**EXAMPLE RESULTS SECTION:**

**Summary of Research on 5-HT5A**

**Results.**

Table 1

*Number of Publications on 5-HT5A per Year*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Year | 1999 | 2000 | 2012 | 2013 | 2015 | 2018 |
| Number of Publications | 1 | 1 | 1 | 2 | 1 | 2 |
| Percent of Total Publications | 12.5% | 12.5% | 12.5% | 25.0% | 12.5% | 25.0% |

Table 2

*Number of Publications on 5-HT5A by Species*

|  |  |  |
| --- | --- | --- |
| Species | Rats | Mice |
| Number of Publications | 7 | 2 |
| Percent of Total Publications | 77.8% | 22.2% |

Table 3

*5-HT5A Publications by Method*

|  |  |  |  |
| --- | --- | --- | --- |
| Method | Antagonist | Immunohistochemistry | Knockout |
| Number of Publications | 6 | 1 | 1 |
| Percent of Total Publications | 75.0% | 12.5% | 12.5% |

Table 4

*5-HT5A Publications by Antagonists*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Antagonist | SB-699551 | ASP-5736 | A-843277 | AS-2030680 | AS-2674723 |
| Number of Publications | 4 | 2 | 1 | 1 | 1 |
| Percent of Total Publications | 44.4% | 22.2% | 11.1% | 11.1% | 11.1% |

Table 5

*5-HT5A Publications by Brain Regions*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Brain Region | Hippocampus | Dentate Gyrus | Olfactory Bulb | Neocortex | Medial Habenula | Ventral Tegmental Area | Prefrontal Cortex |
| Number of Publications | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Percent of Total Publications | 14.3% | 14.3% | 14.3% | 14.3% | 14.3% | 14.3% | 14.3% |

Table 6

*5-HT5A Publications by Topic*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Topic | Memory | Anxiety | Depression | Exploratory | Nociception | Schizophrenia |
| Number of Publications | 4 | 1 | 1 | 1 | 1 | 1 |
| Percent of Total Publications | 44.4% | 11.1% | 11.1% | 11.1% | 11.1% | 11.1% |

**Overview.** The 5-HT5 receptor family is relatively understudied and not as well understood as more researched families such as 5-HT1 or 5-HT2.The 5-HT5A receptor has been identified in the mouse (Plassat, Boschert, Amlaiky, & Hen, 1992; Matthes et al., 1993), rat (Erlander et al., 1993), and human (Pasqualetti et al., 1998) brains, namely in regions including, but not limited to, the hippocampus, cerebral cortex, and hypothalamus. Thomas (2006) reviewed several tentative therapeutic roles for the 5-HT5A receptor, and also summarizes its gene localization, structure, and sequencing.

**Discussion of Results.** The majority of publications studying the role of 5-HT5A on behavior have been published after 2010, likely due to somewhat recent discovery of selective ligands. These publications overwhelmingly focus on rodent (primarily rat) animal models. Furthermore, the method of choice for the majority of studies is antagonism. This is partially because there are no known selective 5-HT5A agonists (Gonzalez, Chávez-Pascacio, & Meneses, 2013). While five different antagonists were found to be used, the selective antagonist SB-699551 was by far the most researched. This may be in part because it was the first selective antagonist developed for 5-HT5A (Corbett et al., 2005). By analyzing the topics of the collected 5-HT5A studies, we also found that nearly half of publications focused on the role of 5-HT5A in memory and learning. This is a departure from the topics typically studied in publications focusing on the 5-HT1 or 5-HT2 families, such as anxiety and depression. In fact, little existing research has looked at the relationship between the 5-HT5A receptor and models of depression (Carr & Lucki, 2010). Such a focus on memory and learning may be due findings suggesting relatively high expression and wide distribution of 5-HT5A receptors in the rat hippocampus (García-Alcocer et al., 2010).

**Select Brief Literature Reviews.**

**Brief Summary of Memory Literature.** The selective antagonist SB-699551 has been linked to decreased responses in associative learning for both short- and long-term memory tasks (Gonzalez et al., 2013), and has been shown to produce amnesia-like effects (Aparicio-Nava, Márquex-García, & Meneses, 2018). However, SB-699551 has also reversed forgetting effects in object recognition (Yamazaki et al., 2018; Nikiforuk, Hołuj, Kos, & Popik, 2016). The 5-HT5A antagonists ASP5736, AS2030680, and AS2674723 were also assessed regarding their role in rodent models of dementia and age-related cognitive dysfunction, and analyses suggest that all three agents actually decreased scopolamine-induced working memory deficits and improved reference memory impairments in mice and rats, respectively (Yamazaki, Okabe, Yamamoto, Yarimizu, & Harada, 2015). Therefore, while SB-699551 may be useful in studying therapeutic approaches for dementia and memory functions, the precise role of 5-HT5A remains unclear in memory and learning remains unclear.

**Brief Summary of Schizophrenia Literature.** Furthermore, 5-HT5A has also been examined in animal models of schizophrenia. Antagonist ASP5736 may reduce positive symptoms associated with animal models of schizophrenia (Yamazaki et al., 2014), and SB-699551 has been demonstrated to reduce ketamine-induced social withdrawal, a negative symptom in animal models of schizophrenia (Nikiforuk et al., 2016). 5-HT5A may be a promising target in therapeutic research on schizophrenia.

**Brief Summary of Anxiety & Depression Literature.** SB-699551 exhibited an anxiolytic-like effect in the vocalization test, and compound A-843277 demonstrated an antidepressant-like property in the forced swim test (though this may have resulted from abdominal irritation; Kassai et al., 2012). Further analyses of 5-HT5A antagonists and their role in antidepressive function should be performed.

**Future Directions.** The 5-HT5A receptor is one of the least studied of the various 5-HT receptor subtypes. Because there are no available selective agonists, the majority of publications focus on antagonist functions. Until a greater number of selective ligands are identified, research on 5-HT5A may be limited in comparison to research on other 5-HT receptor subtypes. However, several recent publications suggest a causal role of 5-HT5A in memory functions, as well as in animal models of schizophrenia, anxiety, and depression. In particular, there is yet to be a clear consensus as to the role of the selective antagonist SB-699551 in any of these topics, and conflicting findings exist in the memory domain.

---------------------------------------------------------------------------------------------------------------------

A FEW COMMENTS/CONCERNS FOR FEEDBACK @ HANNAH & STEVE:

* When compiling tallies for various variables of interest (e.g., species, antagonists, etc.), I think it might be best to “double-count” publications for certain variables (I did this above). This means: although there may be 10 studies for a receptor subtype, the sums of each table may not add up to 10 if the same study does multiple techniques in the same variable type. For example, if a publication looks at ‘mice’ AND ‘rats’, I would count that as +1 for mice and +1 for rats, meaning that the species table will have an n greater than the number of actual publications. Do you think this is a good idea, or should there just be another column for “mice & rats”. The latter, I think, would get more confusing and would certainly be more difficult to code.
* Since we’ll be automating the text mining, I may re-run the PubMed and Web of Science library search and omit the term “behav\*”, since this is a somewhat specific (and sometimes too exclusionary) term that was mostly added to reduce our article count. Any objections?
* For studies that look at multiple receptor subtypes (this happens often), publications will be double-counted – BUT since the ‘results’ tables will be made from receptor-specific subsets, this shouldn’t affect the total number of publications per receptor, which would be untouched in the studies/year table.
* Other than these few concerns, let me know what you think about the general format of the results section. For more studied receptors, this section will obviously be much longer (and have more meaningful statistics in the tables), but I think having *brief* literature reviews for the major topics will be manageable – especially since reviews exist for some of the more studied receptors which can be referenced.

CODING UPDATE & PSEUDOCODE EXAMPLE:

Good news: the text mining required for this paper is doable! I’m still working out the details of some of the R packages I’ll be working with (pdftools, tm, tidytext, etc.), but I’ve spent some time outlining the general structure of extraction.

Here’s the basic outline:

# read in pdf of paper

file <- system.file('pdf', SEROTONIN.pdf', package = 'pdfsearch')

# convert to all lower case, and split by lines and spaces

file<- sapply(file, tolower)

file <- strsplit(file, "\n")

file <- strsplit(file, " ")

# convert to DocumentTermMatrix format

file <- VCorpus(DirSource(file), readerControl = list(reader = readPDF))

Document Term Matrix()

# extract date

for i in 1:lengthofpdf{

if format = %Y{

gsub(“.\*”, “”, file[i])}}

# perform a heading search to locate sections of paper

file <- system.file('pdf', SEROTONIN.pdf', package = 'pdfsearch')

# identify abstract (this will contain receptor(s), method(s), etc.) and subjects (identify species if # not in abstract) (NOTE: I may also include identification of ‘Methods’, but this is named # differently based on publication so this will be more difficult)

heading\_search(file, headings = c('abstract', ‘subjects’, ‘OTHER HEADERS TO INSERT’), path = TRUE)

# grab everything in the abstract

first <- “Abstract”

last <- “Introduction\\.”

abs <- paste0(“.\*”, first, “(.\*)”, “\n”, last, “.\*)

[do same for subjects/methods section]

# keyword text-match search abstract and methods sections for variables of interest

txtmatch <- function(PartOfText) {

keyword\_search(PartOfText, keyword = c(‘HT1A', ‘HT1B’, [etc…]), path = TRUE, split\_pdf = TRUE)

keyword\_search(PartOfText, keyword = c(‘knockout, ‘antagonist’, [etc…]), path = TRUE, split\_pdf = TRUE)

keyword\_search(PartOfText, keyword = c(‘rat\*’, ‘mice’, [etc…]), path = TRUE, split\_pdf = TRUE)

keyword\_search(PartOfText, keyword = c(‘AGONIST1’, ‘AGONIST2’, [etc…]), path = TRUE, split\_pdf = TRUE)

keyword\_search(PartOfText, keyword = c(‘BRAINREG1, ‘BRAINREG2, [etc…]), path = TRUE, split\_pdf = TRUE)

}

# Note: the best way to pull specific agonists, antagonists, and brain regions of study is likely #also text-match. Comprehensive lists of all three of these variables exist, and therefore I do not #think we risk anything significant slipping through the cracks so long as we account for #different precise spellings (e.g., SB-699551 versus SB699551)

# text-match function then produces tibble data frames with keyword, location of match, line of #text match, and tokens associated with line of text match

abs\_tab <- txtmatch(abs)

methods\_tab <- txtmatch(methods)

# then look at the most frequent text matches for each variable

sort(table(abs\_tab[1]))

After this step, I will need to figure out more precisely how to best identify the proper classification for when multiple different variables in the same category are found in given sections. This will probably involve cross-referencing between sections and eliminating anything only mentioned once.

Furthermore, I have yet to determine how to extract an overall ‘topic’ for each paper. While text-matching is a possibility, it would require a less constrained *a priori* list of topics (e.g., aggression, depression, etc. etc.) which would likely result in the omission of several somewhat important minor topics in the field. However, I’m meeting with a friend who does NLP several times in the next several weeks, and hopefully we’ll be able to put together something better. This may include a sentiment analysis of the text, in which elimination of any common words or non-relevant words should produce a refined list of topic-specific tokens which we could then work with.

PLAN FOR THE REST OF THE SEMESTER:

Here’s a tentative timeline for finishing the paper -

4/14: Working pilot code that can accurately extract data from ~10-20 papers that can be checked by hand.

4/28: Generalized & debugged code; run on updated library

5/9: Complete draft of paper with introduction, complete methods, 14 results sections with complete tables and mini-literature reviews, and tentative discussion section

Then I’ll be in New Haven from about June 1 to June 23, and I can continue to refine any sections of the paper in need of revisions before it is submission-ready.