ƒ ‹F | Œfia‹ ˙'e›ß• |–M^D‹ ª¤s›p e ƒ¤"G‡ł§i

« ¥ ´ ˙––¤ |⁄W""¥˝´ †"cƒ¡»P«D †"cƒ¡,Œfi˘⁄§" ^ « "c

»P" †z¤t† ¸‹ ª o(I)

›p e^ §¡G §O«‹›p e

›p e‰ ¸s„¡GNSC92-2524-S-038-001-

ıƒ · ¶¡¡G92 ƒ~05⁄°01⁄ ∅ 93 ƒ~04⁄°30⁄Ø

ıƒ ‡ ƒ ¡G» ¥_´ ˙⁄j ˙´ ˙,Œ T‹ ª¤s'

›p e¥ D ø⁄¡HG‰ – ¥⁄fl

ƒ ∅P¥ D ø⁄¡HG… B B'M

‡ł§i^ «‹¡G" ° †‡ł§i

‡ł§i" ¥ ¡G¥ Xu Œ» • |˜‡‹ ª¤s⁄ –o‡ł§i⁄ ¸ o" ‰⁄

‡B†z⁄Łƒ¡¡G¥»›p eflA⁄ ¸–M§ Q ' ¸ ¤ ¥ L · …z1 ƒ‡‹£Æv¥Ai ⁄ ¶‰}‹d,

⁄⁄ ¥` Œ 93 ƒ~9⁄°9⁄Ø

« ¥ ´ ·—¤ |/W"″¥″´ †″cf ¡»P«D †″cf ¡,Œfi˘/§" ^ « ″c »P″ †z¤t† ¸< ª o

Building a biomedical structural and unstructured data collection and management system for medical education

‰ − ₩/fl

I-Jen Chiang [ab]

[a]¥ ¥_´ ·/j ·´ ·,Œ T< ª¤ ‡ .[b]¥x˘W/j ·´ ·/ u{ ·< ª¤ ‡

¥ D ø/:H‰ − ₩/fl › p ®s:,N,SC-92-2524-S-038-001

/@¡B/// ″K›n

¥»› pef ffibx« ¥ /@" ^ » ' ¿ ª z¤^″ †z" ^ "″¥ ¥x¡ A fl » '¶ ˜ m ¤ ^ t ¥„Œ» fi ¤ • j M¡A¤ ·§ψ ·¥″' ¸ f ⎪ fi ¶i f {§ ·† ¡ C¸¡B¶ É_»P† B / Ł¥‹†″c /˘"″,Œfi˘¡ A fl f œ†¡Ł§i ¯æ¡B/¿ ˜ m ¥‹ ‹ / f rO«D †″cf ¡,Œfi˘¡A›n¤ §U´ · ⎪ · ¥″' ¸ n ⎪ fi v ¥ ¤ft† ¸"″ ·† ¡ A Nƒb' − N ⊕¤,·Œ T¥ † § · / ƒ¤ƒ¡"″† Wh" ^ ¡A • f t ` { § ›fiEyidence-based /§ case-based ƒ¡/§ ·† ¡ A^└N ¤œ"™ ^ ¥H ¥−/˘™ ^ ″ „˘¥₩H§ ⊕{¡A· ‹ ` § ´ fi fl † z L„," ^ « ″c/¸″ †z¤t† ¸¡ Ab› ‾ flf− fi fl · £ ¤ § † ¥B″¸°‰fl‹»P‹ ^ • § ' /"RA„´¥ /@¤·‹ ^ / ˜m"″ƒ î ¥H § @ ¥„flf− w †¤"¸"†B,m¡C

¥»× ª¤s¤t† ¸‹ [″c/ƒ†¤ › ›«›n.. † ¡G ¡]1¡^‹ ƒ °˙ "″ V ₧...† ¡]2¡^‹ ¶ ¥h ƒ¡" ^ / ^...† A‹¥H¤ ˃ (Bayesian)'w†z ‹ ¥D›n/Ł"k †″c¡F¥H¤ • § −N/ ¥ ₩¶q˘ „ ' ¤ ¥¸"″ fiw¶ ƒ /æ„ § ¥Xƒr ^ `p '°¡A§ ₩¸¤t† ¸/"″/ ¥ ƒ °˙ § †N¡A„g ` § − M fi œ ƒ¡ ·† ' †£¥˝"″/ ^ † W k¡A • ·‰T¶ ƒ / ˜m"″/ ^ ¡A· ₤ ¤ ˘ ¥"fl ƒb† † ƒh,Œfi˘/″ ‰T− ø ' » / ˜ ¶C

¥»× ª¤s− ˜ ¥ ¸ , ß"¡.A'• ¥H¤^¥...−N' ƒ†,Œfi˘fiw«D‹ ^ / ˜mfl ˃J/ ¥ ¡]...¸¥»¥ ...˘¡^ ›p″ ¡Aƒ ‹ ΘN,g¥ ^ ` ¸fl «Æ' »sƒ¤"″ ,Œfi˘fiw/˃/ ^ ¤ œV.‰¸₤¥»¡]305 ‰g/ ˜m¡^¡A„g −M• ~` § ´ fiv„ ƒ„¤t† ¸¶ ƒ ...¸¥» ‰m«Æ¡A „ · ø ...¸¥»¡]108 ‰g/ ˜m¡^¶ ƒ ƒß ¡A †"G o†{˘ „ ¤ ¥¿‰T†v "„F95.4¢M¡C

^ ` ¡G" ^ ″ †z, / ¥ − · , ´ ·—¤ |, ´ ·/ ˜m

ABSTRACT

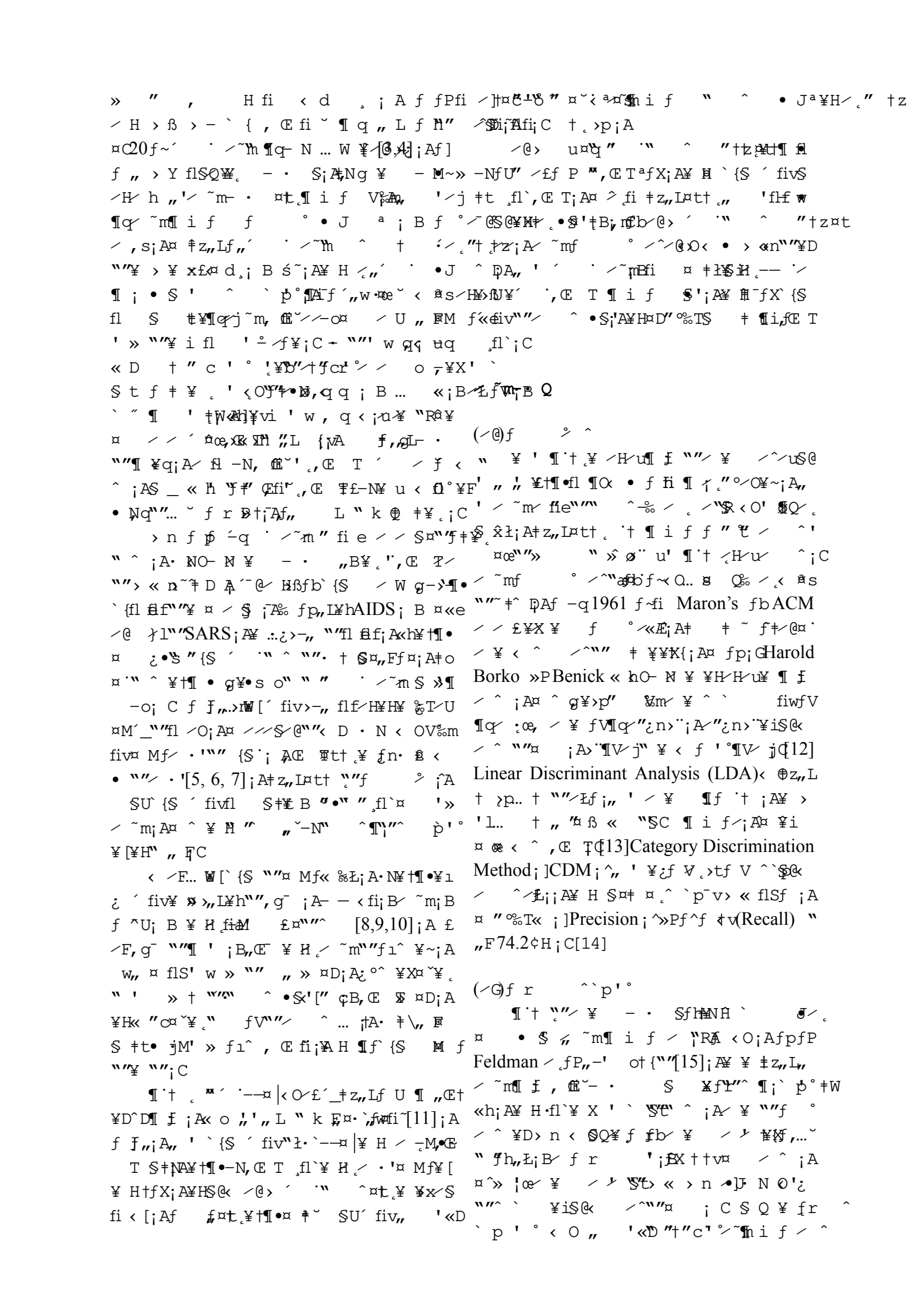
fi Medical knowledge needed for physicians in clinical practice is getting complex and rapidly expanding nowadays. In order to provide an optimal patient cares in clinical practices, physicians need to get concise and precise evidences from information in time. This framework will design and develop a knowledge management system for the structured data management.

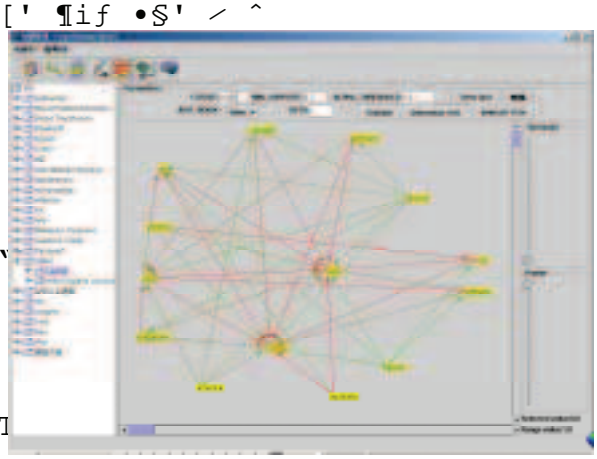
Keywords: Knowledge Management, Text Mining, Medical Education, Medical Literatures.

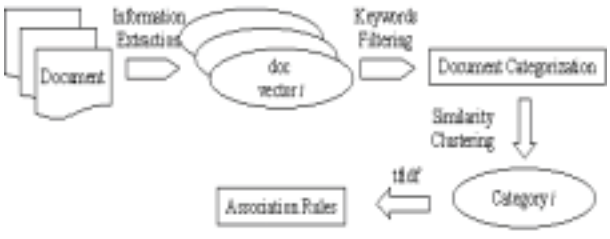
″ ‰

ƒ] ´ · / W / j † ¡ / " ¶E´_»P†B,m¡A • ¥ " / £ ‰ T ' w ' ° / U ¶ i ƒ ¤M´_¡A• ´ fiv − w / U „ F ¶ E ´ _ ¤ M ƒ fi A¶•ƒæ H ƒU' ' ° ¡ A " H ‹ „ / § ₩›¨¡A§@¥ifl ¥ H ¤ M ' w / U / @ ¤ B " ″ ¸ ¯ ' ¸"v Ł†B,m¡E − Ł ¥ † v « ‹ " ″ § P ´ _ „ L{˜Y Mƒ¤«‹ ¡A ‰ ... ƒ ¡ ¥ ¿ ‹ ¢ ƒ n ¥ i ″ ¡ ¤ ‹ ƒ/@ƒb/£‰T'w¡ " p / U ¶ƒ‰/ § / Ł " k ¡ A ¥ » › pe−N¥H¤'/ „ § ‾¸ ‰ ... ƒ ¡ ¥ ¿ ƒ ¡ « ″ c ª › ´ ·—¤ |/W" ´ † ″ c ƒ ¡ » P « D † ″ c ƒ ¡Œfi˘/§" ^ « † z ¤ t † ‹ª o¡C

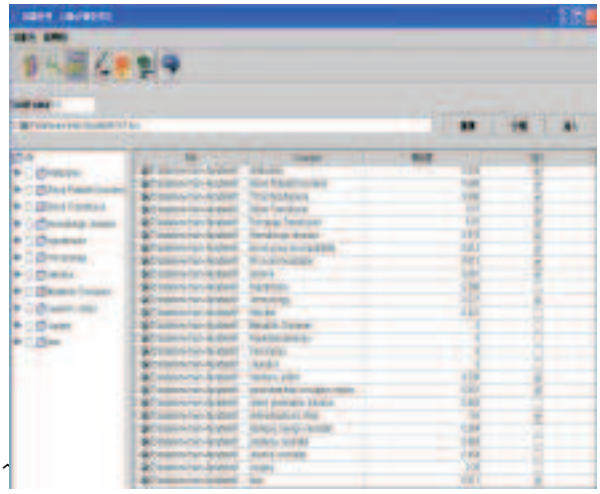
• − d • ¯ U " ″ « ~ ‰ Ł / @‰‹ O†{¥N§ › ¤ D " ″ ¡ A ¤ ƒ n ^ a « h ¤ M ' ₩ ´ fiv"″¤M ¤ M ƒ " ″ ƒ ¤ fi ˜ « h § „ ¥¥¶[¶¡"´ ·" ^ "«´ · " ^ ¶ƒ/¡† ' ¤ †t"″† ¿n" ″ ·/ m̃/¡/ A ´ / ˜m‹ O†z„L† }h˜Y´ "″` { § „Œ ... ¶gƒ ƒ ¤ " ″ « D ″†c'G¤,Œfi˘¡A ˜ ›^/j¶q" ″ " ^ [2₵ H „ q , £ ¥H/ ″¸ » ″,"″· ¶¸¡ A ´ / ˜m, Œ˘fi w /] ₩¥ ¶· †¸¥ œ ″ − ƒ¤‰u/W„ q ₤fif ¡A¤^¥i†z„L″






Information Extraction
Document
doc vector i
Keywords Filtering
Document Categorization
Similarity Clustering
Category i
Association Rules

HTML/ ¸PDF¤ ˆ – N ¸ fl ` † " G ¤ ª ¡A¥»¤ct¤¸f¥ °/ ˆ
/ W › z , Œ fi ˘ fi w /f1736¶ gR‹ ø ˆ ' ¿Ø , ¤ t¸¤ ‹ ¤s" „ / ¥ ‹ f «›¨
f ´ ¨¥H p ¤ ‹ ¡] § t ‹ A L¤ ¡(Cut off)¤]‹¡B …˘›¨¶ f / ˆ ¡ A ‹¤"s ¶•–N
f ¤¥H¸¥ fi £ ¡/¡ Am¤¥† o ¤ ¸ ˜m – / ˆ "/, Œm¡]¥i‹ † ‰ gBfh‰g/ ¸/@
¤ œ / ˜ m §‹/ ¸ V‰mf ¤ t † ¸ † ' ›o"Œ˘§¤¡ ˆ´ ¡ i ´ s ˜ j«¤ ¶tsf¡Af/A
ˆ † Wh¡C ¥ ¡i/ ˆ ¡ j «¶¶s f f °/ ˆ ¡A¤ t¸• |w
„ ‹ ¤s" ' ¿Ø/ J§cut off ›¨¡A¶ f ˆ ¥ ´k
ˆ ¡A¤ ¤ t¸/ ¶–f p˘/ ›' ¥ ¡C
„˘/ › f °/ ˆ / ¶–



ˆ ¥ ¿

¤ ¥ NCBI' ¤ ¥¸"'"MeSH‹ ¥»× [
"¢¡AMeSH‹ O¥H¥ D ˆ D§§@¥ ¶b"'"¶ ¥ h.f¡
" R ¥ DD¥.–¡Afb´ ¨» / /† ¤ – "'"¥D
ˆ D–¥ ¡A‹ Θ˜¥¸¶ ¨† "¸"– – ¤ •J¥ i … W[„
‹ Y §'"'"¥¸ ¶ ¡ „/@A¤° §U¤ ¥¸" §f
• § ' ›«¡Afb/ p ¤ ¿Ø»f / / Af@D–o7/j
ˆ 48/ p̂ ¡Af ¤fl°I«h‹ O¤ •J»P¤ ¥¸"
• § ¤ˆ« D §¥‹ fie¡A– – ¤ •J¥¸ ' „ „/]
/ £ •s¿o¡A/ f]» " ˆ ' » ˆ ¥ / £ ‹
fP¡Af ›P ¥ „ Eh/ ¸/ £ ¤ ‹†{¶A g¥ ,Œ
† 'f fiw/u§@/ H‰¥H/ ¸` § / p ¤‹¥D"v´ fiv
f@fP„ ' ¤ f†ˆ ' ¿ Ø /¸/ p ¤ ‹‹ ˆ
ˆ ¥ ¶ f …W· ¡ A@¿ 10/jˆ 21 › – M~†N
»y§ @ ‹ / ‹ˆ "c¡Af p „¥'| ¥ ¡A¥H† ¯ ¥–M
•~´ fiv„ …¸¥»/ ˜ m " " ¥ ¡C

« f¡ ˙†

¥D›n‹ V ‰¤‹t¥¸• § / ˆ ¡A, g–M
•~` § ´ fiv„ f„¤ t¸¶ i f /H«u f¡ †
…¸¥» V ‰ m] .V‰¥»…˘‹ 305‰ g / ˜m«Æ¡A
/O¤D¤Cˆ ¥ ‹ f†‹ ˆ / ¡Am g¥ ¤ t¸f °
/ ˆ ¡A' –o¤ "'"· œ …¸¥»¡]· œ …¸¥»…˘‹
108‰ g / ˜m¡† G¡fAA¥ ¥ , f – M~` § ´
fiv„ ¤ t ¡' "'"· œ …¸¥»¶ f ß ¡A¤ '
–o "¶G¥ Hß´_, ¤ t¸, g« f¡ ˙ † V‰Æm
' – ø •ˆ § »P, ` § –Mfia"'"– M~• § ‹
f {« ¡A¥ H ß f¤,tg«¸ f¡ † «Æ"'"¥i
«H« ¡C

"¶G

¡]/@¡ ˆ t¤†¥¸i«H« ß

/ " R · œ / ¥ ¶ f / "¡"'"• ‰T' °»P
/@›P' °¡]inter-coder agreement¡ ˆ ßf¡A–˜¥¸
Kappa Statistics /u¤ ª¡AKappa ›¨¥D›n‹ O– "¥¸
¤ ß › 'l„ ¤ t¸¶ f · œ "'"` § ´ fi »P
‰m«Æ¤ t¸„ · œ / ¥ "¶G"'"¥i«H« 'M/@
›P' °¡Cf]f„¡A¥i§ @ '…f „ œ," /§¶ ¡ /
¥ /"fie' †£¥ ˝ " " ‹ˆO§_/@›P¡Cf]f„› p

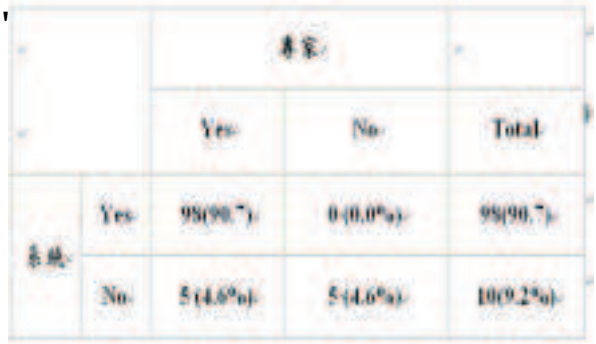
Antibodies
Blood Platelet Disorders
Thrombocytopenia
Blood Transfusion
Exchange Transfusion
Hematologic Disease
Blood Group Incompatibility
Rh Isoimmunization
Anemia
Hypotension
Immunology
Infection
Infant, Newborn Disease
Persistent Fetal Circulation Syndrome
Infant, Premature Disease
Erythroblastosis, Fetal
Epilepsy, Benign Neonatal
Asphyxia, neonatal
Anemia, Neonatal
Surgery
Twin

„ ¥| ¶ ¥ h.f¡ ¥

Kappa › ¨ ¤ ¶ i ƒ ß ƒ » P ⁄ "R'
ƒ p " ⁄ @ ¡ C › p "⁄‰ƒ¡ƒp⁄U¡G

$$K = \frac{Po - \Pr}{1 - \Pr}$$

Po ‹ ˘ [„ ›Pr¡‹A H › ¨Al ‹ ¯[„
›¨† ⁄j›¨¡C



		Yes	No	Total
系統	Yes	98(90.7%)	0(0.0%)	98(90.7%)
	No	5(4.6%)	5(4.6%)	10(9.2%)

" ⁄@ Kappa ›¨ ß ƒ ⁄"+Q› P ' ° ™ƒn^a • – ˜ «ƒ ‹ 0.5¡]cut off¡ 0.5¡ ^ ¶ƒ

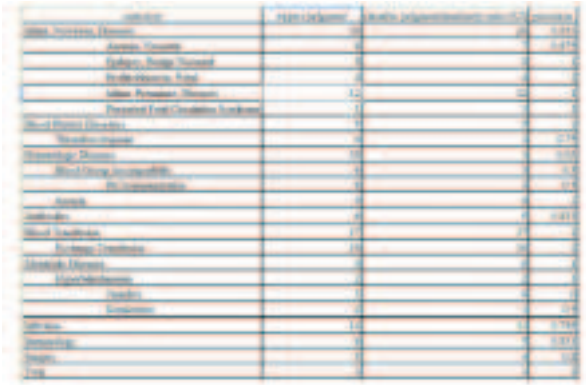
Kappa	¥i«H« ß»ø
0.00	' ƒH
0.01-0.20	• fiz
0.21-0.40	¥i a
0.41-0.60	¥i«H
0.61-0.80	› «n
0.81-1.00	S ƒ„

· œ fi ¡A` { S ´ fiv»{‹ ⁄ †„„ g ¤ t †¶if
« ƒ ¡ † ¡AS^ «Æ o†108 ‰ g ⁄˜m
• S@· œ , …¸f¥›¡A´ fiv»{‹ ƒ†103 ‰ g ⁄˜m
¥ i ¥ H‡Q˜k¡Aƒ ƒ†5 ‰ g ⁄ m´¤ fiv»{‹
ƒ† A‡X" " ¥ ¥i¶ ƒ ´k^ ¡ F ¥‰H‡Q k' ,
^ ¥ ⁄" " ⁄˜ƒm†98 ‰g(89.8%)¡A { S ´ fiv
– o ¤ t † ¿ ¥ ⁄S·""ƒ†5 ‰g(4.6%)¡A M
¥tƒ† 5 ‰ g ⁄˜(4.6%) ¡A¤ ‹ ƒ « ‹ SC'
0.5¡A¤^¥…†Q¤ t † ¸ ¶ ƒ ¡A‹^Œ'‹ ´ fiv»{
‹ «D‹ ⁄ p ¤ ¿ Ø » ^ ⁄ ˜m¡C' ¥H– N
fiv˜— o ¤ t †⁄¥ ‰T""98 ‰ g ⁄˜¥n{W´ fiv
˜—o L"k†Q⁄ ^"5 ‰ g ⁄˜m¥i˘ „ ¤ ¥¿
‰T† v "„Ø5.4¢M

Observed agreement = (98+5)/108 = 0.954
Random agreement = 0.907 * 0.954 + 0.092 * 0.046 = 0.869

Kappa = (0.954 - 0.869)/(1 - 0.869) = 0.649

, g – M • ~ ` { S ´ fi v „ ƒ¤t†¸¶if …¸¥» V‰m
¡]305 ‰ g ⁄ ˜ m ¡ ^ « Æ ' – ø 1"08·œ, …¸¥»¡]
‰ g ⁄ ˜ m^¡A"Æ ¥ ¥ ⁄Ø ƒ – M • ˜{S ¡¡Clever Craft¡ j ƒ¥"™Kappa ›¨⁄j' 0.6¡A
fi v „ ¤ t † ¶if‡A→œ · œ , … ¥»¶if „ ⁄ ¥ ⁄ ^ ¤ t † ¸ g V‰m«Æ»P`{S
ƒ ¡A` { S ´ fi v „ ' ƒ U ‰ g ¥ ⁄ ^ »P¤t†¸g
« ƒ ¡ · † «Æ¶A¤ƒ ⁄' ^ – o†™GƒpP ¡] G¡ØƒÛ ¥ " • ·«
⁄ G '¥¡C ¥ » ¤ ‡¸⁄g ‹ Ø"ý fi v V‰m«Æ»
† ¸ Ø–fiMa„ · œ , …¸¥»¶Uƒ " • ·« ·œ
" ⁄G ` { S ´ fi v » P ¤ t †⁄ ^ fi˜« ß ƒ ¸ ¡A¥ ¤ t † ¸⁄ ^ '–o‡U‰g‰S ¥¥
¡]›Y⁄"fie‹ , ^ ¥ ¡A«h⁄£›› ^ ¥ … ¶q¡ ¥
¥ , ` { S ´fiv_¤ß ⁄^• ·‰T'°¡A † GƒpP
" ⁄T' ¥ ¡A ф{¤ "‰T« ‹ • "¡A⁄ }N‹O
»¡• ¤ t † „g¥Gƒ ` { S ´ fi v ¶f « ƒ¡ ·
† «Æ¡A¤ ⁄^• S ‹ • „pƒ P • " „ ‡† ¸
¶ i ƒ V‰m‡ß"" • S ¡A' ¥H, ¤ t † ¸ ™«
· † ƒ⁄…˜† ¤ ¥i«H« »Æ "¡C



[1] W. M. Tierney, M. E. Miller, J. M. Overhage, C. J. McDonald, Physician order writing on microcomputer workstations. JAMA 1993; 269: 379-383

[2] EBMR ... Available at: http://www.hint.org.tw/family/research/ebm-1.htm Accessed Sep 1, 2003

[3] P.N. Gorman, J. Ash, and L. Wykoff, Can primary care physicians' questions be answered using the medical journal literature?. Bull Med Libr Assoc, 1994. 82(2):140-146.

[4] J. C. Wyatt, Knowledge management and innovation in medicine: how to go beyond practice guidelines? Advances in Clinical Knowledge Management, 2002; 5.

R. Smith, What clinical information do doctors need? BMJ, 1996. 313(7064): p. 1062-1068.

[5] B.E. Barnes, Creating the practice-learning environment: using information technology to support a new model of continuing medical education, Acad Med, 1998. 73(3): 278-281.

[6] M. Frize, F. G. Solven, M. Stevenson, B. G. Nickerson, T. Buskard, K. Taylor, Computer-Assisted Decision-support

Systems for Patient Management in an Intensive Care Unit. Proc. Medinfo '95 1995; Vancouver:1009-1012. 14.

[7] R. C. Shank, Case-based teaching: Four experiences in educational software design, (Technical support No. 7). Institute for Learning Sciences, Northwestern University, 1991.

[8] E.J. Huth, The underused medical literature. Ann Intern Med, 1989. 110(2): 99-100.

[9] D.G. Covell, G.C. Uman, and P.R. Manning, Information needs in office practice: are they being met? Ann Intern Med, 1985. 103(4): 596-599.

[10] P.N. Gorman and M. Helfand, Information seeking in primary care: how physicians choose which clinical questions to pursue and which to leave unanswered. Med Decis Making, 1995. 15(2): 113-119.

[11] B.E. Barnes, Creating the practice-learning environment: using information technology to support a new model of continuing medical education. Acad Med, 1998. 73(3): 278-281.

[12] H. Borko, and M. Bernick,. Automatic document classification. in ACM, 1963. 10(2):131-135.

[13] K. Fukunaga, Introduction to statistical pattern recognition (2nd edition). (New York, 1990).

[14] R. Feldman, Mining unstructured data in ACM SIGIR, pages 182-192, San Diageo, CA,1999

[15] R. Feldman, Y. Aumann, A. Amir, W. Kl´osgen, and A. Zilberstien. Text mining at the term level. In Proceedings of 3rd International Conferenceon Knowledge Discovery, KDD-97, pages 167–172, Newport Beach, CA, 1998.

[16] Saracevic, T. A Study of Information Seeking and Retrieving. Background and Effectiveness. Journal of the American Society for Information Science, 1988. 39(3):177-196.

[17] F. Sebastiani, Machine Learning in Automated Text Categorization. ACM Computer Survey, 2002.34(1):12.

[18] fl ` ⁄ ¤ ⁄ ¸ ¤ ' , Œ T ¸ fl ` ⁄ W " " – · Q ¡ A ¥ x ⁄ j „ ˇ fi ‹ ª ¤ Available at:public.ptl.edu.tw/publish/suyan/36/text_46.html Accessed Sep 1, 2004