How has Computer-Assisted Qualitative Data Analysis Software affected qualitative research?

Megan Woods, PhD, Rob Macklin, PhD, Gemma Lewis, PhD, University of Tasmania Australia
Context and rationale for the research

Computer-Assisted Qualitative Data Analysis software (CAQDAS) packages

- Specialised programs for analysing qualitative data
  - Eg NUD*IST, N Vivo, Atlas ti, HyperResearch, QUALRUS, MaxQDA, QDA Miner, the Ethnograph, Leximancer, Transana

- Now an established tool for qualitative research
  - In use since early 1980s
  - Widespread usage
  - Sometimes divisive
  - Software users and non-users
  - Allegiances to specific programs
  - Advocates and critics
Questions and controversies about CAQDAS

• What can (or can’t) be done with programs? (c.f. Seidel and Clark 1984; Muhr 1991; Di Gregorio 2000; Bazeley 2002; Hutchinson, Johnston and Breckon 2010)?

• Are computer-assisted analyses more rigorous, transparent, credible or trustworthy? (cf Tallerico 1991; Dainty et al 1998; Bong 2002; Smyth 2006)

• How does using software change the process and experience of analysis? (c.f. Richards and Richards 1987; Kelle 1995; Weitzman 1999; Gilbert 2002; Davidson and Skinner 2010)

• Do programs ‘impose’ methodologies or ‘drive’ the analysis? (c.f. Bryman and Burgess 1994; Lee 2002; Seror 2005)
Questions and controversies about CAQDAS

• How do we choose between programs?
  – Are programs comparable in their features and functions? (cf The KWALON 2010 experiment)
  – Does using (any or all) CAQDAS programs compromise creative freedom? (cf DeNardo & Levers 2002)

• How do we guide new researchers through these choices? (cf Kaczynski 2003)
Our research interest and focus

Our focus in this study:
• How has the technology evolved?
• What implications has this had for qualitative research practices for
  - Creating and collecting data?
  - Analysing data?
  - Presenting data?

Our research interest:
• Computer-assisted qualitative data analysis as a form of ‘professional practice’ utilising
  - Technical reasoning and wisdom (techne)
  - Practical reasoning and wisdom (phronesis)
• Experiential learning and collective wisdom
Research method

Analysis of methodological literature from 1980 to 2012

Dataset generated by purposive sampling:
- Initial search for literature for the terms CAQDA, CAQDAS, “qualitative data analysis software”, “qualitative data analysis program” and “computer-aided qualitative data analysis”.
- Subsequently supplemented by program descriptions from manufacturers

Final data set: 163 items
Research method

Analytical strategy:
- Reviewed literature in chronological order to identify debates and trace discussions over time
- Read and wrote memo for each article detailing key points and arguments relevant to research question
- Intended to use N Vivo (version 10) to
  - Record notes about each publication
  - Develop data categorisation system reflecting discussions of
    - Program features
    - techniques supported by programs
  - Develop conceptual model illustrating relationships

• Subsequently used Word to chart the data, N Vivo to develop conceptual model
<table>
<thead>
<tr>
<th>Program feature</th>
<th>Program</th>
<th>Source detailing</th>
<th>Applications</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Reduces, eliminates need for transcription</td>
<td>Lee 2002</td>
</tr>
</tbody>
</table>

Transana
Historical development

• 1970s-early 1980s: introduction of computers to support qualitative analysis of data

• Mid 1980s to mid 1990s: introduction of dedicated CAQDAS programs

• Development of original CAQDAS programs
  - Refinement of original features and functionality
  - Introduction of additional features and functionality

• Development of additional programs
Features for data collection and creation

- Files types/ data sources that can be accommodated by programs
  - Text files eg transcripts of interviews
    - Initial formats – plain text; later formats - Rich text, HTML and PDF
  - Audio and video files
  - Pre-coded survey data
  - Data from online technologies such as web-based communication forums eg Twitter, Facebook
  - Geo data

- Creating data files
  - Text files: Initially created in WP then imported, Later created in programs with text editing functionality and transcription functionality
  - Other files types: create in other programs then import
Impact on data collection and creation

- Formatting of data files
  - Initially:
    - Structure of source documents
    - Formats of text files
    - Size of text unit
    - Number lines of text to create ‘addresses’ for coded sections
  - Subsequently
    - Conversions of formats for use in software programs

- Volume of data collected
- Types of data being used
- Integration and compatibility of programs
Features for data analysis

Marking up data with codes, tags or symbols
- By researcher assigning tags
  - Select text, assign tag
- By program assigning tags (Autocoding)
  - Specified by researcher
  - Specified by program

Indexing, categorising data
- Initially used separate database management programs as file directory
- Subsequently, indexing systems in programs to categorise data
- Editing coding/ indexing systems after applied
- Cross indexing of data
Impact on data analysis

• Retrieval of coded material
  – According to code assigned (by researcher or program)
  – For review in original context

• Identification of ‘key’ concepts

• Boolean searching and linking codes to:
  • Retrieve text fitting set parameters
  • Develop propositional relationships regarding concepts and participant characteristics
  • Investigate extent of data support for hypotheses

• Conversion of data for subsequent analysis
  – Eg converting codes into variables
Features for data analysis

Memoing
- Initially: noted in memos
- Subsequently: hyperlinking of memos to data and other elements (annotating original data source)

Integration of analyses by team members
- Initially: by merging projects
- Subsequently:
  - By supporting multiple users in project
  - By enabling simultaneous working
  - Restricting levels of access

Calculation of coding consistency scores
  Eg percentage agreement between coders, Krippendorf’s alpha
Impact on data analysis

- Logging of project decisions, actions, outcomes
- Documenting chain of evidence between data and conclusions
- Establishment of ‘team rules’ for analysis
- Determining consistency of coding approaches
- Objective determination of coding similarity
Presentation of data

- Data display / visualisations
  - In imported format eg original transcript
  - Data to which code is attached (eg text units)
    - Reviewing content of data categories
    - Coding reports
    - Coding stripes
- Coding matrices
  - Counts of text
  - Displaying text in cells
- Hierarchical systems of major and subsidiary data categories
  - Illustrating data topics eg responses given to question
Presentation of data

• Data display / visualisations
  – Illustrations of networks
    • Linkages between concepts
    • Linkages between sources, project items
  – Graphs and charts
  – Tag clouds
• Key words in context
  • Word trees
  • Clustering
  • Proximity and sequence with other terms
Impact on data presentation

- Presentation in original context
- Illustration of co-occurrences of codes
- Demonstration of data support for propositions
  - Similarity and difference across groups
  - Co-occurrence of concepts
- Demonstration of face validity
  - of coding
  - of conclusions
- Illustration of dynamic analytic processes for handling data
Models of CAQDAS-supported research approaches

1. Homogenisation of program features and functions
2. Expanded functionality

- Data collection and creation
  - Create electronic versions of hard copy data files

- Data analysis
  - Develop electronic equivalents of manual analytic techniques

- Data presentation
  - Electronically present data as with manual techniques

- Models of CAQDAS-supported research approaches
  - Using CAQDAS to replicate ‘manual’ approaches
    - Examine types and volumes of data 'too hard' to manage without CAQDAS support
  - Using CAQDAS to undertake previously 'impractical' approaches
    - Undertake 'theoretically possible but impractical' analyses
  - Using CAQDAS to develop and execute CAQDAS-specific approaches
    - Analytical techniques for which computer functionality and support is essential
  - Presentation techniques for which computer functionality and support is essential

- Integration of data from multiple technologies

- Expanded options for data presentation incl multi-dimensional presentation
Next questions….

- How do the current technologies influence creative freedom?
  - What techniques do they support (or not)?
  - What technical wisdom is required?
  - What practical wisdom is developed?

- At what point do technical requirements dominate?
  - In research planning?
  - In research practice?

- What implications does this have for notions of ‘professional practice’?
Next questions…

- Which features, techniques and forms of wisdom are program-specific?
- Common across programs?
- Common across research approaches?
- Common across user groups?

- How can we best learn from user experiences?
  - What do we want researchers to explain, and how?
  - How do we want people to validate their methodologies?

- How do we best teach new users to develop their wisdom?


References


