Quantifying Keystroke Disambiguation for the Top 200 Medication Formulary from 2025

Gaurav Bhatti, Allen Flynn
December 15, 2025

Abstract

Medication name confusion is a documented patient safety risk. This study utilized the Medication Name Overlap Analyzer (MNOA) to quantify medication name overlap and what it would take to complete "keystroke disambiguation" of names on the Top 200 list of medication names from 2025.

In the context of searching for medications by name by typing one letter at a time into a computer search bar, ***keystroke disambiguation*** is a term for the number of keystrokes required to uniquely identify each medication name on any list of these names.

This analysis compared keystroke disambiguation for three lists of medication names: Generic-Only names from the Top 200 list (2025), Brand-Only names from the Top 200 list (2025), and a Combined list of all Generic and Brand medication names from the Top 200 list (2025).

The left-to-right overlap in medication names on the three lists varies. While any overlap in Brand names was resolved within a maximum of 6 keystrokes, the medication name overlap found in the Generic-Only and Combined lists required a maximum of 12 keystrokes to uniquely resolve every medication name on those lists. The greater number of keystrokes (12) required to disambiguate every name on the Generic-Only and Combined lists stems from the fact that the Generic names on both lists include overlapping entries like *fluticasone* vs. *fluticasone propionate*.

1. Introduction

Accurate medication selection is critical for patient safety. In Automated Dispensing Cabinet (ADC) and Computerized Provider Order Entry (CPOE) systems, look-alike/sound-alike (LASA) names can lead to medication selection errors when the system does not require enough distinct keystrokes to differentiate between similar or somewhat overlapping medication names (e.g., Glucotrol vs. glucophage).

The Medication Name Overlap Analyzer (MNOA) tool quantifies this risk by calculating the number of characters that need to be typed into a computerized search mechanism to achieve complete keystroke disambiguation.

This report presents a comparative analysis of three lists of medication names derived from the US "Top 200" formulary from 2025. These three lists are referred to in the following manner:

1. **Combined:** A list containing both Brand and Generic names.
2. **Brand-Only:** Brand names only.
3. **Generic-Only:** Generic names only.

2. Methods

2.1 Data Source

The analysis utilized the **Top 200 Brand and Generic Drug Names for 2025** (Source: PTCB), representing the most frequently prescribed medications in the United States. Each list included a different number of medication names, as follows:

- **Combined List** (N=447)
- **Brand-Only List** (N=240)
- **Generic-Only List** (N=207)

2.2 Algorithm

As it processes any given list of medication names, the MNOA algorithm first transforms the medication names on an input list of names (e.g., converting all names to lowercase and removing duplicates) and then iterates character-by-character from left to right to determine all instances of left-to-right character overlap in the names on the list, including the **Longest Overlap(s)**. The output provided by the MNOA algorithm includes a character-by-character, round-by-round report of medication name overlap. The MNOA algorithm also computes the most powerful keystroke (from keystroke number 1 to n, where n is the total number of characters required to fully disambiguate all medication names on a list).

3. Results

The comparative analysis reveals the degree of overlap among the medication names on each of the three lists studied (Table 1).

Table 1: Comparative Medication Name Overlap and Keystroke Disambiguation Metrics

Metric	Definition	Combined List	Brand-Only List	Generic-Only List
Total Names Analyzed (N)	The final count of names analyzed.	447 names	240 names	207 names
Total Rounds to Disambiguate	The final number of rounds needed to achieve complete keystroke disambiguation.	12 rounds	6 rounds	12 rounds
Longest Overlap (Characters)	The length of the longest overlapping prefix that still had unresolved names.	11 characters	5 characters	11 characters
Round with Most Powerful Keystroke	The number of the round of analysis with the highest *KPraw* value	Round 3	Round 3	Round 3
Raw Keystroke Power (KPraw)	The number of names resolved during the Most Powerful Keystroke round.	192 names	123 names	89 names
Percentage Keystroke Power (%KP)	The percentage of names resolved during the Most Powerful Keystroke round.	42.95%	51.25%	43.00%
Examples of Greatest Medication Name Overlap	The names sharing the longest character overlap, setting the maximum keystroke requirement for the analyzed list	fluticasone *vs.* fluticasone propionate	glucophage *vs.* glucotrol	fluticasone *vs.* fluticasone propionate

Next, Table 2 below provides more detail on each round of keystroke disambiguation for the 3 lists of medication names. Table 2 includes a row with "0" characters, indicating the number of names on each list before a character-by-character, left-to-right search begins (Round 0).

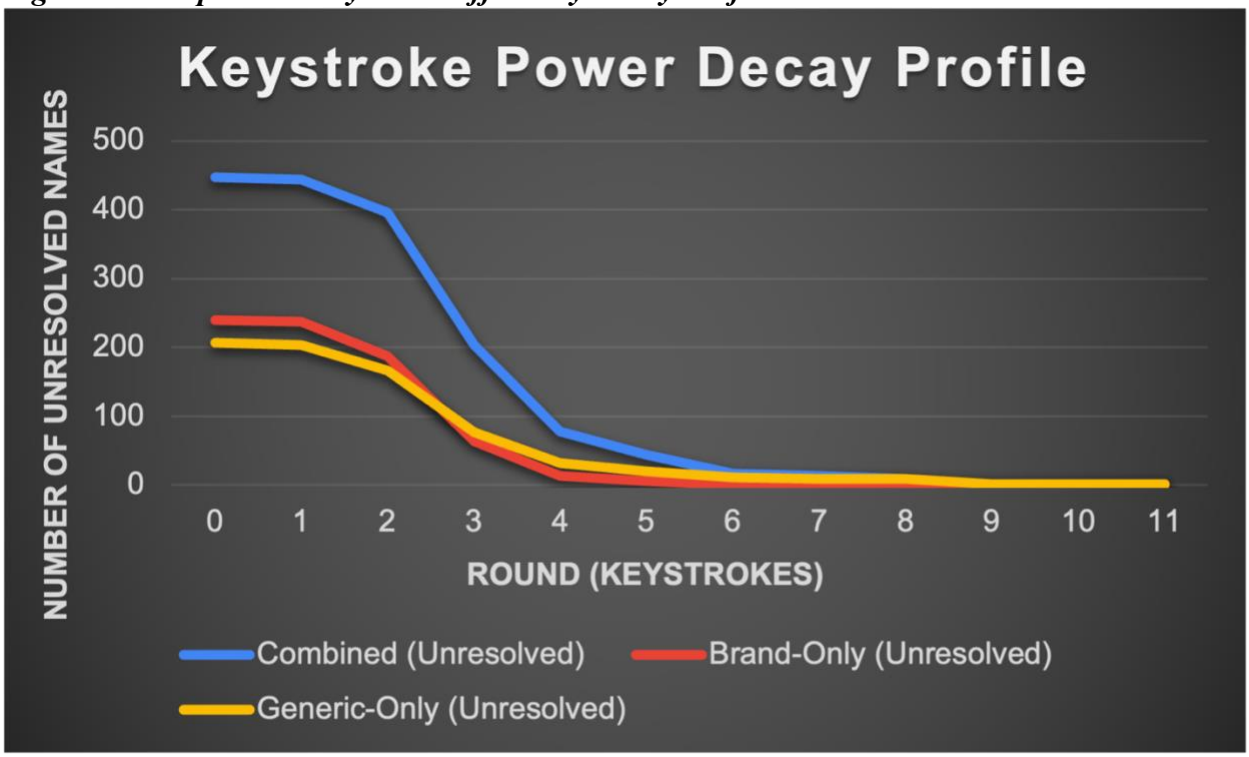
Table 2 also shows the number of medication names remaining to be disambiguated after each keystroke entered up until complete keystroke disambiguation (1 to 12). By Round 6, the Brand-Only list uniquely identified every medication name (0 unresolved), while the Combined and Generic lists of medication names are not fully resolved until Round 12. These rounds where keystroke disambiguation became complete are indicated by the tan shaded cells in Table 2.

Table 2: Keystroke Disambiguation Profile in terms of Unresolved Names by Round

Round	Number of Unresolved Names on the Combined List	Number of Unresolved Names on the Brand-Only List	Number of Unresolved Names on the Generic-Only List
0	447	240	207
1	444	237	203
2	396	187	166
3	204	64	77
4	78	14	32
5	44	6	20
6	17	**0**	11
7	13	0	9
8	9	0	9
9	2	0	2
10	2	0	2
11	2	0	2
12	**0**	0	**0**

Next, in Figure 1 below, the graph shows the "Keystroke Power Decay Profile" when using keystroke disambiguation to resolve all of the medication names on the three study lists. The graph shows how the early keystrokes, especially keystrokes 2 to 4, do most of the work of disambiguating the medication names on the studied lists.

Figure 1: Comparative Keystroke Efficiency Decay Profile



4. Discussion

The most critical finding is the difference in **Total Rounds**. The Brand-Only list is fully disambiguated with **6 keystrokes** (Maximum Name Overlap = 5). In contrast, the Generic-Only list is fully disambiguated when **12 keystrokes <u>are entered during search</u>** (Maximum Overlap = 11). This is not surprising because brand names tend to be more distinct than generic names. Brand names are created to market drug products whereas generic names exist primarily to indicate the active ingredient(s) in drug products.

Because the Combined list includes every name on the Generic-Only list, it inherits the worst case of overlap from the Generic-Only list.

Having examined medication name overlap for the names on the 2025 Top 200 drug list, we find that overlap applies to both Generic and Brand names. Further, we find that overlap increases when Generic names are included in the medication name overlap analysis. The implication of this overall finding is that for the most used drug products in the United States there is reason to be concerned that software search mechanisms could lead computer users to inadvertently search by name, and then find and select the wrong medication.

5. Conclusion

A significant medication name overlap is found in medication names on the 2025 Top 200 drug list. Overlap in medication names is longer for generic names than for brand names alone. With respect to medication safety, continued efforts to prevent medication name confusion are warranted.