**Clinician insights on using telehealth services during COVID19 pandemic in NZ: topic modelling of interview transcripts**

**Abstract**

**Background**. -- COVID19 has plunged the clinicians both into an unprecedented crisis and offered an unprecedented opportunity to incorporate telehealth in general and videoconferencing in their daily practice of patient or client care. Telehealth practice is known to have its own unintended consequences on patients, and impact of telehealth practice on clinicians is also documented; however, little is known about the adoption of telehealth in unprecedented situations and clinician perspective of such adoption. Hence we conducted a series of interviews with clinicians on their usage of telehealth devices and videoconferencing with their clients during covid19 pandemic in New Zealand.

**Materials and methods**. -- We conducted 11 interviews of the clinicians; each clinician was asked a similar set of questions. The questions and answers were audiotaped, and transcribed. All individual identifiers were removed from the interview texts. Machine learning procedures were then used as follows.

Step one, the individual interview texts were edited to remove identifiers and the text was converted to machine readable by expanding abbreviations, and standardising words and expressions, and removing punctuation marks and apostrophes. These documents were then collated to create a corpus of 10 documents. Step two, using gensim, from each of the individual document, words were extracted. Step three, using the stopwords function of the natural language toolkit of python, commonly used words referred to as "stopwords" were removed from the corpus of the documents. Step four, the corpus of the text was used to derive a dictionary of the unique words in the corpus. Step five, the dictionary was used to create a "bag of words" where each word was indexed with a positional identifier from the dictionary. Step Six, the dictionary and bag of words were used to model the corpus of the tokens to model relative frequency of unique words in individual documents over the entire corpus (term frequency - individual document frequency) to identify the relative importance of word in individual documents to indicate its importance in the corpus; to identify meaning within the corpus, we used latent semantic analysis to assign weight to words and themes.

**Results**. -- The tokens from the individual transcripts suggest different emphasis on words from transcripts from clinicians in medical/clinical practice. Specialist practitioner transcripts from Neurologists and Speech language experts transcripts or nursing emphasised words such as patients, therapy winscribe, or practice aspects such as "lsvt" or affirmative signals such as "yeah" (for speech language therapists and general practitioners). In contrast, transcripts from midwifery contained weighted expressions such as "guidance", "rude", "overwhelming", indicating their emotional interaction with the clients or systems with negative connotations and emphasis on the need for structure.

The contrast between these words were highlighted with the themes returned by the latent semantic analysis. The first theme loaded on words "Yeah", "women", "patients", "therapy" indicating that pracitioner or care based theme, and that remote consultations or telehealth was positively suited to clinical practitioners. Themes two through five positively loaded on clinical practice words ("Yeah", "therapy") and negatively loaded on "women", "mask", "families", and "facebook". When combined with the position and context of these words in the original transcripts, they suggested the contrast of negative experiences among the midwives and lactation consultants but overall acceptance among the general practitioners, neurologists, and paediatrician/neonatology practitioners and nurse practitioners. The tables and the equations are presented below.

Table 1. Tfidf scores from the neurology transcript

|  | **term** | **weight** |
| --- | --- | --- |
| 0 | patients | 0.463584 |
| 1 | timaru | 0.219999 |
| 2 | winscribe | 0.18857 |
| 3 | stroke | 0.157142 |
| 4 | systems | 0.153773 |
| 5 | coast | 0.125714 |
| 6 | neurology | 0.125714 |
| 7 | queenstown | 0.125714 |
| 8 | ups | 0.125714 |
| 9 | west | 0.125714 |

Table 2. Tfidf scores from Speech language therapy transcript

|  | **term** | **weight** |
| --- | --- | --- |
| 0 | yeah | 0.624953 |
| 1 | therapy | 0.43497 |
| 2 | sessions | 0.219578 |
| 3 | lsvt | 0.169155 |
| 4 | might | 0.126354 |
| 5 | telemedicine | 0.105779 |
| 6 | technology | 0.0872928 |
| 7 | thinking | 0.0865461 |
| 8 | teams | 0.0769298 |
| 9 | suppose | 0.0758123 |

Table 3. Tfidf scores from one Midwifery transcript

|  | **term** | **weight** |
| --- | --- | --- |
| 0 | guidance | 0.223521 |
| 1 | women | 0.185038 |
| 2 | important | 0.167641 |
| 3 | internet | 0.155659 |
| 4 | access | 0.134573 |
| 5 | telemedicine | 0.133422 |
| 6 | big | 0.117176 |
| 7 | rude | 0.117176 |
| 8 | assessment | 0.116874 |
| 9 | advised | 0.111761 |

Table 4. Tfidf from another midwifery transcript

|  | **term** | **weight** |
| --- | --- | --- |
| 0 | mask | 0.412146 |
| 1 | women | 0.248137 |
| 2 | speaker | 0.176634 |
| 3 | wore | 0.176634 |
| 4 | ppe | 0.123462 |
| 5 | scales | 0.123144 |
| 6 | interested | 0.117756 |
| 7 | overwhelming | 0.117756 |
| 8 | pad | 0.117756 |
| 9 | photos | 0.117756 |

Table 5. Tfidf scores from a neonatology/paediatrics transcript

|  | **term** | **weight** |
| --- | --- | --- |
| 0 | families | 0.308687 |
| 1 | child | 0.169002 |
| 2 | address | 0.129285 |
| 3 | brilliantly | 0.129285 |
| 4 | developmental | 0.129285 |
| 5 | clinical | 0.101401 |
| 6 | pre | 0.101401 |
| 7 | consults | 0.0903667 |
| 8 | guideline | 0.0903667 |
| 9 | remote | 0.0903667 |

Table 6. Tfidf scores from another neonatology/paediatrics transcript

|  | **term** | **weight** |
| --- | --- | --- |
| 0 | feeding | 0.38574 |
| 1 | families | 0.285494 |
| 2 | teleconference | 0.220748 |
| 3 | established | 0.165561 |
| 4 | majority | 0.165561 |
| 5 | telemed | 0.165561 |
| 6 | texts | 0.165561 |
| 7 | premature | 0.115722 |
| 8 | remote | 0.115722 |
| 9 | house | 0.115722 |

Table 7. Tfidf scores from a lactation interview transcript

|  | **term** | **weight** |
| --- | --- | --- |
| 0 | partners | 0.196797 |
| 1 | notes | 0.137556 |
| 2 | appt | 0.131198 |
| 3 | appts | 0.131198 |
| 4 | cracked | 0.131198 |
| 5 | def | 0.131198 |
| 6 | emails | 0.131198 |
| 7 | nipples | 0.131198 |
| 8 | reasonably | 0.131198 |
| 9 | rolleston | 0.131198 |

Table 8. Tfidf scores from another lactation interview transcript

|  | **term** | **weight** |
| --- | --- | --- |
| 0 | facebook | 0.30491 |
| 1 | latching | 0.174491 |
| 2 | mastitis | 0.174491 |
| 3 | milk | 0.174491 |
| 4 | breastfeed | 0.130868 |
| 5 | messaging | 0.130868 |
| 6 | mothers | 0.130868 |
| 7 | opioids | 0.130868 |
| 8 | weight | 0.130868 |
| 9 | breastfeeding | 0.121964 |

Table 9.Tfidf scores from a new nursing graduate interview transcript

|  | **term** | **weight** |
| --- | --- | --- |
| 0 | women | 0.238938 |
| 1 | limiting | 0.158747 |
| 2 | scans | 0.158747 |
| 3 | vaginet | 0.158747 |
| 4 | house | 0.147946 |
| 5 | sense | 0.126344 |
| 6 | notes | 0.110959 |
| 7 | bedroom | 0.105831 |
| 8 | clearly | 0.105831 |
| 9 | employed | 0.105831 |

Table 10. Tfidf scores from a general practitioner interview transcript

|  | **term** | **weight** |
| --- | --- | --- |
| 0 | yeah | 0.580476 |
| 1 | okay | 0.454968 |
| 2 | might | 0.15257 |
| 3 | problem | 0.129098 |
| 4 | patients | 0.128377 |
| 5 | terms | 0.116114 |
| 6 | doctor | 0.112226 |
| 7 | um | 0.0941313 |
| 8 | mean | 0.0938891 |
| 9 | four | 0.089781 |

Table 11. Loading of unobserved themes on words from the transcripts based on latent semantic analysis (the figures indicate weights, positive and negative signs indicate positive and negative correlations with the themes that can be named)

| **Theme** | **Equation** |
| --- | --- |
| Theme 1 | 0.374\*"yeah" + 0.198\*"women" + 0.162\*"patients" + 0.158\*"okay" + 0.134\*"therapy" + 0.112\*"families" + 0.094\*"might" + 0.087\*"feeding" + 0.084\*"mask" + 0.083\*"sessions" |
| Theme 2 | 0.587\*"yeah" + 0.248\*"okay" + 0.210\*"therapy" + -0.198\*"women" + 0.129\*"might" + 0.095\*"sessions" + -0.095\*"mask" + -0.083\*"families" + 0.082\*"lsvt" + -0.068\*"facebook" |
| Theme 3 | 0.341\*"families" + 0.226\*"feeding" + -0.186\*"women" + 0.131\*"child" + -0.127\*"mask" + 0.125\*"teleconference" + 0.117\*"remote" + 0.094\*"majority" + 0.094\*"texts" + 0.094\*"established" |
| Theme 4 | 0.256\*"patients" + -0.245\*"facebook" + -0.133\*"milk" + -0.133\*"latching" + -0.133\*"mastitis" + 0.126\*"timaru" + 0.108\*"winscribe" + -0.107\*"feeding" + -0.100\*"messaging" + -0.100\*"weight" |
| Theme 5 | -0.355\*"patients" + -0.171\*"timaru" + -0.146\*"winscribe" + 0.124\*"families" + -0.123\*"facebook" + -0.122\*"stroke" + -0.115\*"systems" + 0.103\*"yeah" + -0.098\*"queenstown" + -0.098\*"coast" |

**Discussion**. -- This machine generated analysis of the transcripts of interviews of a range of clinicians who used telehealth tools and videoconferencing facilities suggest that interactions and patterns of practice with telehealth tools vary with the context of practice. Some practitioners and patients adapt to the existent tools while others need guidance and better technological solutions. As technology adoption is likely to vary across disciplines, this preliminary semantic modelling of usage transcripts bring out such contrasts. However the results of this analysis need to be interpreted in the light of detailed qualitative analysis of the transcripts themselves.

**Introduction**

**Methods**

**Results**

**Discussion**

**References**