

## Supplementary material

Maud Bernisson<sup>1</sup>  
Sergio Sismondo<sup>1, 2</sup>

<sup>1</sup> Institute for Science in Society, Radboud University, Nijmegen, Netherlands,

<sup>2</sup> Department of Philosophy, Queen's University, Kingston, Canada

## Materials and Methods

### *The Archive*

The documents in the Industry Documents Library come from Mallinckrodt litigations (US Bankruptcy Court for the District of Delaware) and were made public following subpoenas and information requests from several Attorney general and civil investigative demands. The archive concerning Mallinckrodt includes documents from 2002 to 2020 and encompasses 1 398 993 documents (see Caleb Alexander et al., 2022).<sup>1</sup> These documents were provided by the US Courts to the UCSF industry documents library, hence their public availability. This archive is made of diverse documents, e.g., invoices, presentations, or e-mails, that concern marketing and sales. It also includes contracts, proposals, and invoices, which we retrieved to document Mallinckrodt's implementation of ghost management through its contractual relationships with a diverse range of stakeholders.

### *Search queries summary*

Search queries were built departing from the type of contracts, which are template-based. For example, "author agreement." Nevertheless, a strong limitation was discovering enough different templates to be able to find more different templates. Consequently, our search strategy was to also use very broad queries, e.g., "*statement of work*" "*manuscript*".

**Supplementary Table 1 Search queries and results**

	Queries	Date of the query	#Result
1	"author agreement"	2022_11_28	437
2	"statement of work" "manuscript"	2022_11_28	819
3	reconciliation "working title"	2022_12_15	549
4	"INVESTIGATOR - SPONSORED STUDY AGREEMENT (ISS)"	2022_12_15	81
5	"research agreement"	2022_12_15	393
6	"Research organization master services agreement"	2022_12_15	12
7	"Research grant agreement"	2022_12_15	19
8	medlogix agreement	2023_01_11	984
9	Journal Publishing Agreement -"Subject: " -"Jurnista" -"ppt" - powerpoint -"Dahm et al Clin J Pain 1998.txt"	2023_01_24	154
10	grant compliance committee approval publish* -"Subject: " -"Jurnista" -"ppt" -powerpoint -"Dahm et al Clin J Pain 1998.txt"	2023_01_24	160
11	consulting agreement "Consultant possesses expertise in the area" → corrected to [consulting agreement "possesses expertise"]	2023_01_24	257
12	"statement of work" publication -ppt -"FW: " -"Cc: "	2023_03_06	1187
13	#11 corrected to [consulting agreement "possesses expertise"]	2023_03_06	328
14	"Mallinckrodt" consulting agreement speaker -ppt -"FW: " -"Cc: " -"daily news report"	2023_04_26	1,127
15	clinical trial* "statement of work" -ppt -"FW: " -"Cc: " -"daily news report"	2023_05_01	302
16	Mallinckrodt "statement of work" "speaker program" -ppt -"FW: " -"Cc: " -"daily news report"	2023_05_01	34
17	Grant Compliance Committee "Request for Funding"	2023_05_02	265

In addition to contracts, we found key documents like emails and proposals through serendipity. We did not search for these types of documents although they appeared to provide key information.

We removed from the selection all documents that had the same ID. We were left with 3,862 documents. We went manually through these documents and removed duplicates and alternative versions of documents (1305 out of 3862). Among all versions of a document, we privileged the oldest one, which usually includes the signatures of the parties involved in the contract. The final number of documents selected amounts to 876.

*The figure (Figure 1)*

In Figure 1, there are 374 nodes, that is, different actors involved in the network. While the data set contains 876 contracts, the figure shows 915 edges. This is because one sub-contractors might be many. If one contract mentions several clients and describes the budget of all clients, then, there is one edge per relationship among each of the clients (e.g., yjfc0254 describes services and costs per pharmaceutical company, then, edge 1 will represent pharma company 1 linked to the service contractor, edge 2 will represent pharma company 2 linked to the service contractor, etc.)

CROs and MECCs often offer similar activities (e.g., medical writing). Thus, the category "CRO" includes companies that offer to produce studies for Mallinckrodt. Companies that offer medical writing and other marketing services are categorized as MECCs.

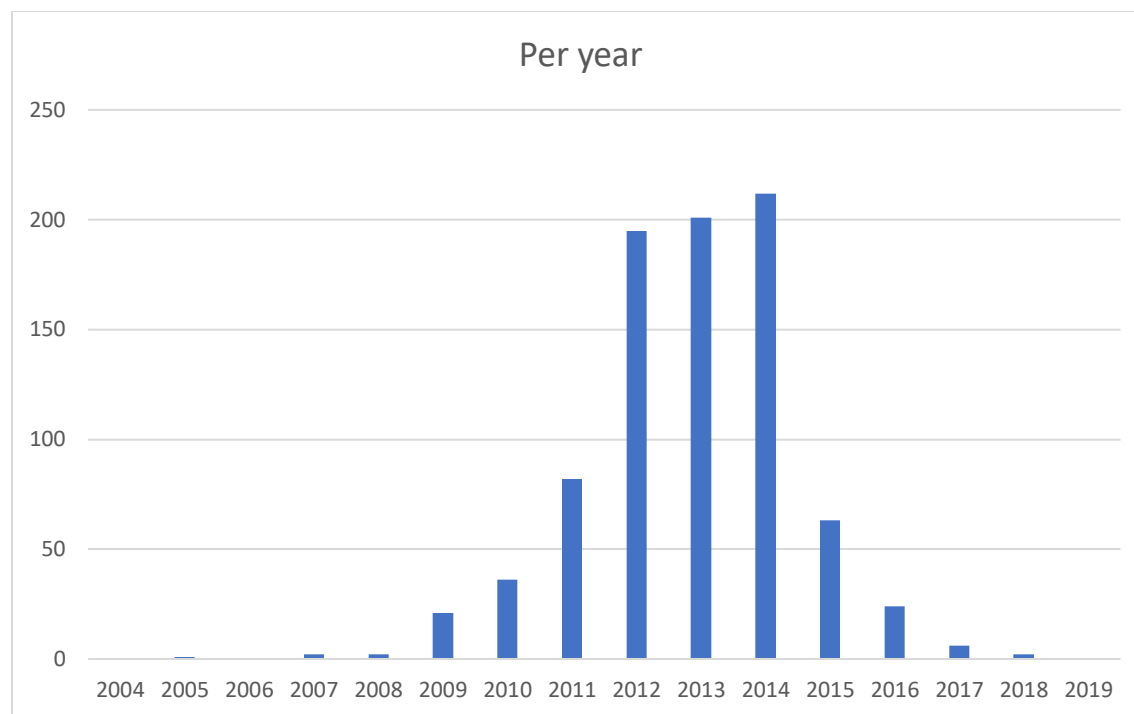
We coded Covidien as Mallinckrodt. We also coded individuals working for Mallinckrodt and who have had an author agreement with Synchrony, as Mallinckrodt.

Finally, edges do not represent the number of activities per contract, only the amount of relationships between different actors. One contract can mention several activities, though. For example, an outlier

specifies 145 speaker programs (e.g., lfvw0243). However, it is quite common to find in contracts several posters or abstracts (2 to 16). However, if a document contains several contracts, the different contracts were listed in the final data set (e.g., xnyw0232).

### *Limitations of the data*

The number of documents included in the data set does not represent the overall publications by Mallinckrodt. Other documents provide information on many more publications (NSAIDs included), however, the corresponding contracts did not appear in the SERPs. The dates of the documents show limitations to the data set:



*Supplementary Figure 1 Number of contracts per year*

Other limitations concerned OCR, which is problematic to retrieve all documents relating to a search query. In addition, the librarians are working regularly on the database, which means that results might change over time. Finally, the search was also limited to the documents that came to our knowledge. There might be documents that can be retrieved with specific keywords that we were not aware of.

### *Anonymization*

All individuals were anonymised on the figure. The results and the analysis are not altered by the anonymization, since we focused on the ghost management system rather than individuals.

## **References**

---

<sup>1</sup> Alexander GC, Mix LA, Choudhury A, Taketa R, Tomori C, Mooghali M, Fan A, Mars A, Ciccarone D, Patton M, Apollonio DE, Schmidt L, Steinman MA, Greene J, Knight KR, Ling PM, Seymour AK, Glantz S, and Tasker K.

