Sensitivity and specificity of the URAM scale for Dupuytren contracture: A systematic review and meta-analysis

Diego Gomez-Herrero (diegogohe@hotmail.com)
Hospital 9 de Octubre

Rafael Sanjuan-Cervero
Hospital de Denia https://orcid.org/0000-0002-1958-0987

Pedro Vazquez-Ferreiro
Hospital Virxe da Xunqueira

Francisco Javier Carrera-Hueso
Hospital de La Plana

Jaime Eduardo Poquet-Jornet
Hospital de Denia

Emilio Garcia-Jimenez
Universidad de Granada

Research

Keywords: "Dupuytren Contracture", "Meta-Analysis", "ROC Curve", "URAM"

DOI: https://doi.org/10.21203/rs.3.rs-30621/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background

The URAM (Unité Rhumatologique des Affections de la Main) scale is a patient-reported functional outcome measurement tool validated for exclusive use in Dupuytren contracture (DC) a condition in which one or more fingers become permanently bent in a flexed position. The current study’s purposes were to evaluate: How sensitive and specific is the URAM scale for defining quality of life in patients with in DC and how sensitive to change is the URAM scale after treatment.

Methods

We performed a meta-analysis of all relevant articles published in PubMed, Embase, Cochrane, Google Scholar, LILACS and in various gray literature databases that describe the use of the URAM. We built three models: a hierarchical summary receiver operating characteristic (HSROC) model to determine the optimal threshold for defining DC, a difference in means models to assess the magnitude of the effect of different treatment modalities, and a meta-regression model to determine the effect on patient quality of life questionnaires according to variations in Tubiana scores after treatment (URAM).

Results

The HSROC and bivariate models showed a sensitivity of 80.23% (95%CI: 75.66 to 84.14) and an overall specificity of 2.61% (95%CI: 1.11–6.05). The second model showed an overall difference in means of 1.95 (95%CI: -2.86 to -1.04). The coefficient obtained in the meta-regression model was −1.666 (95%CI: -4.183 to 0.851).

Conclusion

The URAM scale is highly sensitive to changes in DC but has low specificity. It also showed a strong correlation with worsening of finger contracture.

Introduction

Dupuytren contracture (DC) is a fibroproliferative disease of the palmar fascia that can affect one or both hands. It causes progressive digital contracture that prevents patients from straightening their fingers, affecting hand function, performance of basic activities of daily living, and quality of life (Eaton et al. 2014).

Treatments include surgical procedures, such as fasciotomy, partial fasciectomy (FSC), and dermofasciectomy, and minimally invasive procedures, such as collagenase Clostridium histolyticum (CCH) injections and needle aponeurotomy (NA) (Eaton 2017). None of these treatments are curative and therefore many patients develop recurrent disease and require repeat treatment (Warwick 2017). Much research has been done on treatment outcomes in DC, with studies evaluating functional outcomes, patient satisfaction, and perceived quality of life using physical measures and self-report questionnaires (Engstrand et al. 2016). Patient satisfaction with treatment or with treatment outcomes is a multidimensional concept that can be difficult to define and measure and it also depends on aspects related to healthcare structure and delivery (Engstrand et al. 2016). In addition, the studies published to date have used different methods and asked different questions. Some authors have recommended using the URAM (Unité Rhumatologique des Affections de la Main) scale to monitor disease progression and treatment success and enable comparisons between different treatment modalities in DC (Ball et al. 2013).

The URAM scale was specifically designed to evaluate hand function in DC by the Rheumatology Department at Hospital Lariboisere in Paris, France in 2011. It is the first patient-reported functional outcome measurement tool validated for exclusive use in DC and has been shown to have adequate psychometric properties (Beaudreuil et al. 2011). It consists of nine multiple-choice questions, meaning it is sufficiently short and easy to use in both daily practice and clinical trials (Stromberg et al. 2016, Verstreken et al. 2016). It assesses patients’ perceived ability to perform a range of activities, including activities of daily living, and also addresses symptoms such as stiffness and loss of strength (Bernabé et al. 2014). Several studies have evaluated the reliability (good to excellent) and responsiveness of the scale (Binhammer 2018), which in addition has been adapted and validated for use in different languages (Beaudreuil et al. 2011).

The aim of this study was to perform a systematic review and meta-analysis of the sensitivity and specificity of the URAM scale in terms of its ability to define quality of life in patients with DC and its responsiveness to changes in disease severity following treatment.

Methods

The systematic review was performed according to the recommendations of Eden et al. (Eden et al. 2011) on review methods, data sources, and search strategies. We addressed two review questions:

1. Using the Tubiana scale as a reference test, how sensitive and specific is the URAM scale for defining quality of life in patients with in DC?
2. How sensitive to change is the URAM scale after treatment with FSC and CCH?

We performed a systematic search of PubMed, EMBASE, Cochrane, Google Scholar, LILACs, and Web of Science for articles published between January 1, 1990 and June 1, 2019.

The search criteria used in all the databases were combinations of the terms "Unité Rhumatologique des Affections de la Main", "URAM", "Dupuytren Contracture", and "Dupuytren".

Two reviewers (PVF and DGH) independently searched the databases and reviewed the articles retrieved. They also hand searched the reference lists of relevant articles and reviewed the gray literature to identify clinical trial reports and conference proceedings.

Clinical trials, cohort studies, and case-control studies that had used the URAM scale to evaluate DC were included. Authors were contacted when specific information on the use of this scale was missing. To minimize publication bias, no language constraints were placed.

**Study Selection**

Two researchers (PVF and DGH) independently screened the titles and abstracts to identify suitable texts, which they then reviewed in depth. When the researchers disagreed on whether a particular article should be included or excluded, the article was reviewed by a third researcher (FJCH) to break the tie.

**Data Extraction and Risk of Bias Assessment**

Working separately, PVF and DGH transferred all relevant data from the selected articles into standardized forms. The reliability of the entries was checked by another researcher (JEPJ). In addition to effect variables (mean [SD] pre- and post-intervention URAM and Tubiana scores), the data recorded included demographic variables (age, gender, and hand and radius affected) and variables for the stratification analyses in the meta-analysis (e.g., quality, language, study type).

As the studies included in the meta-analysis differed in type, their quality was assessed using the STROBE (STrengthening the Reporting of OBservational studies in Epidemiology) checklist (von Elm et al. 2014) applied separately by two researchers for each article. To minimize bias, a score of 15 or higher was used to identify high-quality studies. Discrepancies (i.e., differences in scores that placed a given study above or below the cutoff of 15) were resolved by a third researcher (RSC).

**Statistics**

Three meta-analysis models were used to answer the research questions: a hierarchical summary receiver operating characteristic (HSROC) model, a difference in means model for pre- and post-treatment URAM scores, and a meta-regression model adjusted for time since treatment.

For the HSROC model, tables summarizing Tubiana and URAM scores reported in each of the studies were created. In both cases, it was assumed that the scores were normally distributed. The data were then presented in $2 \times 2$ contingency tables with the URAM scale as the index test and the Tubiana scale as the reference test. The respective thresholds used were 2.5 and 1. Prevalence of DC was established at 100%. In other words, it was assumed that there were no true negatives, that it is that all the negative results for the reference test were false negatives. Enabling continuity correction, we then built a hierarchical multinomial regression HSROC model (Rutter et al. 2001), which converts the distribution of the two variables, allowing calculation of the overall ROC curve under the assumption that there is an underlying curve for each of the studies included. Each curve is determined by two parameters, $\alpha$ and $\beta$, which denote accuracy and asymmetry, respectively. Using these parameters and a $\theta$ parameter to denote the positivity threshold, distribution tables were generated for each study assuming that while the distribution of parameters would vary between studies, it would be normal and random (random-effects model). We then estimated the overall ROC curve together with the optimal threshold and corresponding confidence interval. The bivariate model was applied to directly model specificity and sensitivity based on the assumption that the Napierian logarithm of the odds ratio had a normal bivariate distribution in the different studies analyzed (Reitsma et al. 2005).

For the second model, standardized mean differences in pre- and post-treatment URAM scores were computed using Cohen's D and appropriate weighting. The most conservative model was selected in each case (Higgins et al. 2011). Differences of over 10% were considered to be clinically significant and the results were stratified by type of intervention (FSC or CCH). Each group was finally assigned an overall value.

For the meta-regression model, the dependent variable was change in URAM scores after treatment (differences in means before and after FSC or CCH) and the independent variables were Tubiana scores, time since treatment, type of treatment, age, and sex. The model with the greatest explanatory power was selected.

Heterogeneity between studies was investigated using the $I^2$ statistic, with high heterogeneity defined as a value of over 50% (Higgins et al. 2002). Potential sources of heterogeneity were investigated by subgroup analyses (study setting, language, ethnic origin), and the effect of outliers was analyzed in a sensitivity analysis in which studies were excluded one by one.

Analyses were conducted using the metan, metacum, metafunnel, and metandi features in Stata version 15. Differences in means were considered to be significant when the confidence intervals did not cross 0 and clinically significant when there was a difference of at least 10%. Publication bias was assessed using funnel plots and the Begg-Mazumdar test (Begg et al. 1994).
Funding and potential conflicts of interest

The whole investigation has been financed by own funds. Each author certifies that he has no commercial associations that might pose a conflict of interest in connection with the submitted article. All authors have completed and submitted the Conflict of Interest Disclosure Form and none were reported.

Results

Our search strategy retrieved 384 articles (Fig. 1) but 50 of these were excluded due to duplication. After screening the titles and abstracts of the remaining 334 articles, 313 were excluded (inadequate study design, missing data, different definitions of disease or disease severity, and publication in a language that could not be translated). Of the 21 articles selected for full-text review, 11 were excluded as they did not contain the information needed for our calculations (URAM scores, Tubiana scores, or degrees of contracture). Ten articles thus were included in the HSROC model, nine in the difference in means model, and 10 in the meta-regression model. The main characteristics of the studies are shown in Table 1.

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>N</th>
<th>Country</th>
<th>Type</th>
<th>Language</th>
<th>Age</th>
<th>Male</th>
<th>NA</th>
<th>CCH</th>
<th>Hand</th>
<th>R5</th>
<th>R4</th>
<th>R3</th>
<th>R2</th>
<th>R1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beaudreuil et al.</td>
<td>2011</td>
<td>53</td>
<td>France</td>
<td>CH</td>
<td>English</td>
<td>63.2</td>
<td>83.02</td>
<td>0</td>
<td>100</td>
<td>54.72</td>
<td>N.A</td>
<td>N.A</td>
<td>N.A</td>
<td>N.A</td>
<td>N.A</td>
</tr>
<tr>
<td>Scherman et al.</td>
<td>2013</td>
<td>93</td>
<td>Sweden</td>
<td>CH</td>
<td>English</td>
<td>67</td>
<td>86.75</td>
<td>50</td>
<td>50</td>
<td>53.49</td>
<td>55.81</td>
<td>37.21</td>
<td>6.98</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Warwick et al.</td>
<td>2013</td>
<td>254</td>
<td>International</td>
<td>CH</td>
<td>English</td>
<td>60</td>
<td>87.79</td>
<td>0</td>
<td>100</td>
<td>75.59</td>
<td>75.20</td>
<td>68.90</td>
<td>29.53</td>
<td>09.06</td>
<td>2.36</td>
</tr>
<tr>
<td>Bernabé et al.</td>
<td>2014</td>
<td>83</td>
<td>France</td>
<td>CH</td>
<td>English</td>
<td>63</td>
<td>83.02</td>
<td>0</td>
<td>0</td>
<td>N.A</td>
<td>N.A</td>
<td>N.A</td>
<td>N.A</td>
<td>N.A</td>
<td>N.A</td>
</tr>
<tr>
<td>Stromberg et al.</td>
<td>2016</td>
<td>140</td>
<td>Sweden</td>
<td>CT</td>
<td>English</td>
<td>67.5</td>
<td>85.26</td>
<td>51</td>
<td>49</td>
<td>58.57</td>
<td>51.43</td>
<td>35.00</td>
<td>06.43</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Broekstra et al.</td>
<td>2017</td>
<td>233</td>
<td>Holland</td>
<td>CH</td>
<td>English</td>
<td>65.5</td>
<td>65.35</td>
<td>0</td>
<td>0</td>
<td>N.A</td>
<td>N.A</td>
<td>N.A</td>
<td>N.A</td>
<td>N.A</td>
<td>N.A</td>
</tr>
<tr>
<td>Harrison et al.</td>
<td>2017</td>
<td>71</td>
<td>UK</td>
<td>CH</td>
<td>English</td>
<td>65.7</td>
<td>76.06</td>
<td>54</td>
<td>46</td>
<td>80.28</td>
<td>57.75</td>
<td>12.68</td>
<td>29.58</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Stromberg et al.</td>
<td>2018</td>
<td>152</td>
<td>Sweden</td>
<td>CT</td>
<td>English</td>
<td>66.5</td>
<td>85.27</td>
<td>50</td>
<td>50</td>
<td>58.34</td>
<td>51.28</td>
<td>41.67</td>
<td>07.05</td>
<td>N.A</td>
<td>N.A</td>
</tr>
<tr>
<td>Beaudreuil et al.</td>
<td>2018</td>
<td>30</td>
<td>France</td>
<td>CH</td>
<td>English</td>
<td>72</td>
<td>83.34</td>
<td>100</td>
<td>0</td>
<td>N.A</td>
<td>N.A</td>
<td>N.A</td>
<td>N.A</td>
<td>N.A</td>
<td>N.A</td>
</tr>
<tr>
<td>Wiseman et al.</td>
<td>2018</td>
<td>136</td>
<td>Australia</td>
<td>CH</td>
<td>English</td>
<td>66</td>
<td>80.15</td>
<td>0</td>
<td>100</td>
<td>75.18</td>
<td>N.A</td>
<td>N.A</td>
<td>N.A</td>
<td>N.A</td>
<td>N.A</td>
</tr>
</tbody>
</table>

Study: study author; Year: date of publication; N: number of patients; Country: country in which study was performed; Type: study design (CT: clinical trial, CH: cohort study); Age: mean age of patients; Male: percentage of male patients; Fasciotomy: percentage of patients who underwent fasciotomy; NA: needle aponeurotomy; CCH: percentage of patients who received Clostridium histolyticum injections; percentage of right hands operated on; R5: percentage of interventions involving the fifth radius; R4: percentage of interventions involving the fourth radius; R3: percentage of interventions involving the third radius; R2: percentage of interventions involving the second radius; R1: percentage of interventions involving the first radius.

Application of the HSROC (Rutter et al. 2001) and bivariate models showed an overall sensitivity of 80.23% (95% CI: 75.66 to 84.14) and specificity of 2.61% (95% CI: 1.11 to 6.05). The diagnostic odds ratio was 0.109, (95% CI: 0.041 to 0.292), with a positive predictive value of 0.824 and a negative predictive value of 7.546 (Table 2). The reference value together with its 95% prediction region is shown in Fig. 2.
Table 2
Parameters computed for the HSROC and bivariate models.

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter</th>
<th>Coef.</th>
<th>Std. Err.</th>
<th>z</th>
<th>P &gt; z</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bivariate</td>
<td>E(logitSe)</td>
<td>1.401</td>
<td>0.136</td>
<td>1.134</td>
<td>1.669</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E(logitSp)</td>
<td>-3.616</td>
<td>0.446</td>
<td>-4.490</td>
<td>2.742</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Var(logitSe)</td>
<td>0.036</td>
<td>0.051</td>
<td>0.002</td>
<td>0.577</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Var(logitSp)</td>
<td>0.350</td>
<td>0.652</td>
<td>0.009</td>
<td>13.436</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Corr(logits)</td>
<td>1.000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>HSROC</td>
<td>Lambda</td>
<td>0.428</td>
<td>2.376</td>
<td>-4.228</td>
<td>5.085</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Theta</td>
<td>2.261</td>
<td>0.160</td>
<td>1.948</td>
<td>2.574</td>
<td></td>
</tr>
<tr>
<td></td>
<td>beta</td>
<td>1.138</td>
<td>1.034</td>
<td>1.100</td>
<td>3.164</td>
<td></td>
</tr>
<tr>
<td></td>
<td>s2alpha</td>
<td>0.449</td>
<td>0.580</td>
<td>0.036</td>
<td>5.631</td>
<td></td>
</tr>
<tr>
<td></td>
<td>s2theta</td>
<td>0.000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Summary pt.</td>
<td>Se</td>
<td>0.802</td>
<td>0.022</td>
<td>0.757</td>
<td>0.841</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sp</td>
<td>0.026</td>
<td>0.011</td>
<td>0.011</td>
<td>0.061</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DOR</td>
<td>0.109</td>
<td>0.055</td>
<td>0.041</td>
<td>0.292</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LR+</td>
<td>0.824</td>
<td>0.027</td>
<td>0.773</td>
<td>0.878</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LR-</td>
<td>7.546</td>
<td>3.601</td>
<td>2.961</td>
<td>19.228</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1/LR-</td>
<td>0.133</td>
<td>0.063</td>
<td>0.052</td>
<td>0.338</td>
<td></td>
</tr>
</tbody>
</table>

Covariance between estimates of E(logitSe) & E(logitSp) 0.0174904; Log likelihood −32.050571

In the differences in mean model, we obtained a value of -1.30 (95% CI: -1.77 to -0.83) for FSC and -2.75 (95% CI: -4.73 to -0.78) for CCH. The overall value was -1.95 (95% CI: -2.86 to -1.04) (Fig. 3).

The only significant variable in the meta-regression analysis of the influence of variations in Tubiana scores on URAM scores was time between treatment and completion of the URAM questionnaire (Table 3). The coefficient for this model was 1.666 (95% CI: -4.183 to 0.851).

Table 3
Meta-regression.

<table>
<thead>
<tr>
<th>% residual variation due to heterogeneity</th>
<th>Tau² = 3.18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of between-study variance explained</td>
<td>( \hat{R}² = 75.40% )</td>
</tr>
</tbody>
</table>

| Coef. | Std. Err. | t      | P > |t| | 95% Conf. Interval |
|-------|-----------|-------|-----|---|-------------------|
| Tubiana | -1.666 | 0.979 | -1.70 | 0.150 | -4.183 | 0.851 |
| Time | 0.014 | 0.009 | 1.55 | 0.182 | -0.009 | 0.037 |
| Constant | -13.392 | 2.713 | -4.94 | 0.004 | -20.366 | -6.417 |

Discussion
Our meta-analysis showed high sensitivity (80.2%, 95% CI: 75.7–84.1%) and low specificity (2.6%, 95% CI: 1.11–6.1%) at the optimal threshold for the URAM scale as a diagnostic test for DC. The scale was also very sensitive to change after treatment with FSC and CCH, with an overall difference in means of -1.95 (95% CI: -2.86 to -1.04). It was more responsive to changes after CCH. The scale was also able to capture the effect of time since treatment, although to a lesser extent (nonsignificant coefficient). The high sensitivity observed is to be expected as the URAM scale contains several questions on hand mobility. It is also consistent with the strong correlation observed between URAM and Tubiana scores in previous validations (Beaudreuil et al. 2011) and with high sensitivity values reported for the DASH (Disabilities of the Arm, Shoulder and Hand) (82%) and QuickDASH (79%)
questionnaires, although in these cases, meaningful change was measured using a very different methodology (Franchignoni et al. 2014). The low diagnostic specificity observed (2.6%) is much lower than the rate reported for DASH (overall specificity, 74%) (Franchignoni et al. 2014), although to our knowledge, the diagnostic accuracy of DASH has not yet been analyzed in the specific setting of DC, but rather in studies analyzing different diseases (Gummesson et al. 2006) or responsiveness to changes after corrective surgery for DC (Rodrigues et al. 2017). Its diagnostic specificity for DC thus is unknown and is probably much lower than rates reported for general upper arm disabilities. Notwithstanding, the URAM scale has been reported to outperform other scales in terms of its specificity for DC (Knobloch et al. 2012). Although we do not have data to confirm this superior performance, we did find a relatively good balance between sensitivity and specificity (around 60%) in the underlying ROC curve (Fig. 2), suggesting good disease-specific performance in DC.

The low specificity observed for the URAM scale in our meta-analysis has several explanations. On the one hand, our model was based on a series of assumptions, including the absence of false positives (perfect specificity) and the thresholds used to define DC. There is currently no agreement on where the line between disease and recurrence lies, although some progress is being made (Harrison et al. 2017). On the other hand, the URAM scale was specifically designed for DC, but it could theoretically be used in other diseases such as carpal tunnel syndrome, as it addresses hand mobility problems that are not specific to DC. Question 9, for example, evaluates problems with pinch, which is generally a greater problem in patients with carpal tunnel syndrome than in those with DC, who face more difficulties straightening their fingers. Additional sources of heterogeneity in our model are the diverse criteria used to measure contracture (Pratt et al. 2016), the different time points at which treatment outcomes were measured, and even doubts about the applicability of the scale in different languages and cultures, although high consistency has been reported for the validated versions of the URAM scale in several languages.

The limitations of this study are linked to the varying degrees of contracture severity in the samples analyzed, potential selection bias, and potential information bias as the URAM scale was not designed as a diagnostic test.

Conclusions

This meta-analysis shows that the URAM scale has high sensitivity and low specificity for DC, although it was sensitive to clinically significant changes following treatment.

Abbreviations

DC
Ducpuytren’s Contracture
URAM
Unité Rhumatologique des Affections de la Main
CCH
Collagenase Clostridium Histolyticum
NA
Needle Aponeurotomy
FSC
Fasciectomy
STROBE
STrengthening the Reporting of OBservational studies in Epidemiology
HSROC
Hierarchical Summary Receiver Operating Characteristic
Declarations

Competing interests

Not applicable

Funding

Not applicable

Authors' contributions

Conception and design: PVF, DGH, FJCH. Collection and assembly of data: DGH, PVF. Analysis: PVF. Interpretation of the data: PVF, DGH, FJCH, RSC, JEPJ, EGJ. Drafting of the manuscript: DGH, PVF. Critical revision and final approval of the article: PVF, DGH, FJCH, RSC, JEPJ, EGJ.

Acknowledgements

This article is part of the Doctoral Thesis of Diego Gómez-Herrero within the Doctoral Programme in Pharmacy at Granada University in Spain.

Availability of supporting data

Data relative to this work is available under request to the correspondence author.

Ethical Approval and Consent to participate

Not applicable

Consent for publication

All we (the authors) understand that the information will be published without we/we child or ward's/our relative's (circle as appropriate) name attached, but that full anonymity cannot be guaranteed. We understand that the text and any pictures or videos published in the article will be freely available on the internet and may be seen by the general public. The pictures, videos, and text may also appear on other websites or in print, may be translated into other languages or used for commercial purposes. We have been offered the opportunity to read the manuscript.

References


Figures
Figure 1

Flowchart of study inclusion

383 records identified through database searches
1 additional record identified through other sources
334 records after removal of duplicates
334 records screened
313 records excluded
21 full-text articles assessed for eligibility
11 full-text articles excluded
10 studies included in qualitative synthesis
10 studies included in quantitative synthesis (meta-analysis)

Figure 2

HSCROC curve with 95% confidence region and empirical Bayes estimate and 95% prediction region
Figure 3

Meta-analysis of differences in URAM scores according to procedure

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- PlainEnglishSummary.docx
- PRISMA2009checklist.doc