a huge decrease from 2020 and 2021 (6.9 to 5.38). Interestingly, in 2020 when in middle of the pandemic company financial health score is good which is even better than 2019.

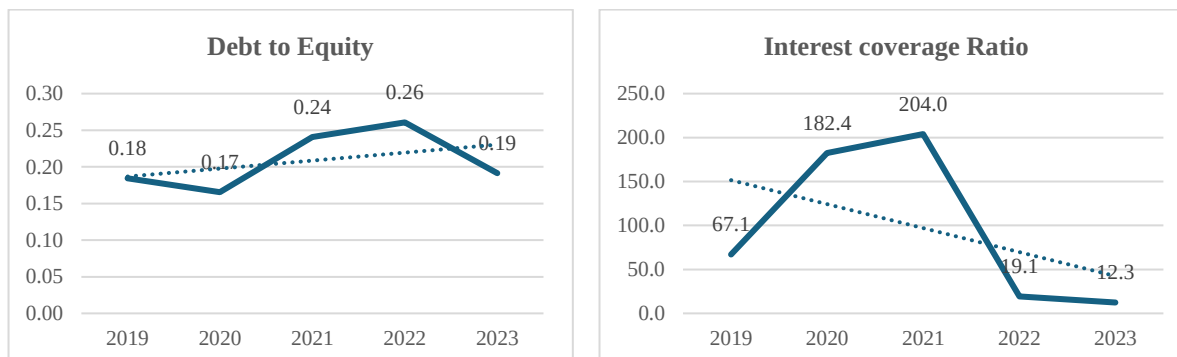


Figure 19: Debt management related ratios Gedeon Richter 2019-2023

Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)

On the debt management ratios, the debt-to-equity ratio showed an increase in debt load and the interest coverage ratios diminished. From 2019 to 2020, there is little change in debt to equity (2020: 0.17, 2019: 0.18) but there is a rising in interest coverage ratio which derive from the significant increase in the operating profit (EBIT). In 2021 and 2022 there is a positive shift in debt ratio (0.17 to 0.24) which imply that there is more debt in this period, which is validated by the issuance of bond amount to 70 billion HUF, and it will affect the Altman Z-score, despite the CEO's assertion that it is funding in favorable situation. which led to the decrease in the interest coverage in the subsequent years. The debt is reduced in 2023 with the debt to equity

(0.26 to 0.19). However, the interest coverage ratio in the same year is not improved but is worsen (19.1 to 12.3) because the non-current liabilities is not decrease but the change is largely in current liabilities which is from the liabilities directly associated with assets classified as held for sale when Gedeon Richter made the decision to sell the Romanian subsidiaries to Dr. Max Group in 2022 amount to 49044 million HUF. Moreover, the debt on issue of bonds is increase in 2023 lead to the increase in interest expense. In general, the debt of Gedeon Richter is low and most of the funding is from equity, which is not surprising that the interest coverage ratios are also high even some change cause significant movement in the ratios.

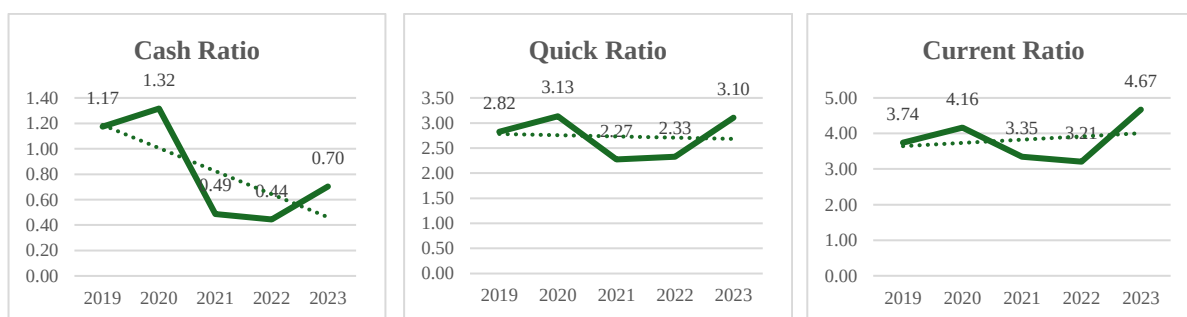


Figure 20: Liquidity ratios Gedeon Richter 2019-2023

Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)

Regarding the liquidity ratios, the fluctuation is exhibited from 2019 to 2023. There is an improve in in current assets, quick assets and cash from 2019 to 2020, on the contrary there is a big drop in 3 types of assets in 2021 and 2022 (cash ratio: 1.32 to 0.49, quick ratio: 3.13 to 2.27, current ratio: 4.16 to 3.35) but they rebounded in 2023 (cash ratio: 0.44 to 0.7, quick ratio: 2.33 to 3.1, current ratio: 3.21 to 4.67). In 2020, there is a surge in cash and cash equivalent from the operation income, also quick and current assets increase from built up inventories to diminishing supply risk by the pandemic and asset hold for sale from divesture of some none core wholesales and retails businesses with little change in current liabilities. In 2021, the company decide to invest its free cash into government securities and corporate bonds, make it into non-current assets which make the liquidity ratios decrease, attribute to the lessen of the ratio is the purchasing of EVRA right, which cost 73198 million HUF. In 2022, there is an increase in inventories due to the development of specialty pharma (EVRA, DROVELIS, RYEQO), the transformation causes inventory more than 24% compared to last year. In the same year, Gedeon Richter paid excess income tax and a huge dividend amount to 42146 million HUF approximately the same as 2021. In 2023, the current liabilities are decrease significantly due to reduce of the liabilities directly associated with assets classified as held for sale (49,044 million HUF) from the Romanian subsidiaries divesture and settlement of windfall tax in 2022, while the assets classified as held for sale in current assets also gone (67,014 million

HUF) but the total current assets is not change much compared to last year because of the increase in inventories (2023: 177,767 million HUF, 2022: 153,335 million HUF) and trade receivables (2023: 204,968 million HUF, 2022: 175,182 million HUF). The inventories and trade receivable grew partially because of high sale levels and partially because of the recover stock in RUB from the Russian-Ukrainian war.

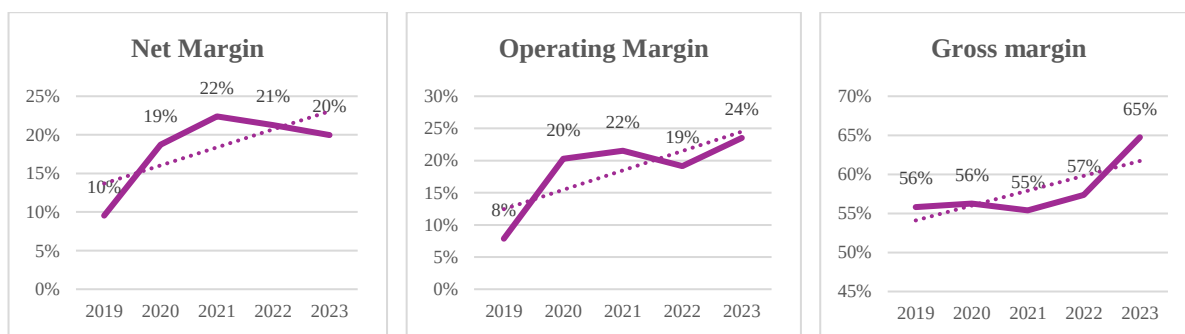


Figure 21: Profitability ratios Gedeon Richter 2019-2023

Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)

About profitability, the ratios indicate an increase in all three gross profit margin, net profit margin and operating margin. For gross profit margin, it changes the most from 2022 to 2023 (57% to 65%) which driven by the growth in Pharma segment sales from which are noticeable WHC products like Drovelis and Ryeqo alongside the growth of Vraylar on contrary there is a drop from divestiture of Romanian operations. For operating margin and net margin, they dramatically change in 2019 to 2020 as operating margin changes from 8% to 20% and net margin changes from 10% to 19%. That indicates a decrease in the expenses of sales and marketing as the promotion activities is affected by the pandemic, besides the number of staff is reduce because of the adverse situation. Also, in the note, Gedeon Richter emphasize the change in other income and expense which contribute to the increase profit, the income from licensing Cariprazine to South Korean partner and tocilizumab to Mochida. On the other hands, the net income of 2023 is not that increase compared to 2022 like its gross income, the main reason is from the income tax and income interest, for the income tax it is increasing by the significant decrease of deferred tax asset as it has been used to reduce tax in 2022, while the income interest is increasing from the increase in debt.

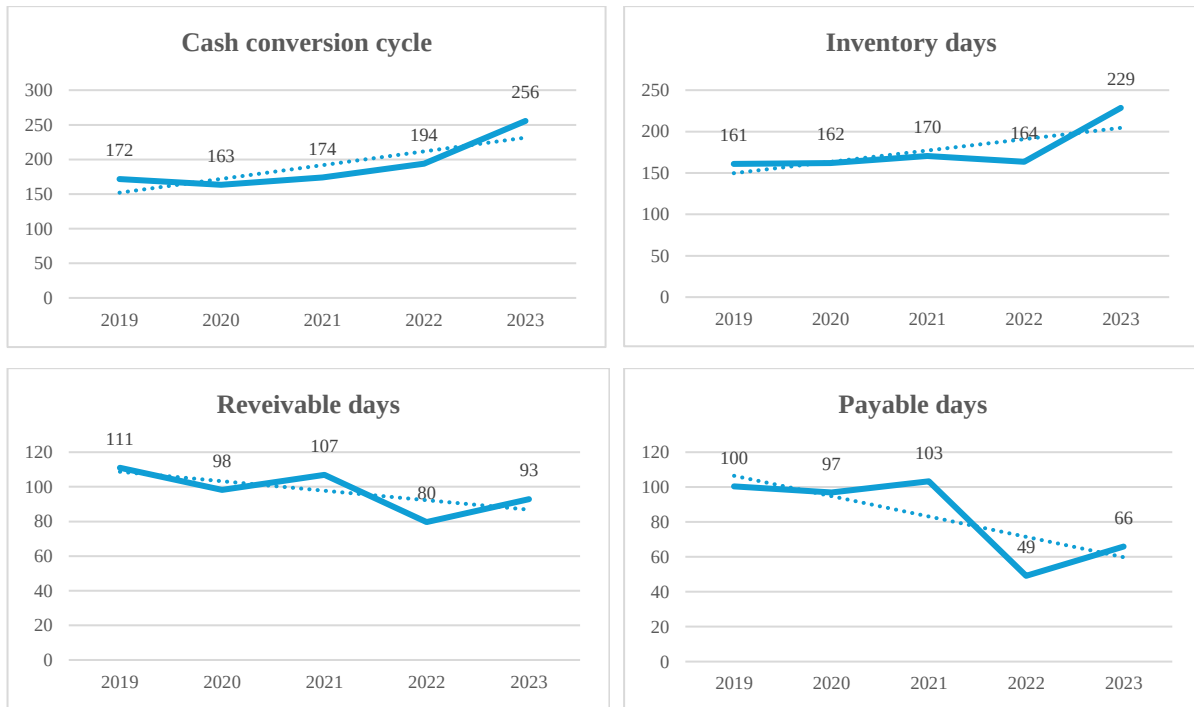


Figure 22: Cash flow related ratios Gedeon Richter 2019-2023

Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)

From cash flow related ratios and sale growth, there is an increase trend in cash conversion cycle, the inventory days increase the most in 2023 (164 days to 229 days), the payable days decrease the most in 2022 (103 days to 49 days), the receivable days seems to fluctuate with notable decrease in 2022 (107 days to 80 days). The decrease in trade payable and trade receivable days in 2022 is due to the classification of Romanian Wholesale and retail business of the Group receivable and payable to asset and liability held for sales. Conversely, the inventories in 2022 did increase but Gedeon Richter had 27% in sales growth which drove inventory day down.

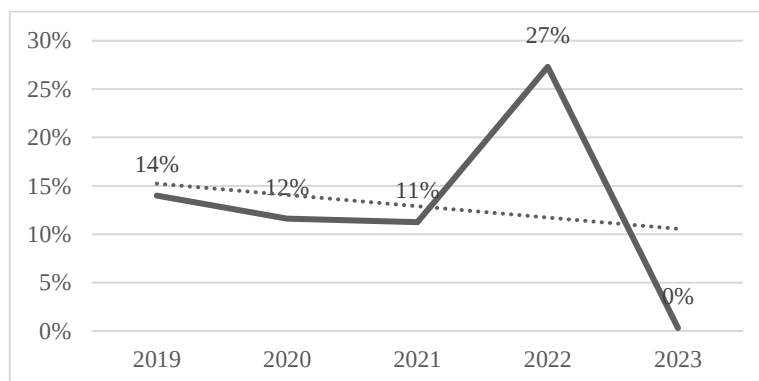


Figure 23: Sale growth of Gedeon Richter 2019-2023

Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)

Regarding sales performance, the company's revenue primarily comes from the sale of pharmaceuticals, revenue from service and royalty income. While the overall trend of sale growth is positive through 5 years, there is a slowdown in 2023 with the sales growth near 0%. The highest growth rate is 2022 with 27%. The sale growth is contributed mostly by the sales of pharmaceuticals and royalty, especially from royalty of VRAYLAR to AbbVie throughout 5 years.

Finally, from calculation of correlation of all ratios to Altman Z-score, the ratios moving in the same direction or opposite with the score will be assessed. The results show that the ones that have strong positive correlations are liquidity ratios (quick and cash ratios: 0.81 and 0.95) while debt management ratios (debt to equity) have strong negative correlation (-0.94). Other ratios are of medium to weak correlation especially inventory day and gross profit margin which mostly have no correlations.

Table 2: Correlation of ratios with Altman Z-score

Variable	Correlation with Altman Z-score
Sales growth	-0.45
Receivable days	0.52
Quick ratio	0.81
Payable days	0.63
Operating margin	-0.26
Net profit margin	-0.47
Inventory days	-0.08
Interest coverage	0.39
Gross profit margin	-0.06
Debt to equity	-0.94
Current ratio	0.59
Cash ratio	0.95
Cash conversion cycle	-0.29

Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)

4.2. Textual analysis results

For textual analysis, first word clouds will show the general content of CEO's letters from 2019 to 2023, and the main themes thought the representation of words with high frequency. From *Figure 24*, there are words which appear with high frequency like “product”, “market”, “sale”, “company”, “agreement”, “treatment” ... which shows the focus on the products of

Gedeon Richter. Besides, a significant change in 2023 compared to other years is observed, in this year the words like “growth”, “revenue”, “increase” which emphasize the grow instead of focusing on products like before. Outsidess of the general high frequency, the variety of other less common words show variety of topics like “strategy”, “organizational”, “global”, “progress”, ...



Figure 24: Word clouds of CEO's letters from 2019-2023

Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)

To have a better understanding of the main content of CEO's letters from 2019 to 2023, the bigrams will show the combination of 2 adjacent terms with the phi correlations is set here more than or equal 0.5.

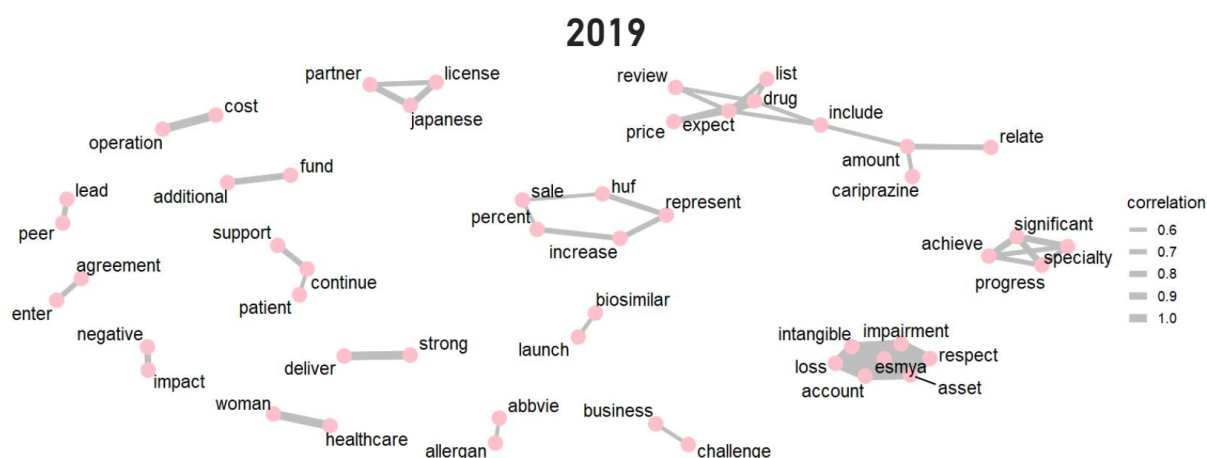


Figure 25: **Bigram of CEO's letter Gedeon Richter in 2019**

Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)

In 2019, on the bigram (*Figure 25*), the operation progressing is demonstrated like “enter agreement”, “continue support”, “Japanese partner”, “strong deliver”, “launch biosimilar”, “significant progress” On the other hand, the financial side of the company in 2019 is highlighted with “impairment”, “loss” of “intangible asset”, “operation cost”, “sale increase”, “drug price”, “additional fund”. Even with the increase in sales, the cost, impairment and drug price are also noticeable with high correlation. The negative sentiment combination is also detected like “negative impact”, “business challenge”.

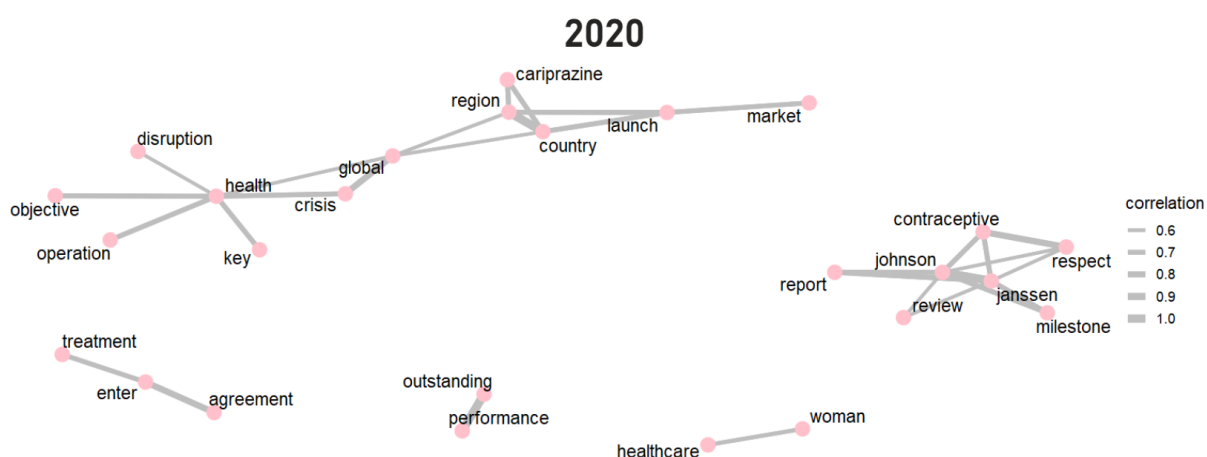


Figure 26: **Bigram of CEO's letter Gedeon Richter in 2020**

Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)

increase” in 2021 and “profit increase” in 2022. Especially, in 2022, the bigram graph shows the event of Russian-Ukraine war by connections of “military”, “conflict”, “February”, and in 2021 is the “anniversary” of 150th birthday of the company.

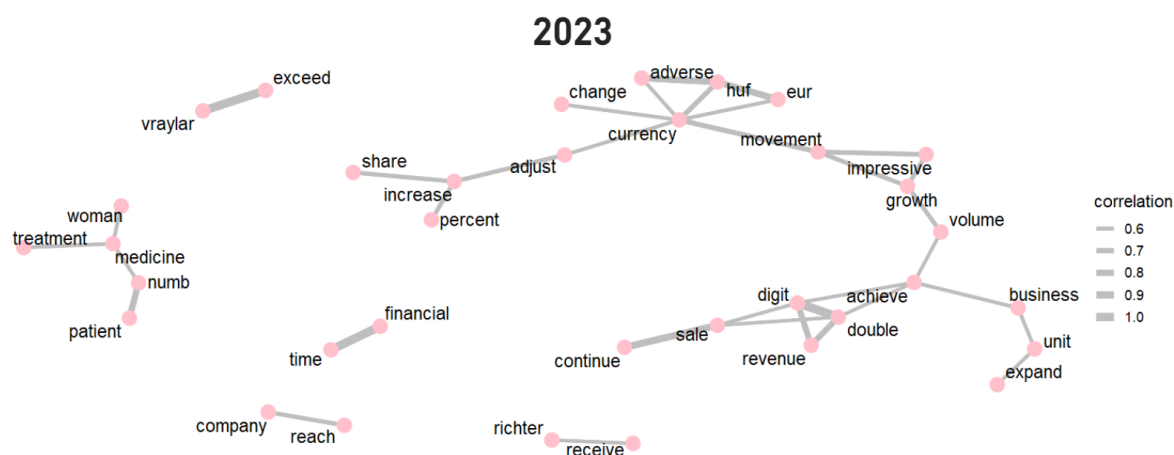


Figure 29: **Bigram of CEO's letter Gedeon Richter in 2023**

Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)

In 2023, the growth and achievement of Gedeon Richter is shown in its bigrams (*Figure 29*) with pairs like “percent increase”, “vraylar exceed”, “company reach”, “impressive growth”, “double digit”, “unit expand”. In financial point of view, the double-digit revenue is positive to Gedeon Richter financial performance, and the overall operation is very promising in this period. However, there is “currency change” related to HUF and EUR is “adverse” to the company. The bigram graph in 2023 is less dense than 2021 and 2022 bigram plots which show more varied usage of language, besides its focus of topics is changed to growth and achievement which show the same with the representation of the word clouds.

From the TF-IDF graph (*Figure 30*), the differentiate information between the years with words with high TF-IDF score is illustrated. In 2019, the words with highest IF-IDF score are “negative”, “esmya”, “authority”, “level” and “allergan”. The words imply suspension of ESMYA related to the case of liver injury after taking the medicine, the event needs the involvement of legal entity and cause an impairment to the ESMYA intangible assets. Besides, “allergan” as one of partner to support VRAYLAR product, now acquired by AbbVie so the relationship needed to be assessed. In 2020, “disruption” is the words with highest TF-IDF score and on the fourth place is “crisis” which is understandable due to the pandemic. “Johson” and “janssen” refers to 2 organizational entities: Johnson – Johnson and its subsidiary Janssen Pharmaceutica NV which sells Gedeon its product assets which is a significant deal. In 2021, “programme” is the highest TF-IDF score indicate the programs initiate in 2021 on R&D, leadership, operation, structural change, other words are also related to the assertion of

operational process and goal like “ensure”, “meet”, “solid”, “excellent”, ... In 2022, the attentions are on organization with words like “team”, “culture”, “department”, on product and operation like “therapy”, “Hormone”, “preclinical”, on external market situation like “military”, “conflict”, and on financial performance which is “tax” with second highest TF-IDF score. In this year, “tax” is directly related to financial health due to the extra profit tax. In 2023, “currency” is the word with highest score, which from the bigram (Figure 29), imply the “adverse” in exchange rate of EUR and HUF, the “double” and “digit” is also top words in TF-IDF score, refer to the record increase in revenue from company products.

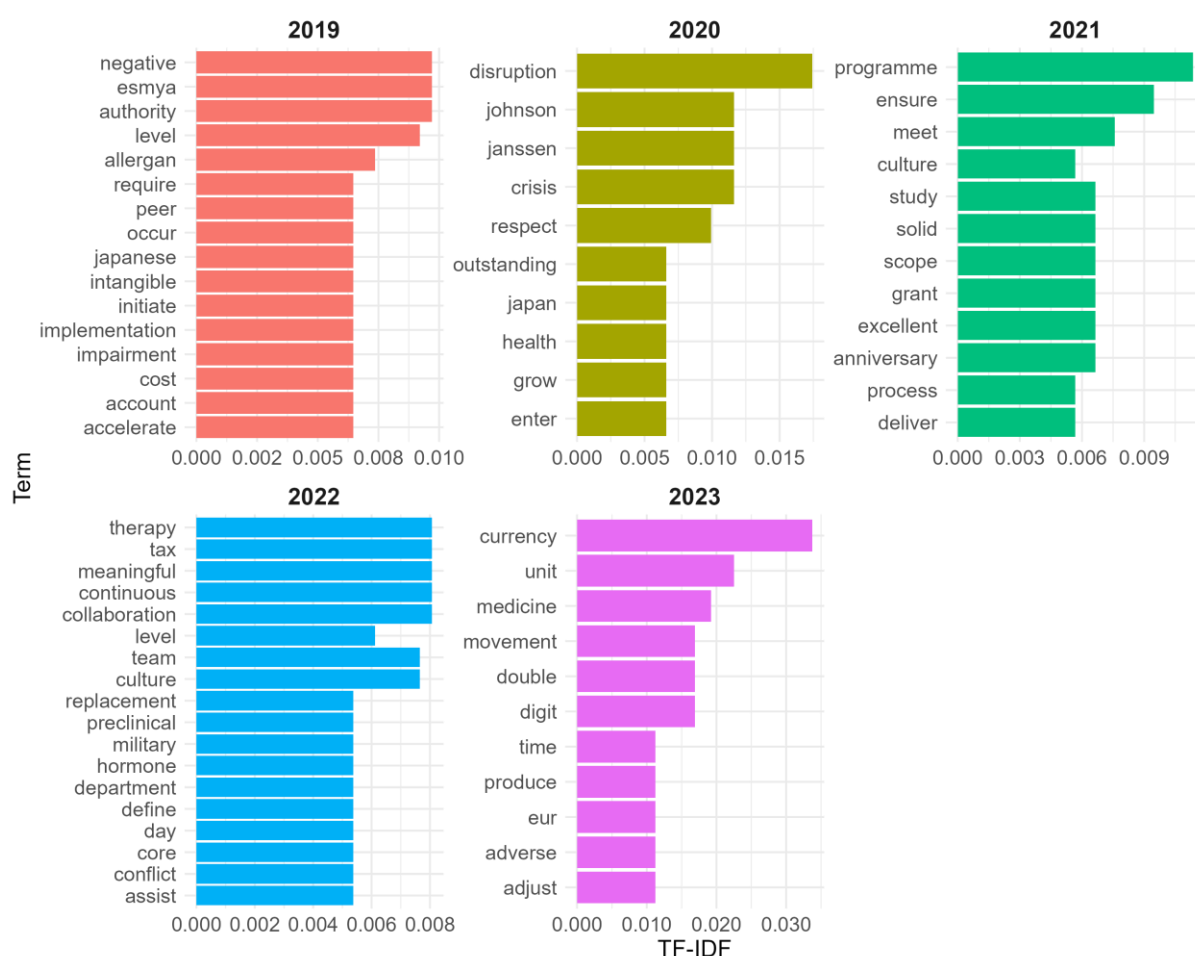


Figure 30: TF-IDF of Gedeon Richter CEO's letters 2019-2023

Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)

From the summarization and feature extraction of the 5 years CEO's letters, the changes in topic, word variety have been shown on word clouds, bigram plots and TF-IDF graphs. The cosine similarity matrix (Table 3) shows the quantify the similarity of document through the similarity score of the CEO's letters with each other. The similarity score is low (0.05-0.16) from year to year, which means the difference between word use in each year are very different. On the one hand, the content of CEO's letters would show the highlight information of each

year, not the general information so it would be highly distinctive from each other. Moreover, the complexity of the company and market could bring about changes both internal and external situation of the company, in the case of Gedeon Richter in pharmaceutical industry is not an exception. On the other hand, the changing in word use could also affect the similarity score without the need of changing the meaning.

From the score, the similarities between 2019-2020 and 2020-2021 are both 0.1, 2021-2022 is slightly more similar with score of 0.16. The most noticeable is year 2023 with a similar score is low with all other year which indicates a significant change in word use and topic of attention, which is validated by the word clouds.

Table 3: Cosine similarity matrix of Gedeon Richter CEO's letters 2019-2023

Year	2019	2020	2021	2022	2023
2019	1.00	0.10	0.10	0.12	0.04
2020	0.10	1.00	0.10	0.09	0.02
2021	0.10	0.10	1.00	0.16	0.04
2022	0.12	0.09	0.16	1.00	0.05
2023	0.04	0.02	0.04	0.05	1.00

Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)

To mine the tone of CEO's letters, the word count could be analyzed according to the Loughran-McDonald Master Dictionary (*Table 4*), with 6 categories: constraining, litigious, negative, positive, strong modal, uncertainty and weak modal.

Table 4: Sentiment word count in Gedeon Richter CEO's letters 2019-2023

Year	Constraining	Litigious	Negative	Positive	Strong modal	Uncertainty	Weak modal
2019	7	2	18	21	1	8	3
2020	3	0	11	11	0	0	0
2021	2	0	6	32	8	2	2
2022	2	2	13	31	4	2	1
2023	0	0	8	29	2	2	1

Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)

From the Table 4, in constraining category, 2019 have the highest word count (7) compared to other years, from the “impairment”, “impair”, “require”, “restrict” which come

from the suspension of ESMYA with authority involvement which is mentioned in DF-IDF analysis.

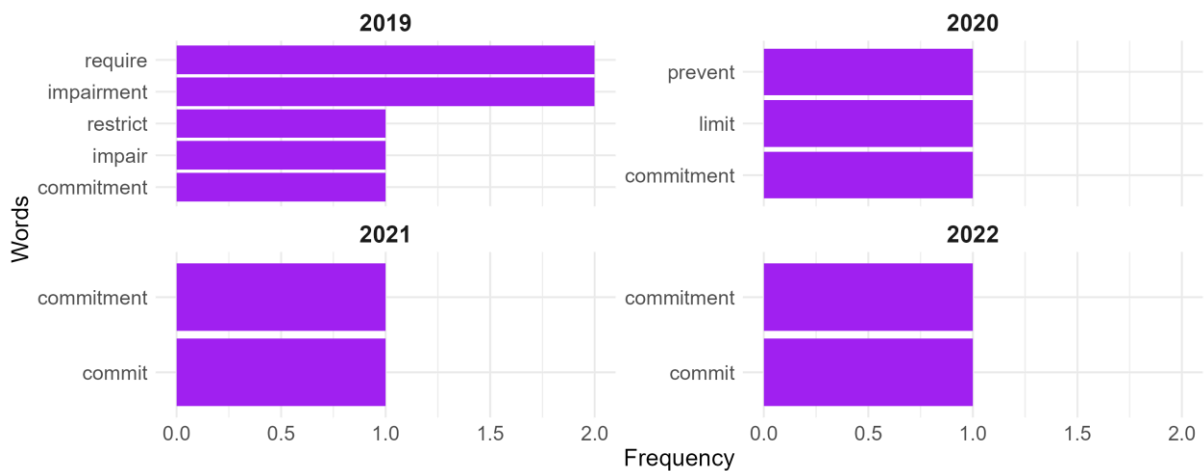


Figure 31: Gedeon Richter CEO's letters constraining sentiment words 2019-2023
Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)
In litigious category only 2019 and 2022 have words, while 2019 is due to the ESMYA suspension and Allergan acquisition process, in 2022, the “decree” refers to government regulation lead to the increase of tax and process treatment application of RYEQO product.

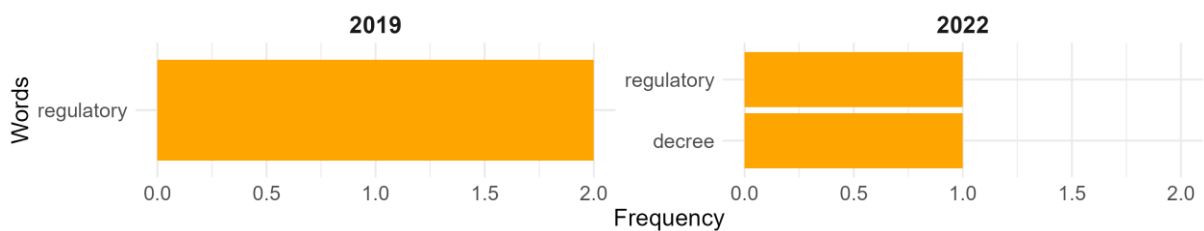


Figure 32: Gedeon Richter CEO's letters litigious sentiment words 2019-2023
Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)
The two categories: negative and positive is used to calculate the tone (*Figure 33*), with the increasing trend, the sentiment is low in 2019 (0.6%) and 2020 (0.0%) then increasing substantially in 2021 (5.2%), then slightly drop in 2022 (3.6%) and increase back in 2023 (4.2%).

In positive category, in all 3 years 2021, 2022 and 2023 have high number of word count while in 2019 (18) and 2020 (11) the number of words is a lot less, the words in the category imply the excellent in operation and achievement with words like “achieve”, “good”, “progress”...In negative category, 2019 have the highest words count (18) with words related to the operation like “impairment”, “loss”, “termination”, “suspense”, in 2021 and 2023 the negative word count is low. In 2022 and 2020, the number of negative words is subsequently 13 and 11, also the having negative sentiments, the high negative in these years is affected by

the external factors in the market like the Covid-19 and Russian-Ukrainian war with words like “crisis”, “conflict”, “disruption”. Overall, the sentiment shows a positive expectation of the future despite the fluctuation of the sentiment score. The positive words confirm the good internal operation of Gedeon Richter and the negative words show that the negative is mostly from external factors in the market which is uncontrollable.

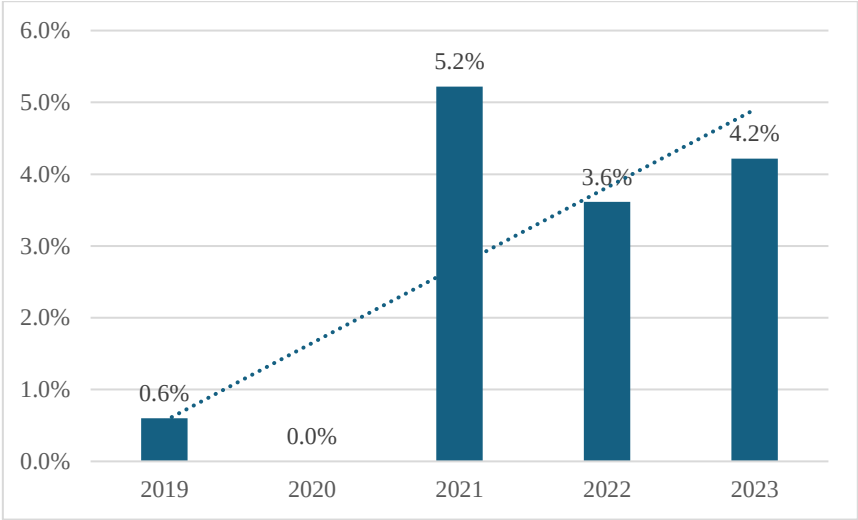


Figure 33: Gedeon Richter CEO's letters tone (%) 2019-2023

Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)

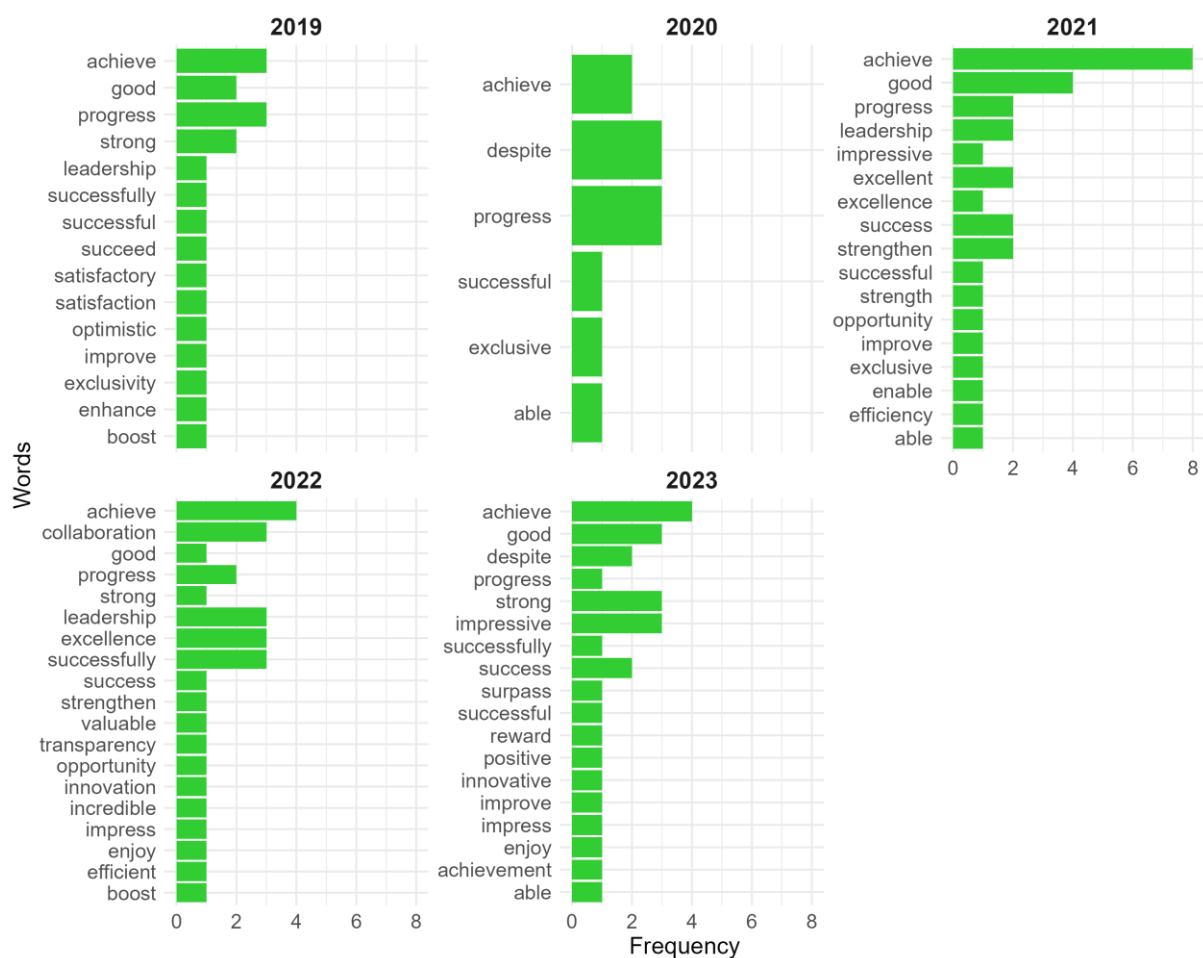


Figure 34: Gedeon Richter CEO's letters positive sentiment words 2019-2023

Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)

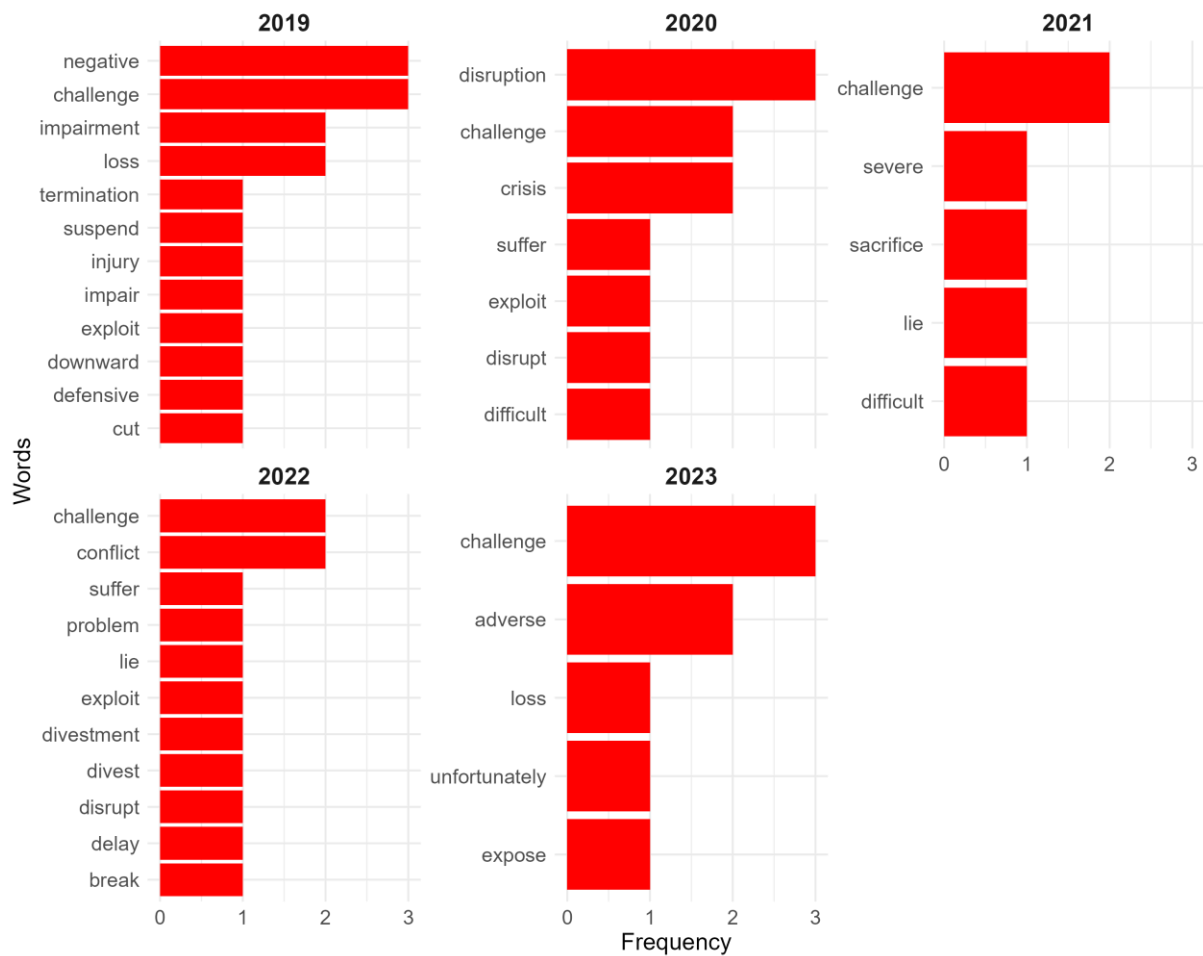


Figure 35: Gedeon Richter CEO's letters negative sentiment words 2019-2023

Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)

Other categories like uncertainty, strong modal and weak modal, in 2019 the number of uncertainty words (8) (*Figure 36*) are more than other year, in 2021, the words count in strong modal (8) (*Figure 37*) are higher than other year, while the number of weak modals (*Figure 38*) is not significantly many or different in any year. In short, there is an uncertainty tone in 2019, which is understandable by the situation in the year and high number of negative words, while there is a confident tone in 2022 which is strengthened by the high sentiment score.

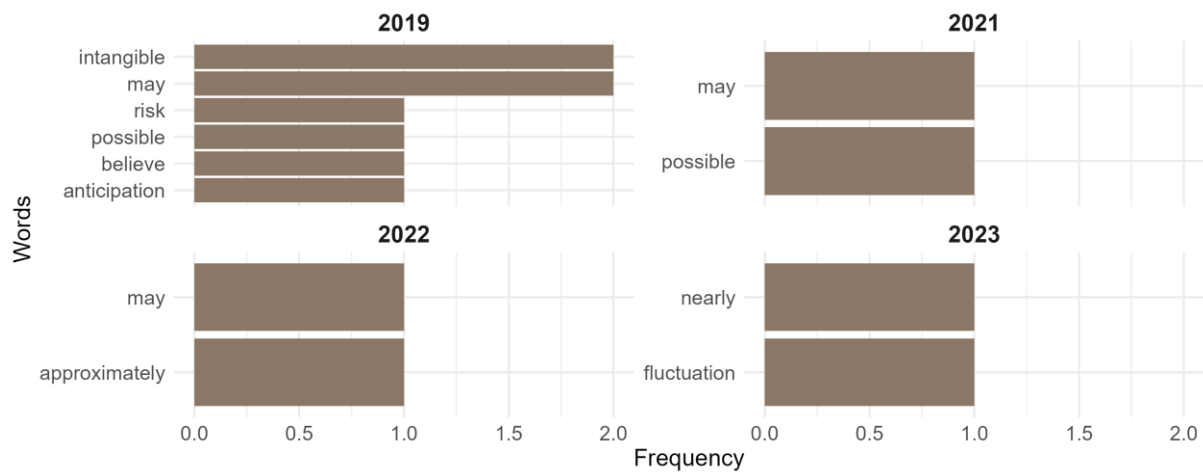


Figure 36: Gedeon Richter CEO's letters uncertainty sentiment words 2019-2023

Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)

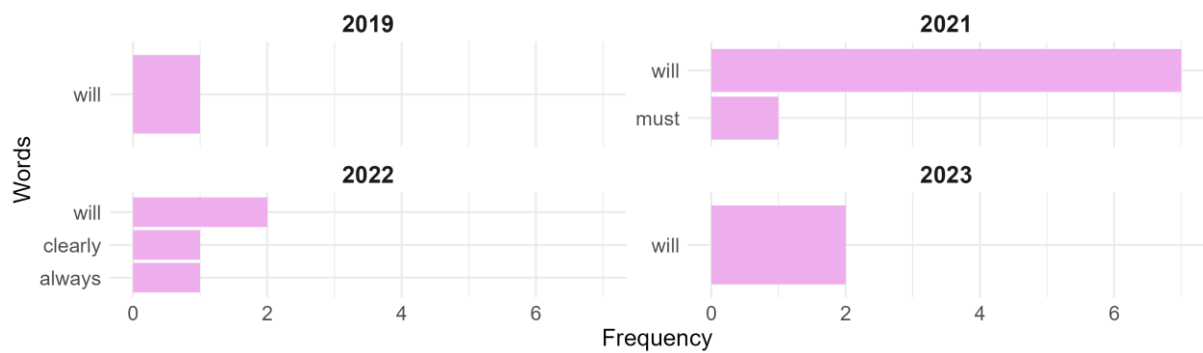


Figure 37: Gedeon Richter CEO's letters strong modal words 2019-2023

Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)

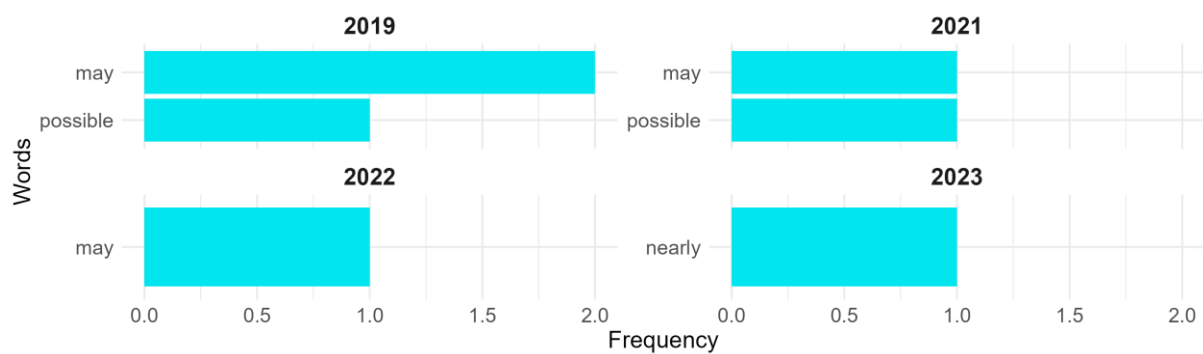


Figure 38: Gedeon Richter CEO's letters weak modal words 2019-2023

Source: Own calculation based on data from Gedeon Richter's annual reports (2019–2023)

4.3. Integration of financial ratios and textual findings

Due to the fundamental differences in the origins of data sources, textual analysis results and financial ratios analysis methods should not be expected to be correlated with each other directly. For textual analysis, the data is from unstructured and subjective data which is the text of CEO's letters in this case which financial ratios is from quantitative structured data on financial statements. Moreover, textual data here does not contain only pure financial information but also comprise of content related to the market (Covid-19, Russian-Ukrainian war), forward-looking statement about future expectation, operation, human resource and other non-financial area. These topics won't always directly affect the financial performance and position of the company, while financial ratios only contain historical financial information on the past year without forward-information or market influence provided.

For example, the Altman Z-score is high in 2020 indicate strong financial health, but the sentiment score is low which reflecting of the market condition. A similar situation is applied to the case of the sentiment score decreases from 2021 to 2022 due to market condition of the war even though the Altman Z-score stays high with slight increase.

On the contrary, in 2021, the Altman Z-score decrease due to the additional funding from bond issuance to prepare for future acquisition opportunity which increase in financial risk, while the sentiment score takes the event as positive tone for word "opportunity" which show that even though there is an increase in debt but not because the company is in difficult condition.

In conclusion, Gedeon Richter's financial health is very strong with a promising future. The consistency of high Altman Z-score in the safe range in 5 years period. In addition to that, the textual analysis has given insight into the underlying information (tone, main topics) which complement financial ratios analysis with high positive sentiment score and significant focus on the grow topics of the company from the CEO's letters in 2023. Besides, the textual analysis result show that, Gedeon Richter's financial health is good at managing through the difference crisis with excellent operation and even though the financial decision of Gedeon Richter bring about more risk at certain time, it is for good strategic purpose, not because the company is in distress situation.

5. CONCLUSIONS AND RECOMMENDATIONS

5.1. Methodological contributions

According to IFRS (*Figure 6*), an entity's financial position and performance is not only dependent on resource and relationship, business model, risk or strategy but also be affected by the external environment. The traditional financial ratios analysis method has been used to assess the risk like Altman Z-score and financial ratios to assess the financial operation of the company. But not all factors could be captured using historical financial information.

Textual analysis is a tool that could be used in this case to address the gap of financial ratios analysis by assessing the textual information in annual report, which contain more information than only financial data. However, the analysis of textual data is challenging due to its complexity and inconsistency compared to the objective financial data.

In our case study, two methods have been used alongside and give an insightful and complementary result into financial health of the company Gedeon Richter. Textual analysis has exposed external financial risks related to market condition and latent tone with outlook of the management, while financial ratios ensure the objective financial risk assessment of the company financial past performance. The use of textual analysis alongside the traditional method of financial statements analysis gives a holistic view of company financial health and reduce the shortcoming of each method.

Using both methods in practice to evaluate the financial health of corporations is suggested for comprehensive assessment, especially the one with complex structure, operating in complex industries with volatile market. With that type of objects, the analysis of textual information would be important.

5.2. Limitations of the study

The study has some limitations that need to be addressed. Firstly, the generalization of the methodology, even though it shows good result in the case study, it cannot be sure to be performed as well in any other special case studies due to the inconsistent of textual data as the content is varied and could be changed easily and affect the result, besides the source of textual data is not guaranteed in the reports as it is not required by the law and even company inside the same industry could yield very different textual data to be compared.

Secondly, about the subjectivity of textual analysis methodology, even though it helps mining the latent information and trend inside text in more objective way, but the data source is still subjective by its nature, so the interpretation of textual analysis result is still have depend

on the researchers, for example the categories of sentiment analysis is still subjective, there could be other way to categorize the tone and it depend on the need of researchers.

Thirdly, the textual analysis could capture the negative and positive in the tone related to the market or operation, but it cannot quantify the impact of the effect as the data is not quantitative data.

Fourthly, the textual data here is from the viewpoint of the CEO which could be biased, it would be better if the viewpoint is from different stakeholders.

Fifthly, the textual analysis method and financial ratios analysis method are affected by industry-specific challenges, so the characteristics of the company and its industry is very important in the analysis, so the methodologies are needed to be adjusted in different ways when assessing different companies in different industries.

Finally, the time frame for the studies is limited to a 5-year period, working on a longer time frame could provide better trends and reveal the financial health of the company better.

5.3. Suggestions for future research

There is a need for further exploration of the topic of textual analysis application as there are more sophisticated methods to get more insight of financial health of the company. As the textual data is complex, a classification or topic modelling method of the textual content to have information on each topic separately like finance, operation, market... could be created, which could provide more targeted analysis.

Additionally, more advanced sentiment analysis techniques like machine learning or NLP (natural language process) could be deployed to measure the sentiment with clarification on which factor causes the sentiment change like operation or market situation instead of only giving the sentiment score. The name entity recognition (NER) is very helpful if it could extract exactly the name appear in the text data like the countries, companies, currencies, products. This is very important in the pharmaceutical industry when both financial-specific and pharmaceutical-specific vocabularies need to be properly recognized and analyzed as they are multifaceted and quickly changed and updated.

It is necessary to research on the importance of textual data and should it be made a required part with required content on the annual reports. And if the textual content is a required parts on annual report, what information should be included and what effect it could have on the financial report practice.

SUMMARY

The study has the primary objective of assessing if the two methods: financial ratios analysis and textual analysis give a more comprehensive view of a company's financial health by analyzing both quantitative and qualitative data. The financial ratios analysis method is expected to give quantitative assessment of financial conditions of the company while textual analysis is to reveal the qualitative information which could not be capture by number like external risks, sentiment, strategic focuses.

The company Gedeon Richter was used as a case study to assess financial health in the period (2019-2023) with two mentioned methods. The examined period is marked with significant market disruptions like Covid-19 and Russian-Ukrainian war which influenced many entities in the market, particularly the pharmaceutical market. Financial ratios analysis is used to calculate Altman Z-score, profitability ratios, liquidity ratios, debt management related ratios, cash flow related ratios and sale growth, while textual analysis focused on evaluating CEO's letters.

The key findings highlighted that financial ratios reflect the past financial operation of the company and overlook indirect risks like market risks, regulatory risks.... On the other hands, textual analysis takes into consideration of not only financial operation but also other important topics like outlook, internal operations, market conditions, it could capture the management's tone which is affected by both non-financial and financial aspects.

Using only one of two methods could be misleading or incomplete. For example, in some cases even though company gain more risk it could be positive due to the alignment with the strategy and conversely, in other cases, the negative tone is not reflecting the financial operation as it is affected by market situation which not directly affect financial operations, but it is still a risk to be considered as it may affect company in the long run.

While the two methods are fundamentally different, they complement each other in providing a holistic view of the financial health of the company through addressing both quantitative, historical data and subjective outlook and sentiment.

There are several limitations to be considered regarding the subjectivity, complexity and unstructured nature of textual data and the bias could be from the only point of view of CEO and the generalization of the method due to the different in the textual data is not always available and the finance and industry specific vocabulary could cause a lot of difficulty of analyzing.

In conclusion, the use of combined methods is recommended and there is a need for further research to develop better tools for textual analysis. Exploring the potential of incorporating textual part of annual reports as a required component by law could significantly enhance the financial health assessment.

REFERENCES:

Books, Articles, Studies:

- Alhasanko, R. (2024): Assessing the Financial Health of the Company and its Relationship to its Market Value Added: Evidence from Industrial Companies Listed on the Amman Stock Exchange for the Period (2013-2021). *International Journal of Finance and Accounting*. Vol. 9. issue 3. pp. 37–51. <https://doi.org/10.47604/ijfa.2698>
- Altman, E. I. (1968): Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*. Vol. 23. issue 4. pp. 589–609. <https://doi.org/10.1111/j.1540-6261.1968.tb00843.x>
- Altman, E. I. – Haldeman, R. G. – Narayanan, P. (2002): ZETATM analysis: A new model to identify bankruptcy risk of corporations. *Journal of Banking – Finance*. Vol. 1. issue 1. pp. 29–54. [https://doi.org/10.1016/0378-4266\(77\)90017-6](https://doi.org/10.1016/0378-4266(77)90017-6)
- Altman, E. I. – Iwanicz-Drozdowska, M. – Laitinen, E. K. – Suvas, A. (2014): Distressed Firm and Bankruptcy Prediction in an International Context: A Review and Empirical Analysis of Altman's Z-Score Model. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2536340>
- Anand, V. – Bochkay, K. – Chychyla, R. – Leone, A. (2020): Using Python for Text Analysis in Accounting Research. *Foundations and Trends® in Accounting*. Vol. 14. issues 3–4. pp. 128–359. <https://doi.org/10.1561/14000000062>
- Antalóczy, K. – Gáspár, T. – Sass, M. (2019): The Specialties of the Pharmaceutical Value Chains in Hungary. *Acta Oeconomica*. Vol. 69. issue S2. pp. 41–72. <https://doi.org/10.1556/032.2019.69.s2.3>
- Frank, M. Z. – Antweiler, W. (2001): Is all that talk just noise? The information content of internet stock message boards. AFA 2002 Atlanta Meetings, Sauder School of Business Working Paper. <https://doi.org/10.2139/ssrn.282320>
- Arkan, T. (2016): The Importance of Financial Ratios in Predicting Stock Price Trends: A Case Study in Emerging Markets. *Zeszyty Naukowe Uniwersytetu Szczecińskiego Finanse Rynki Finansowe Ubezpieczenia*. Vol. 79. pp. 13–26. <https://doi.org/10.18276/frfu.2016.79-01>
- Balakrishnan, R. – Qiu, X. Y. – Srinivasan, P. (2010): On the predictive ability of narrative disclosures in annual reports. *European Journal of Operational Research*. Vol. 202. issue 3. pp. 789–801. <https://doi.org/10.1016/j.ejor.2009.06.023>

- Barnes, P. (1987): The Analysis and Use of Financial Ratios: A Review Article. *Journal of Business Finance – Accounting*. Vol. 14. issue 4. pp. 449–461. <https://doi.org/10.1111/j.1468-5957.1987.tb00106.x>
- Beaver, W. H. (1966): Financial Ratios As Predictors of Failure. *Journal of Accounting Research*. Vol. 4. pp. 71. <https://doi.org/10.2307/2490171>
- Berk, J. B. – DeMarzo, P. M. (2017): *Corporate finance* (4th ed.). Pearson.
- Bertoneche, M. – Knight, R. (2001): Assessing financial health. In *Financial Performance*. pp. 74–105. Elsevier. <https://doi.org/10.1016/B978-075064011-4.50003-X>
- Bologna University, Italy – Aureli, S. (2017): A comparison of content analysis usage and text mining in CSR corporate disclosure. *The International Journal of Digital Accounting Research*. Vol. 17. pp. 1–32. https://doi.org/10.4192/1577-8517-v17_1
- Brown, S. V. – Tucker, J. W. (2011): Large-Sample Evidence on Firms’ Year-over-Year MD&A Modifications. *Journal of Accounting Research*. Vol. 49. issue 2. pp. 309–346. <https://doi.org/10.1111/j.1475-679X.2010.00396.x>
- Capo, F. – Brunetta, F. – Boccadelli, P. (2014): Innovative Business Models in the Pharmaceutical Industry: A Case on Exploiting Value Networks to Stay Competitive. *International Journal of Engineering Business Management*. Vol. 6. pp. 23. <https://doi.org/10.5772/59155>
- Chakraborty, B. – Bhattacharjee, T. (2020): A review on textual analysis of corporate disclosure according to the evolution of different automated methods. *Journal of Financial Reporting and Accounting*. Vol. 18. issue 4. pp. 757–777. <https://doi.org/10.1108/JFRA-02-2020-0047>
- Courtis, J. K. (1978): Modelling a financial ratios categoric framework. *Journal of Business Finance – Accounting*. Vol. 5. issue 4. pp. 371–386. <https://doi.org/10.1111/j.1468-5957.1978.tb01059.x>
- Crowley, R. (2019): Text analytics, NLP, and accounting research. Presentation given at Peking University. pp. 1-27.
- Danzon, P. M. (2006): Economics of the pharmaceutical industry. *NBER Reporter*. Fall 2006. pp. 14–17. National Bureau of Economic Research.
- Das, S. R. – Chen, M. Y. (2007): Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web. *Management Science*. Vol. 53. issue 9. pp. 1375–1388. <https://doi.org/10.1287/mnsc.1070.0704>

- Devi, S. S. – Radhika, Y. (2018): A Survey on Machine Learning and Statistical Techniques in Bankruptcy Prediction. *International Journal of Machine Learning and Computing*. Vol. 8. issue 2. pp. 133–139. <https://doi.org/10.18178/ijmlc.2018.8.2.676>
- Elliott, B. – Elliott, J. (2009): *Financial accounting and reporting* (13th ed.). Pearson.
- Fairclough, N. (2003): *Analysing discourse: Textual analysis for social research*. Routledge.
- Green, B. F. – Stone, P. J. – Dunphy, D. C. – Smith, M. S. – Ogilvie, D. M. (1967): The General Inquirer: A Computer Approach to Content Analysis. *American Educational Research Journal*. Vol. 4. issue 4. pp. 397. <https://doi.org/10.2307/1161774>
- Guan, S. – Tian, S. – Deng, G. (2021): Revenue diversification or revenue concentration? Impact on financial health of social enterprises. *Public Management Review*. Vol. 23. issue 5. pp. 754–774. <https://doi.org/10.1080/14719037.2020.1865439>
- Gulácsi, L. – Brodszky, V. – Péntek, M. – Varga, S. – Vas, G. – Boncz, I. (2009): History of health technology assessment in Hungary. *International Journal of Technology Assessment in Health Care*. Vol. 25. issue S1. pp. 120–126. <https://doi.org/10.1017/S0266462309090527>
- Gulácsi, L. – Dávid, T. – Dózsa, C. (2002): Pricing and reimbursement of drugs and medical devices in Hungary. *Eur J Health Econom*. Vol. 3. pp. 271–278. <https://doi.org/10.1007/s10198-002-0148-z>
- Hajek, P. – Olej, V. – Myskova, R. (2014): Forecasting Corporate Financial Performance Using Sentiment in Annual Reports for Stakeholders' Decision-Making. *Technological and Economic Development of Economy*. Vol. 20. issue 4. pp. 721–738. <https://doi.org/10.3846/20294913.2014.979456>
- Henry, E. (2006a): Are investors influenced by how earnings press releases are written? *SSRN Electronic Journal*. <https://ssrn.com/abstract=933100>
- Henry, E. (2006b): Market Reaction to Verbal Components of Earnings Press Releases: Event Study Using a Predictive Algorithm. *Journal of Emerging Technologies in Accounting*. Vol. 3. issue 1. pp. 1–19. <https://doi.org/10.2308/jeta.2006.3.1.1>
- Henry, E. – Leone, A. J. (2009): Measuring Qualitative Information in Capital Markets Research. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1470807>
- Herath, S. K. – Albarqi, N. (2017): Financial Reporting Quality: A Literature Review. *International Journal of Business Management and Commerce*. Vol. 2. pp. 1–14.
- Faello, J. (2015): Understanding the limitations of financial ratios. *Academy of Accounting and Financial Studies Journal*. Vol. 19. issue 3. pp. 75–86.

- Kerpel-Fronius, S. (2013): The Development of Pharmaceutical Medicine in Hungary. *Pharmaceutical Medicine*. Vol. 27. issue 5. pp. 289–295. <https://doi.org/10.1007/s40290-013-0031-5>
- Khanna, I. (2012): Drug discovery in pharmaceutical industry: Productivity challenges and trends. *Drug Discovery Today*. Vol. 17. issues 19–20. pp. 1088–1102. <https://doi.org/10.1016/j.drudis.2012.05.007>
- Koryński, P.: Small business financial health scorecard. Microfinance Center.
- Kwartler, T. (2017): Text mining in practice with R. John Wiley – Sons Ltd. <https://doi.org/10.1002/9781119282105>
- Li, F. (2010): The Information Content of Forward-Looking Statements in Corporate Filings—A Naïve Bayesian Machine Learning Approach. *Journal of Accounting Research*. Vol. 48. issue 5. pp. 1049–1102. <https://doi.org/10.1111/j.1475-679X.2010.00382.x>
- Lőrinczy, M. (2013): Impact of the crisis on the pharmaceutical market in the Czech Republic and Hungary. *Forum Scientiae Oeconomia*. Vol. 1. issue 2. pp. 21–33. Retrieved from <https://ojs.wsb.edu.pl/index.php/fso/article/view/92>
- Loughran, T. – McDonald, B. (2011): When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance*. Vol. 66. issue 1. pp. 35–65. <https://doi.org/10.1111/j.1540-6261.2010.01625.x>
- Loughran, T. – McDonald, B. (2014): Measuring Readability in Financial Disclosures. *The Journal of Finance*. Vol. 69. issue 4. pp. 1643–1671. <https://doi.org/10.1111/jofi.12162>
- Loughran, T. – McDonald, B. (2024): Measuring Firm Complexity. *Journal of Financial and Quantitative Analysis*. Vol. 59. issue 6. pp. 2487–2514. <https://doi.org/10.1017/S0022109023000071>
- Luo, X. – Liu, Q. – Qiu, Z. (2021): A Correlation Analysis of Construction Site Fall Accidents Based on Text Mining. *Frontiers in Built Environment*. Vol. 7. pp. 690071. <https://doi.org/10.3389/fbuil.2021.690071>
- Malerba, F. – Orsenigo, L. (2015): The evolution of the pharmaceutical industry. *Business History*. Vol. 57. issue 5. pp. 664–687. <https://doi.org/10.1080/00076791.2014.975119>
- Martin, J. – Siehl, C. (1990): Organizational culture: A key to financial performance? In B. Schneider (Ed.), *Organizational climate and culture*. pp. 241–281. Jossey-Bass.
- McKee, A. (2001): *Textual analysis: A beginner's guide*. Sage Publications.

- Mehl, M. R. (2006): Quantitative Text Analysis. In M. Eid – E. Diener (Eds.), Handbook of multimethod measurement in psychology. pp. 141–156. American Psychological Association. <https://doi.org/10.1037/11383-011>
- Moon, S. – Vieira, M. – Ruiz, A. A., et al. (2022): New business models for pharmaceutical research and development as a global public good: Considerations for the WHO European region. World Health Organization. <https://www.ncbi.nlm.nih.gov/books/NBK587828/>
- Morse, M. H. (2003): Product Market Definition in the Pharmaceutical Industry. Antitrust Law Journal. Vol. 71. issue 2. pp. 633–676. <http://www.jstor.org/stable/40843599>
- Ohlson, J. A. (1980): Financial Ratios and the Probabilistic Prediction of Bankruptcy. Journal of Accounting Research. Vol. 18. issue 1. pp. 109. <https://doi.org/10.2307/2490395>
- Österreichisches Bundesinstitut für Gesundheitswesen. (2006): Surveying, assessing and analysing the pharmaceutical sector in the 25 EU member states: Country profiles. European Commission - DG Competition.
- Ricoeur, P. (1971): What is a Text? Explanation and Interpretation. In Mythic-Symbolic Language and Philosophical Anthropology. Springer, Dordrecht. https://doi.org/10.1007/978-94-011-9327-6_7
- Roberts, C. W. (2000): A conceptual framework for quantitative text analysis. Quality – Quantity. Vol. 34. issue 3. pp. 259–274. <https://doi.org/10.1023/A:1004780007748>
- Saus-Sala, E. – Farreras-Noguer, À. – Arimany-Serrat, N. – Coenders, G. (2021): Compositional DuPont Analysis. A Visual Tool for Strategic Financial Performance Assessment. In P. Filzmoser – K. Hron – J. A. Martín-Fernández – J. Palarea-Albaladejo (Eds.), Advances in Compositional Data Analysis. pp. 189–206. Springer International Publishing. https://doi.org/10.1007/978-3-030-71175-7_10
- Scherer, F. M. (2000): The pharmaceutical industry. In F. M. Scherer (Ed.), Handbook of health economics. Vol. 1. Elsevier.
- Short, J. C. – Palmer, T. B. (2008): The Application of DICTION to Content Analysis Research in Strategic Management. Organizational Research Methods. Vol. 11. issue 4. pp. 727–752. <https://doi.org/10.1177/1094428107304534>
- Smith, M. – Taffler, R. J. (2000): The chairman's statement - A content analysis of discretionary narrative disclosures. Accounting, Auditing – Accountability Journal. Vol. 13. issue 5. pp. 624–647. <https://doi.org/10.1108/09513570010353738>

- Stanton, P. – Stanton, J. – Pires, G. (2004): Impressions of an annual report: An experimental study. *Corporate Communications: An International Journal*. Vol. 9. issue 1. pp. 57–69. <https://doi.org/10.1108/13563280410516500>
- Szalkai, Z. (2004): Relationship marketing in the pharmaceutical market: Analysis of the Hungarian case. *Periodica Polytechnica Social and Management Sciences*. Vol. 12. issue 2. pp. 177–188.
- Tepperman, A. (2009): Market definition and the characteristics of pharmaceutical markets. Presented at the Canadian Bar Association 2009 Competition Law Fall Conference. Charles River Associates.
- Theodossiou, P. T. (1993): Predicting Shifts in the Mean of a Multivariate Time Series Process: An Application in Predicting Business Failures. *Journal of the American Statistical Association*. Vol. 88. issue 422. pp. 441–449. <https://doi.org/10.2307/2290323>
- Van Beest, F. – Braam, G. – Boelens, S. (2009): Quality of financial reporting: Measuring qualitative characteristics (NiCE Working Paper 09-108). Nijmegen Center for Economics.
- Weida, E. B. – Phojanakong, P. – Patel, F. – Chilton, M. (2020): Financial health as a measurable social determinant of health. *PLOS ONE*. Vol. 15. issue 5. pp. e0233359. <https://doi.org/10.1371/journal.pone.0233359>
- Wilcox, J. W. (1976): The gambler's ruin approach to business risk. *Sloan Management Review*. Vol. 18. issue 1. pp. 33.
- Yan, J. (2022): Text mining with R: A tidy approach, by Julia Silge and David Robinson. *Natural Language Engineering*. Vol. 28. issue 1. pp. 137–139. <https://doi.org/10.1017/S1351324920000649>
- Yuthas, K. – Rogers, R. – Dillard, J. F. (2002): Communicative action and corporate annual reports. *Journal of Business Ethics*. Vol. 41. issues 1–2. pp. 141–157. <https://doi.org/10.1023/A:1021314626311>

Data Collection on the Internet:

- Evaluate Pharma (2023, August 14): Total global spending on pharmaceutical research and development from 2014 to 2028 (in billion U.S. dollars). Retrieved from <https://www.statista.com/statistics/309466/global-r-and-d-expenditure-for-pharmaceuticals/> (accessed: 1 August 2024).

- Gedeon Richter. (2019–2023): Annual reports. Gedeon Richter Plc. Retrieved from <https://www.gedeonrichter.com/en/investors/company-reports> (accessed: 1 August 2024).
- Hungarian Central Statistical Office. (2024a): A regisztrált vállalkozások száma nemzetgazdasági ág szerint. https://www.ksh.hu/stadat_files/gsz/hu/gsz0003.html (accessed: 1 August 2024).
- Hungarian Central Statistical Office. (2024b): Retail sales by major commodity groups [million HUF]. https://www.ksh.hu/stadat_files/bel/en/bel0005.html (accessed: 1 August 2024).
- Hungarian Central Statistical Office. (2024c): Value and value indices of external trade in goods, HUF, by main group of commodities and by group of commodities, January–July 2024. https://www.ksh.hu/stadat_files/kkr/en/kkr0050.html (accessed: 1 August 2024).
- Hungarian Central Statistical Office. (2024d): Value of industrial production by sub-sections monthly, cumulative from the beginning of the year [at current prices, million HUF]. https://www.ksh.hu/stadat_files/ipa/en/ipa0035.html (accessed: 1 August 2024).
- Hungarian Central Statistical Office. (2024e): Pharmacies. https://www.ksh.hu/stadat_files/ege/en/ege0021.html (accessed: 1 August 2024).
- International Financial Reporting Standards. (2010): IFRS practice statement 1: Management commentary. <https://www.ifrs.org/content/dam/ifrs/project/management-commentary/ifrs-practice-statement-management-commentary.pdf> (accessed: 1 August 2024).
- International Financial Reporting Standards. (2021): IFRS® practice statement: Exposure draft ED/2021/6 management commentary. <https://www.ifrs.org/projects/work-plan/management-commentary/> (accessed: 1 August 2024).
- IQVIA (2024): Revenue of the worldwide pharmaceutical market from 2001 to 2023 (in billion U.S. dollars). <https://www.statista.com/statistics/263102/pharmaceutical-market-worldwide-revenue-since-2001/> (accessed: 1 August 2024).
- OGYÉI. (2024): Process of Marketing Authorisation. Retrieved from https://ogyei.gov.hu/uj_beadvany_en (accessed: 1 July 2024).
- Pugatch Consilium (2015): Measuring the global biomedical pulse. <https://www.pugatch-consilium.com/reports/BCI%202015%20-%20Measuring%20the%20Biomedical%20Pulse.pdf> (accessed: 1 August 2024).