

ORIGINAL ARTICLE

J Korean
Neuropsychiatr Assoc
2015;54(2):228-235
Print ISSN 1015-4817
Online ISSN 2289-0963
www.jknpa.org

북한어 정신건강 평가도구 개발 및 그 신뢰도

울산대학교 의과대학 서울아산병원 정신건강의학교실,¹ 국립서울병원 정신건강의학과,²
통일부 하나원,³ 국립중앙의료원 정신건강의학과,⁴ 서울대학교 의과대학 정신건강의학교실,⁵
성균관대학교 의과대학 삼성서울병원 정신과학교실⁶

이태엽¹ · 안병희¹ · 전진용² · 한정미³ · 이소희⁴ · 함봉진⁵ · 조맹제⁵ · 홍진표⁶

Development of North Korean Version of the Composite International Diagnostic Interview

Taeyeop Lee, MD¹, Myung-Hee Ahn, MD¹, Jin Yong Jun, MD²,
Jeong Mee Han, PhD³, So Hee Lee, MD, PhD⁴, Bong-Jin Hahm, MD, PhD⁵,
Maeng Je Cho, MD, PhD⁵, and Jin Pyo Hong, MD, PhD⁶

¹Department of Psychiatry, College of Medicine, University of Ulsan, Asan Medical Center, Seoul, Korea

²Department of Psychiatry, Seoul National Hospital, Seoul, Korea

³Settlement Support Center for Dislocated North Koreans, Ministry of Unification, Seoul, Korea

⁴Department of Psychiatry, National Medical Center, Seoul, Korea

⁵Department of Psychiatry, Seoul National University College of Medicine, Seoul, Korea

⁶Department of Psychiatry, Sungkyunkwan University School of Medicine, Samsung Medical Center, Seoul, Korea

Objectives The aim of the study was to develop a North Korean version of the Composite International Diagnostic Interview (NK-CIDI) and to evaluate the reliability and validity of NK-CIDI.

Methods Subjects were 100 North Korean defectors, aged 18 to 65, who were recruited in a center for defectors in the National Medical Center or who were on an education curriculum in the defector settlement support center. NK-CIDI was developed based on the Korean version of CIDI considering the sociocultural background of North Korea. Inter-rater reliability, test/retest reliability, and validity of NK-CIDI were evaluated and kappa value was calculated.

Results Different vocabularies and awkward expressions were adjusted, and additional explanations were supplemented for difficult phases. The target for development of NK-CIDI was early defectors who entered South Korea less than 3 months ago. Inter-rater reliability (n=30, kappa value 0.35–1.00), test/retest reliability (n=30, kappa value 0–0.86), and diagnostic validity (n=100, kappa value -0.02–0.42) were evaluated.

Conclusion Compared to the general population, defectors experience more psychiatric issues, and a standardized tool for diagnosis is needed. In this study, NK-CIDI was developed and the results suggest that reliability is acceptable but validity needs further verification. NK-CIDI could be utilized in future epidemiologic studies. **J Korean Neuropsychiatr Assoc 2015;54(2):228-235**

KEY WORDS Psychiatric diagnosis · Reliability and validity · Epidemiology · Democratic People's Republic of Korea.

Received January 8, 2015
Revised January 23, 2015
Accepted March 10, 2015

Address for correspondence

Jin Pyo Hong, MD, PhD
Department of Psychiatry,
Samsung Medical Center,
81 Inwon-ro, Gangnam-gu,
Seoul 135-710, Korea
Tel +82-2-3410-3585
Fax +82-2-3410-0050
E-mail suhurhong@gmail.com

서 론

북한이탈주민의 수는 빠르게 증가하고 있다. 1990년대 후반까지 천 명에도 미치지 못했던 북한이탈주민의 수는 급격히 증가하기 시작하여 2014년 6월에는 2만6천 명을 넘어섰다.¹⁾ 북한사회의 불안정성 증가 등으로 인해 이 수는 향후 더욱 증가할 것으로 예상된다.

북한이탈주민은 탈북 과정에서 심각한 정신적 외상을 경험하는 경우가 많고, 탈북 후 제3국에서 체포를 피해 도망 다니며 폭력 및 가난에 시달린다. 어려움 끝에 남한에 정착하더라도 북한이탈주민이 가지는 독특한 정치문화적 특성으로 인해 이질감, 상호불신, 적대적 긴장감, 문화 충격 등을 경험하게 되고, 이로 인한 스트레스로 정신건강문제를 경험할 가능성이 높다.²⁾ 북한이탈주민의 정신건강과 관련한 실태조사

들이 소규모로 이루어진 바 있는데, 일반 남한 주민에 비해 우울증, 외상 후 스트레스장애, 알코올 사용장애가 더 많이 관찰되었다.³⁾

더욱 늘어날 것으로 예상되는 북한이탈주민의 정신건강문제에 대해 정확히 평가하는 한편 그 대처 방안을 마련하기 위해서는 북한이탈주민의 정신질환에 대한 역학 연구가 필요하다. 역학 연구를 통해 정신질환의 분포와 원인, 위험 및 보호 인자를 알아낼 수 있기 때문이다.⁴⁾ 이러한 역학 연구가 원활히 이루어지기 위해서는 표준화된 진단 도구가 우선적으로 필요하다.

과거에 이루어진 역학 연구들에서도 표준화된 진단 도구가 사용되었다. Diagnostic Interview Schedule(이하 DIS)⁵⁾은 미국의 국립정신보건원(National Institute of Mental Health)이 Epidemiologic Catchment Area Study⁶⁾에 사용하기 위하여 1981년에 개발한 진단 도구로서 한국어를 비롯한 24개 국어로 표준화되어 대규모 지역사회 역학 연구에 널리 활용되었다.^{7,8)} 특히 대규모의 지역사회 역학 연구에서는 비용-효과 측면을 고려해야 하기 때문에, 임상가가 아닌 적절한 교육을 받은 일반인도 수행할 수 있도록 진단 도구를 완전히 구조화할 필요가 있다. DIS의 경우 이러한 점에서는 적합하였으나 미국정신의학회(American Psychiatric Association)의 진단기준인 정신장애의 진단 및 통계 편람(Diagnostic and Statistical Manual of Mental Disorders, 이하 DSM)에 의한 진단만 가능하다는 한계점을 가지고 있었다.

이후 세계 각국의 역학 결과를 비교할 수 있는 국제적인 진단 도구 개발의 필요성이 강조되면서 세계보건기구(World Health Organization)와 미국의 Alcohol, Drug Abuse and Mental Health Administration은 국가 간의 횡문화적 역학 연구에 적합한 역학조사용 면담도구인 Composite International Diagnostic Interview(이하 CIDI)를 개발하였다.⁹⁾ CIDI의 두 가지 주된 특성은 첫째, 특정 증상의 존재 여부가 전적으로 질문에 대한 피면담자의 반응에 근거하여 판정된다는 점과, 둘째, 면담을 위한 질문들이 완전히 구조화되어 있다는 점이다. 면담을 전적으로 구조화함으로써 증상 평가의 일관성을 높일 수 있으며, 진단의 신뢰도를 높일 수 있다. CIDI는 역학 연구시 가장 많이 사용되고 있는 진단 도구 중 하나이며, 그 신뢰도와 타당도를 검토해 보았을 때 국제적인 연구 및 소규모의 지역 연구에서 수용할 정도의 일치도를 보였고 여러 국가나 상황에서 사용하기에 적합하였다.^{10,11)}

한국에서는 Cho 등¹²⁾이 한국어판 CIDI(Korean version of the CIDI, 이하 K-CIDI)를 개발하였고, Lee¹³⁾가 지역사회 일반인구를 대상으로 한국어판 Structured Clinical Interview for DSM Disorders(이하 SCID)를 이용하여 K-CIDI의 진단

적 타당도를 검증한 바 있다. 또한 2001년, 2006년, 2011년 전국정신질환 실태조사에서 K-CIDI를 사용한 바 있다.

그러나 K-CIDI를 북한이탈주민이나, 통일 이후 북한주민들에게 그대로 적용하기는 어렵다. 남한과 북한 모두 한국어를 제1언어로 사용하고 있기는 하나 분단의 세월만큼 의사소통의 장벽이 점차 커지고 있다. 북한이탈주민의 경우 남한사회에서 흔히 사용하는 언어, 특히 감정과 관련한 다양한 단어의 정의 및 용례에 익숙하지 않아서 의사소통에 어려움을 겪고 있다.¹⁴⁾ 이러한 언어의 문제는 정신건강문제의 진단 및 실태 조사에 영향을 미칠 수 있으며 결과를 왜곡할 가능성이 있다. 또한 언어로 인한 의사소통의 장벽 외에도 남한과 북한의 사회 구조 및 정치 이념의 차이로 인해 발생하는 사회문화적 격차도 정신건강문제의 평가에 영향을 미칠 수 있다. 실제로 표준화된 과정을 거쳐 CIDI를 번역하여서 의미의 차이는 없음에도 불구하고 문화적 차이가 있다면 진단의 일치도가 낮아지는 경우가 있다.^{15,16)} 따라서 북한에서 상용되는 언어 및 정서적 소통 방법, 북한의 사회문화적 배경을 반영한 정신질환 진단 도구의 개발이 필요하고, 이는 향후 북한이탈주민을 대상으로 하는 정신질환실태에 대한 체계적인 역학조사를 위한 필수적인 선행 작업이라 할 수 있다.

본 연구에서는 북한이탈주민 및 북한주민의 정신질환에 대한 역학조사를 위해 K-CIDI를 북한의 사회문화적 배경에 적합하게 번역 개발하고 그 타당도와 신뢰도를 검증하고자 했다.

방 법

대 상

본 연구는 국립중앙의료원 북한이탈주민 진료 센터를 방문한 환자, 국립중앙의료원에 입원해 있는 환자, 북한이탈주민 진료 센터의 북한이탈주민 상담실 내의 홍보를 통하여 모집된 북한이탈주민 및 북한이탈주민 지원시설에서 교육받는 국내 입국 6개월 이내의 북한이탈주민들을 대상으로 하였다. 18세 이상 65세 이하의 북한이탈주민들 중 연구에 대한 구두 설명을 듣고 참여 동의서에 자발적으로 서명한 경우를 대상으로 하였다. 면담 및 설문조사를 수행할 수 없는 지적장애, 운동장애가 있는 경우는 연구에서 배제하였다. 전체 연구 프로토콜은 임상시험심사위원회(Institutional Review Board)의 승인을 받았다. CIDI 및 정신과 의사의 임상적 면담을 시행한 100명의 자료를 분석에 이용하였다.

도 구

CIDI는 국가 간의 횡문화적 역학 연구와 정신병리의 비교연구를 목적으로 개발된 완전히 구조화된 검사 도구로 대규모

의 역학조사에 매우 유용하다.⁹⁾ CIDI는 423개의 기본 문항으로 이루어져 있으며, 진단 범주별로 15개 부분으로 구성되어 있어 필요한 부분만을 따로 적용할 수 있도록 되어 있다. 전체 CIDI의 적용시간은 75분 정도이나 증상이 없는 경우에는 그 다음 몇 문항을 건너뛰도록 되어 있어 시간이 단축된다. 18세 이상의 응답자용으로 고안되었으며 다양한 교육적 문화적 배경과 지적 능력을 가진 대상들이 이해할 수 있도록 개발되었다.¹⁷⁾

방 법

북한어판 CIDI(North Korean version of the Composite International Diagnostic Interview, NK-CIDI)의 제작 과정

북한어 정신건강 평가도구(North Korean version of the Composite International Diagnostic Interview, 이하 NK-CIDI)는 다음과 같은 과정을 거쳐 제작되었다. 북한어를 전공한 1인과 한국 표준어와 북한어에 모두 익숙한 국어학자 2인, 북한이탈주민 진료의 경험이 많은 하나원 소속 정신건강의학과 전문의 1인으로 번역위원회를 구성하였다. 번역위원회에서 K-CIDI를 북한어로 번역하였다. 한국에 입국한지 3~4개월이 경과한 북한이탈주민(대학교졸 1명, 인민학교졸 1명) · 입국한지 6~7개월이 경과한 북한이탈주민(남성 1명, 여성 1명) · 입국한지 1년이 경과한 북한이탈주민(남성 1명, 여성 1명) 등 총 6명이 1차 정독 후 검토해야 할 항목을 정리하여 일차 보완 번역을 하였다. 한국 입국 3~4개월이 경과한 또 다른 북한이탈주민이 일차 번역본을 검토하여 이차 보완 번역을 하였다. 탈북자 출신으로 현재 가정의학과에서 수련 중인 의사가 이차 번역본을 검토하면서 삼차 점검 및 감수를 하였다.

NK-CIDI의 번역 결과

NK-CIDI의 제작 과정에서 북한의 사회문화적 배경을 고려하여 이루어진 대표적인 조정 작업은 다음과 같다.

우선 남북한 간의 어휘 차이로 인해 발생할 수 있는 문제들을 고려하였다. 남북한의 생활 어휘가 차이가 나는 경우 단어를 북한의 실정에 맞게 변경하였다. 예를 들어 혼인신고는 결혼등록으로, 터널은 차굴, 엘리베이터는 승강기, 채장염은 취장염, 예는 레 등으로 변경하였다. 물질관련 장애의 L4 문항의 약물 종류의 경우 북한에서 실제 사용 중인 어휘인 역삼(대마), 얼음(암페타민), 현초알약(아편), 페로인(헤로인), 고강도접착제(본드) 등으로 번역하였다. 어휘가 북한에 존재하지 않거나 이해가 어려운 경우, 예를 들어 사별, 별거, 탈진, 에피소드 등의 경우 괄호 안에 추가 설명을 기술하였다. 남

북한의 사회문화적 차이 혹은 정치적인 이유로 이해 불가능한 어휘는 삭제하거나 추가 설명을 기술하였다. 예를 들어 K22.1a의 “전투를 겪었다면 어떤 전쟁입니까?”에서 기술한 걸프전, 이라크전의 경우 삭제하였고, 방송통신대의 경우 통신학교, 양성소 등으로 추가 설명을 기술하였다.

다음으로 표현을 오해할 수 있는 경우 이를 조정하였다. 일상적인 표현이 아닌 경우 다음과 같이 표현을 변경하였다. “갑자기 둔해지거나 저리다”는 “몸이 굳거나 몸의 일부분이 저린 느낌이 있다”로 변경하였으며 “남의 주목을 받을 수 있는 상황에 처했거나”는 “다른 사람의 시선이 내게 집중되는 상황에 처했거나”로 변경하였다. “속이 거북한 것 때문에”는 “속이 미식거리어서”로 변경하였다. 표현이 어렵거나 길어 이해도가 떨어질 경우 문장을 수정하였다. 예를 들어 “걱정을 떨쳐버리기가 어려웠습니까?”의 경우 “걱정을 안 하는 게 어려웠습니까?”로, “상황을 좋게 할 방법이 없다는 생각을 자주 했습니까?”는 “현재보다 내 처지가 좋아질 거 같지 않다는 생각을 자주 했습니까?”로 이해하기 쉽게 변경하였다. “남들에게 불안해 하는 것을 보이거나 창피하게 보일까 봐”를 “남들에게 불안해 하는 것을 보이기 싫어서 혹은 창피하게 생각되어서”로, “심한 운동을 하거나 폐경기에 접어들 때”는 “심한 운동을 하거나 폐경기(나이가 들어 월경이 끝나는 시기)에 접어들 때”로 변경하였다.

NK-CIDI의 타당도 및 신뢰도 검사

NK-CIDI를 적용할 조사요원 총 2인을 선발하여 세계보건기구가 제공한 교육자료를 이용하여 표준화된 교육을 실시하였다. 조사요원은 일반인(대학생) 2인으로 구성되었다. 조사요원 교육을 위한 훈련프로그램은 세계보건기구가 제시하는 지침을 준수하되, 추가되는 도구에 대한 교육일정을 추가하였으며 실제 현장 조사 상황을 가정하여 예상 질문에 대해 대답을 해보는 모의면담(mock interview)과 같은 실습을 강화하였다. 또한 실제 환자에 대해 2인 1조의 면담(live interview)을 하고 CIDI에 기입하는 과정을 진행하였다. 면담 후에는 CIDI 교육담당자와 조사요원이 모여서 회의를 하며 면담 및 기입 과정을 점검하였고 훈련생 간에도 서로 교정하도록 하여 교육에 적극적으로 참여하도록 하였다.

상기 조사요원 2인이 100명의 대상자에게 NK-CIDI를 시행하였다. 평가자 간 신뢰도(inter-rater reliability)를 평가하기 위하여 100명 중 30명을 대상으로 한 명의 조사요원은 직접 면담을 하면서 NK-CIDI를 채점하였고, 다른 한 명의 조사요원은 면담을 관찰하면서 NK-CIDI를 채점하였다. 검사/재검사 신뢰도(test/retest reliability)를 평가하기 위하여 일차 검사를 마친 100명 중 30명의 대상자에게 1~2주일 후

재검사를 실시하였다. 또한 타당도(validity)를 평가하기 위하여 NK-CIDI로 검사를 받은 100명 모두를 대상으로 정신건강의학과 의사가 임상적 면담을 시행하여 진단을 내리도록 하였다.

자료처리 및 분석

참여자 100명의 자료는 다음과 같은 방법을 통하여 처리 및 분석하였다. 우선, NK-CIDI 검사지를 검토하여 부적절한 코딩을 교정하고 조사 결과를 자료입력 프로그램을 이용하여 정리하였다. 이어서 DSM-IV 진단을 산출하기 위해 정리된 NK-CIDI의 문항들을 진단 알고리즘에 대입하였다. 그 결과 각 진단은 “진단 미확정(0)”, “진단 음성(1)”, “배제 기준을 제외한 진단 양성(3)”, “진단 양성(5)”의 네 가지로 분류되었다. “진단 미확정(0)”은 어떤 이유로 해당되는 부분을 건너뛰었기 때문에 해당 정신장애에 대한 진단 과정이 이루어지지 않은 경우이며, “진단 음성(1)”과 “진단 양성(5)”은 각각 진단 프로그램을 적용하여 해당 정신장애가 없는 것과 있는 것으로 판정된 경우이고, “배제 기준을 제외한 진단 양성(3)”은 배제 기준을 무시하면 정신장애가 있는 것으로 판정되는 경우이다.

NK-CIDI의 평가자 간 신뢰도, 검사/재검사 신뢰도는 상기 네 가지 분류를 기준으로 비교하였다. NK-CIDI의 진단 타당도를 분석할 때는 임상가의 진단과 비교가 가능하도록 “진단 미확정(0)”의 경우 “진단 음성(1)”에, “배제 기준을 제외한 진단 양성(3)”의 경우 “진단 양성(5)”에 포함시켜 두 가지 분류를 기준으로 비교하였다.

NK-CIDI의 평가자 간 신뢰도, 검사/재검사 신뢰도, 진단 타당도를 확인하기 위하여 코헨의 카파계수(Cohen's kappa coefficient)를 산출하였다. 자료 분석은 SPSS version 18.0 (SPSS Inc., Chicago, IL, USA)을 사용하였다. 모든 분석에서 $p < 0.05$ 인 경우를 통계적으로 유의한 것으로 간주하였다.

결 과

NK-CIDI의 타당도 및 신뢰도 검사

NK-CIDI 일차 면담을 완료한 100명(남자 21명, 여자 79명) 중(표 1), 나이는 남녀 모두 40대가 가장 많았으며(남자 33.3%, 여자 41.8%), 교육기간은 10~12년이 남성 9명(42.9%), 여성 41명(51.9%)으로 가장 많았다. 이혼 또는 별거 중인 경우(35.4%), 결혼한 경우(34.3%)가 각각 전체의 삼분의 일 가량을 차지하였다. 현재 직장이 없는 경우(84.5%)가 직장이 있는 경우(15.5%)보다 5배 이상 많았다.

평가자 간 신뢰도 검사는 30명을 대상으로 이루어졌고, 그

결과는 진단에 따라 kappa값이 0.35~1.00으로 평가되었다(표 2). 물질남용, 강박장애, 신체형장애, 식사장애 등에서 사례가 부족하여 kappa값이 산출되지 않았다. 특정공포증의 kappa값이 0.35로 가장 낮았고 그 외 진단에서는 0.73 이상으로 높게 평가되었다.

30명에서 이뤄진 검사-재검사 신뢰도 결과는 진단에 따라 kappa값이 0~0.86으로 평가되었다(표 3). 정신병적 장애, 외상 후 스트레스장애에서 사례가 부족하여 kappa값이 산출되지 않았다. Kappa값은 알코올 남용, 알코올 사용 장애, 광장공포증에서 0.86, 0.65, 0.64로 우수하게 평가되었고 불안장애, 공황장애, 특정공포증에서 0.26, 0, 0.19로 낮게 평가되었다.

타당도는 조사요원에 의한 NK-CIDI의 진단과 임상가 면담을 통하여 시행한 진단명 사이의 일치도를 분석한 것으로 100명 모두를 대상으로 하였다. 진단에 따라 kappa값이 -0.02~0.42로 평가되었다(표 4). 알코올 남용, 광장공포증, 사회공포증, 특정공포증 등에서 사례가 부족하여 kappa값이 산출되지 않았다.

고 찰

CIDI는 횡문화적 비교 연구를 위해 개발되었기 때문에 여

Table 1. Demographic characteristics of the NK-CIDI study participants (n=100)

| | Male (n=21) | Female (n=79) | Total (n=100) |
|--------------------|----------------|------------------|------------------|
| Age | | | |
| 18-29 | 2 (9.5) | 6 (7.6) | 8 (8.0) |
| 30-39 | 3 (14.3) | 7 (8.9) | 10 (10.0) |
| 40-49 | 7 (33.3) | 33 (41.8) | 40 (40.0) |
| 50-59 | 6 (28.6) | 23 (29.1) | 29 (29.0) |
| ≥60 | 3 (14.3) | 10 (12.7) | 13 (13.0) |
| Education | | | |
| 0 | 1 (4.8) | 0 (0.0) | 1 (1.0) |
| 1-6 | 1 (4.8) | 2 (2.5) | 3 (3.0) |
| 7-9 | 5 (23.8) | 8 (10.1) | 13 (13.0) |
| 10-12 | 9 (42.9) | 41 (51.9) | 50 (50.0) |
| 13-16 | 5 (23.8) | 18 (22.8) | 23 (23.0) |
| ≥17 | 1 (4.8) | 9 (11.4) | 10 (10.0) |
| Marital status | | | |
| Married | 6 (28.6) | 28 (35.9) | 34 (34.3) |
| Widowed | 1 (4.8) | 11 (14.1) | 12 (12.1) |
| Divorced/separated | 9 (42.9) | 26 (33.3) | 35 (35.4) |
| Single | 5 (23.8) | 13 (16.7) | 18 (18.2) |
| Employment status | | | |
| Yes | 4 (20.0) | 11 (14.3) | 15 (15.5) |
| No | 16 (80.0) | 66 (85.7) | 82 (84.5) |

NK-CIDI : North Korean version of the Composite International Diagnostic Interview

Table 2. Inter-rater reliability of the NK-CIDI (n=30)

| Diagnoses | Rater 1 | | | | Rater 2 | | | | Kappa | p value |
|---------------------------|---------|----|----|----|---------|----|----|----|-------|---------|
| | 0* | 1† | 3‡ | 5§ | 0* | 1† | 3‡ | 5§ | | |
| Alcohol use disorders | 13 | 15 | 0 | 2 | 13 | 15 | 0 | 2 | 1 | <0.01 |
| Alcohol dependence | 4 | 24 | 0 | 2 | 6 | 22 | 0 | 2 | 0.83 | <0.01 |
| Alcohol abuse | 15 | 15 | 0 | 0 | 16 | 14 | 0 | 0 | 0.93 | <0.01 |
| Mood disorders | 5 | 12 | 0 | 13 | 5 | 12 | 0 | 13 | 1 | <0.01 |
| Major depressive disorder | 3 | 16 | 0 | 11 | 3 | 16 | 0 | 11 | 1 | <0.01 |
| Bipolar disorder | 2 | 26 | 0 | 2 | 2 | 26 | 0 | 2 | 1 | <0.01 |
| Dysthymic disorder | 3 | 18 | 9 | 0 | 0 | 21 | 9 | 0 | 0.73 | <0.01 |
| Psychotic disorders | 0 | 27 | 1 | 2 | 0 | 27 | 1 | 2 | 1 | <0.01 |
| PTSD | 0 | 25 | 0 | 5 | 0 | 25 | 0 | 5 | 1 | <0.01 |
| Anxiety disorders | 15 | 0 | 0 | 15 | 17 | 0 | 0 | 13 | 0.73 | <0.01 |
| Panic disorder | 3 | 25 | 2 | 0 | 4 | 24 | 2 | 0 | 0.89 | <0.01 |
| Agoraphobia | 26 | 1 | 1 | 2 | 26 | 1 | 1 | 2 | 1 | <0.01 |
| Social phobia | 22 | 4 | 1 | 3 | 22 | 4 | 1 | 3 | 1 | <0.01 |
| Specific phobia | 4 | 18 | 1 | 7 | 16 | 9 | 1 | 4 | 0.35 | <0.01 |

* : Diagnosis indeterminate, † : Diagnosis negative, ‡ : Diagnosis positive except exclusion criteria, § : Diagnosis positive. NK-CIDI : North Korean version of the Composite International Diagnostic Interview, PTSD : Posttraumatic stress disorder

Table 3. Test-retest reliability of the NK-CIDI (n=30)

| Diagnoses | Test | | | | Retest | | | | Kappa | p value |
|---------------------------|------|----|----|----|--------|----|----|----|-------|---------|
| | 0* | 1† | 3‡ | 5§ | 0* | 1† | 3‡ | 5§ | | |
| Alcohol use disorders | 17 | 11 | 0 | 2 | 15 | 11 | 0 | 4 | 0.65 | <0.01 |
| Alcohol dependence | 5 | 23 | 0 | 2 | 5 | 21 | 0 | 4 | 0.45 | <0.01 |
| Alcohol abuse | 19 | 11 | 0 | 0 | 19 | 11 | 0 | 0 | 0.86 | <0.01 |
| Mood disorders | 4 | 9 | 0 | 15 | 0 | 14 | 0 | 14 | 0.5 | <0.01 |
| Major depressive disorder | 1 | 17 | 0 | 12 | 1 | 18 | 0 | 11 | 0.41 | <0.01 |
| Bipolar disorder | 0 | 26 | 0 | 4 | 1 | 27 | 0 | 2 | 0.37 | <0.01 |
| Dysthymic disorder | 3 | 17 | 9 | 1 | 1 | 23 | 5 | 1 | 0.41 | <0.01 |
| Psychotic disorders | 1 | 26 | 1 | 2 | 2 | 28 | 0 | 0 | - | - |
| PTSD | 0 | 28 | 0 | 2 | 0 | 30 | 0 | 0 | - | - |
| Anxiety disorders | 16 | 0 | 0 | 14 | 17 | 0 | 0 | 13 | 0.26 | 0.15 |
| Panic disorder | 8 | 21 | 1 | 0 | 3 | 26 | 1 | 0 | 0 | 0.98 |
| Agoraphobia | 27 | 0 | 2 | 1 | 27 | 0 | 2 | 1 | 0.64 | <0.01 |
| Social phobia | 25 | 1 | 0 | 4 | 22 | 2 | 3 | 3 | 0.46 | <0.01 |
| Specific phobia | 4 | 19 | 0 | 7 | 0 | 18 | 1 | 11 | 0.19 | 0.17 |

* : Diagnosis indeterminate, † : Diagnosis negative, ‡ : Diagnosis positive except exclusion criteria, § : Diagnosis positive. NK-CIDI : North Korean version of the Composite International Diagnostic Interview, PTSD : Posttraumatic stress disorder

러 문화에서의 사용을 고려하여 제작되었다. 그러나 각국에서 번역할 때 지역의 특수성을 반영하여 추가적인 조정이 이루어졌고, 이러한 조정에는 어휘의 차이, 문화적 차이, 문장 길이의 차이 등이 포함되었다.^{9,12)} 본 연구에서도 NK-CIDI는 K-CIDI의 각 질문이 의도하는 바에 대한 이해를 근거로 북한의 사회문화적 맥락에 맞게 제작되었다. 특히 분단 이후 생긴 어휘의 차이의 보정, 비일상적인 표현 및 긴 문장의 수정 등의 조정 작업이 이루어졌다.

NK-CIDI 번역시 북한 탈출 후 제3국 보호시설에서 보호를 받고 있는 주민들이나, 한국에 입국한지 2~3개월 가량이 경과한 입국 초기의 북한이탈주민들이 검사 내용을 이해할

수 있도록 하였다. 따라서 북한이탈주민의 학습 능력과 문장 이해력을 참고하여 번역할 필요가 있었다. 2014년 6월 기준, 한국에 입국한 북한이탈주민 26864명의 74%는 20~40대로,¹⁾ 이들이 교육을 받아야 했던 시기는 1994년 전후 고난의 행군이 시작될 즈음으로, 학교 출석률이 매우 저조했던 시기여서 학습의 기회가 많지 않았다. 또한 이탈주민은 북한에서의 직업이 무직(50%)이거나 단순노동직(38%)인 경우가 대부분이었다. 따라서 이들 대다수는 사회에 대한 이해도가 낮거나, 감정과 관련된 언어 학습의 기회가 적어 이에 대한 고려가 필요하였다.

K-CIDI를 제작할 때 한국어권 문화의 특수성을 고려해

Table 4. Diagnostic validity of the NK-CIDI (n=100)

| Diagnoses | Clinical diagnosis | | NK-CIDI | | Kappa | p value |
|---------------------------|--------------------|----|---------|----|-------|---------|
| | 1* | 5† | 1* | 5† | | |
| Alcohol use disorders | 96 | 4 | 95 | 5 | 0.42 | <0.01 |
| Alcohol dependence | 96 | 4 | 95 | 5 | 0.42 | <0.01 |
| Alcohol abuse | 98 | 1 | 99 | 0 | - | - |
| Mood disorders | 59 | 41 | 52 | 48 | 0.42 | <0.01 |
| Major depressive disorder | 61 | 39 | 63 | 37 | 0.41 | <0.01 |
| Bipolar disorder | 98 | 2 | 91 | 9 | 0.15 | <0.01 |
| Dysthymic disorder | 99 | 1 | 69 | 31 | -0.02 | 0.5 |
| Psychotic disorders | 98 | 2 | 91 | 9 | 0.34 | <0.01 |
| PTSD | 95 | 5 | 94 | 6 | 0.14 | 0.18 |
| Anxiety disorders | 96 | 4 | 53 | 47 | 0.05 | 0.25 |
| Panic disorder | 99 | 1 | 96 | 4 | -0.02 | 0.84 |
| Agoraphobia | 100 | 0 | 86 | 14 | - | - |
| Social phobia | 100 | 0 | 85 | 15 | - | - |
| Specific phobia | 100 | 0 | 71 | 29 | - | - |

* : Diagnosis negative, † : Diagnosis positive. NK-CIDI : North Korean version of the Composite International Diagnostic Interview, PTSD : Posttraumatic stress disorder

번역을 하였는데,¹²⁾ 이러한 K-CIDI를 바탕으로 하여 NK-CIDI를 제작하였다. 영어권 문화와 비교하였을 때는 남북한 문화가 가진 공통점들이 더 많기 때문에 이러한 제작 과정은 적절하다고 판단되었다. 연령 체계는 남한 및 북한의 전통적인 연령체계에 따른 나이가 아닌, 만 나이를 기록하도록 하여 영어권의 연령체계와 일치하도록 하였다. 또한 C45 문항에서 성생활에 대해 질문시 '성생활 없이도 잘 지냈습니까?'라는 영어권 표현 대신, '섹스에 관심이 없거나 부부관계를 싫어했습니까?'로 의역한 K-CIDI의 관점을 유지하였다. 영어권에서는 성에 대한 적극적인 표현이 정상으로 받아들여지는 반면 남한과 북한의 경우 문제가 없다는 정도의 표현이 정상으로 여겨질 수 있는 점을 반영하였다.

물질 관련 장애에서 L4 문항의 약물 종류의 경우 북한에서 실제 사용 중인 어휘인 역삼(대마), 얼음(암페타민), 현초(아편), 페로인(헤로인), 고강도접착제(본드) 등으로 번역하였다. 북한에서는 마약류의 경우 불법 약물로 분류되어 있기는 하나 그 단속이 철저하지 않은 경우들이 있어 남한과 비교하여 물질 사용 장애의 유병률이 높을 수 있다. 관한 어휘는 향후 지속적으로 개정하여 북한 주민의 물질 사용 장애에 대해 정확히 평가할 필요가 있다.

진단 도구 제작 후 그에 대한 신뢰도 및 타당도 평가는 필수적으로 이루어져야 한다. 진단 도구 평가시 타당도가 궁극적인 기준이 되지만 신뢰도가 떨어진다면 실제로 유용하지 않기 때문에 진단 기준이나 진단 도구 개발에는 신뢰도 검사가 중요하다. 정신과 진단의 신뢰도는 진단 과정을 구조화하고 평가자를 교육함으로써 향상시킬 수 있으며, CIDI의 경우 구조화된 진단 도구이며 평가자에 대한 교육 역시 표준화되

어 있다. 그러나 인구 집단마다 언어, 교육 수준, 증상의 표현 등이 다르기 때문에 새로운 인구 집단에 CIDI를 적용할 때는 신뢰도에 대한 평가가 다시 이루어져야 한다.¹⁸⁾ 신뢰도를 평가하는 방법으로는 평가자 간 신뢰도 검사와 검사-재검사 신뢰도 검사가 있다. 각 방법은 장단점이 있는 것으로 알려져 있다.¹⁹⁾

본 논문에서는 일치도를 평가하기 위해 Cohen's kappa를 사용하였다. Kappa값은 우연에 의한 일치를 보정한 일치도를 평가하는 통계학적 평가 방법으로 정의된다.²⁰⁾ Kappa값은 Fleis²¹⁾의 척도에 따르면 0.4 미만일 경우 낮은 일치도(poor), 0.4~0.75일 경우 타당한(satisfactory) 또는 우수한(good) 일치도, 0.75를 초과할 경우 탁월한(excellent) 일치도를 보이는 것으로 평가할 수 있다.

NK-CIDI의 평가자 간 신뢰도의 kappa값은 특정공포증(0.35) 외에는 0.73~1.00으로 매우 높게 평가되었다. 따라서 NK-CIDI의 평가자 간 신뢰도는 우수하다고 평가할 수 있다. 다만, 본 연구에서 평가자 간 신뢰도를 평가할 때 두 명의 조사원 중 한 명은 직접 면담을 하면서 NK-CIDI를 채점하였고, 다른 한 명은 면담을 관찰하면서 NK-CIDI를 채점하였기 때문에 관찰자가 면담자의 판단을 일부 알 수 있어 신뢰도가 실제보다 높게 평가될 가능성을 염두에 두어야 한다.¹²⁾

NK-CIDI의 검사-재검사 신뢰도의 kappa값은 0~0.86으로 평가자 간 신뢰도보다 다소 낮게 평가되었다. 검사-재검사 신뢰도는 재검사의 면담이 여러 가지 변수에 의해 영향을 받는다는 단점이 있다.¹⁸⁾ 재검사시 대상자의 상태에 변화가 있을 수 있고 이 경우 신뢰도는 실제보다 낮게 평가될 수 있다. 반면 재면담시 첫 번째 면담의 답변을 기억할 경우 같은

답변을 하여 신뢰도가 높게 평가될 가능성도 있다. 따라서 재검사 기간을 선정할 때 상태 변화가 있지 않을 정도로 짧고 기억에 의한 영향이 크지 않을 만큼 충분히 긴 시간이 필요하다. K-CIDI 개발시에는 재검사를 1주 간격을 두고 시행하였으며 본 연구에서는 1~2주 간격을 두고 검사를 시행하였다. 본 연구에서 검사-재검사 신뢰도의 kappa값은 평가자 간 신뢰도보다 낮았는데, 검사 시간 단축을 위해 대상자들이 증상이 없는 것으로 보고하였을 가능성도 고려할 필요가 있어 보인다. 과거 연구에서도 첫 번째 검사에서 증상이 있는 것으로 대답했을 때 질문이 길어진다는 것을 알게 되면 두 번째 검사에서는 검사 시간 단축을 위해 증상이 없는 것으로 보고하는 경향이 생길 수 있다는 점을 제시한 바 있다.²²⁾ 본 연구에서도 재검사시 14개의 진단 범주 중 7개의 진단에서 재검사시 양성률이 줄어드는 것을 확인할 수 있었다.

요약하자면 NK-CIDI의 평가자 간 신뢰도는 특정공포증 외의 13개의 진단에서 매우 높게 평가되었고, 검사-재검사 신뢰도는 8개의 진단에서 타당하거나 우수한 수준, 4개의 진단에서 낮은 수준으로 평가되었다. 검사-재검사 신뢰도가 일부 항목에서 낮게 평가되었으나, 평가자 간 신뢰도가 검증되었기 때문에 NK-CIDI의 진단도구로서의 신뢰도는 충분하다고 판단되었다.

NK-CIDI의 진단 타당도의 kappa값은 -0.02~0.42로 낮은 수준으로 평가되었다. 특히 불안장애를 비롯한 절반 이상의 진단의 타당도는 충분히 확보되지 못하였다. 타당도 검사에서는 임상 경험이 있는 정신과 의사의 비구조화된 면담을 통한 진단을 그 기준으로 사용하였는데 환자의 주관적 증상에 의존하는 경향으로 인하여 각 정신과적 진단의 기준을 빠짐없이 확인하지 못했던 것으로 보인다. 또한 정신과 의사들이 북한어나 북한 문화에 익숙하지 않은 관계로 인하여 효과적인 면담이 이뤄지지 못했을 가능성도 있다. 기존의 DIS의 타당도에 대한 연구에서도 정신과 의사의 진단을 기준으로 하였을 때 낮은 타당도를 보였던 바 있다.^{23,24)}

연구 과정에서 다음과 같은 제한점들이 발견되었다. 첫째, 구조화된 면담이 아닌 정신건강의학과 의사의 비구조화된 면담을 진단 타당도 검증의 기준으로 사용하였다는 점이다. 향후 추가 연구를 통하여 SCID 등 구조화된 면담도구의 북한어판을 개발하고 사용하여 타당도 검증이 적절히 이루어질 것으로 보인다. 둘째, 탈북자 중에 함경도 출신이 많아서 북한어 표현이 함경도 방언 중심으로 이루어졌고, 따라서 다른 북한 지역 출신은 이에 익숙하지 않을 수 있겠다. 셋째, 대상자의 수가 충분하지 못하여 일부 진단의 경우 신뢰도나 타당도를 검증하기에 사례가 부족했다. 본 연구에서 표본수가 충분했던 진단명에서는 신뢰도가 높게 나타난 점을 고려할

때, 추후 보다 큰 규모의 연구를 통해 추가적으로 신뢰도를 확인할 필요가 있다.

결론

북한이탈주민의 수는 이미 빠르게 증가하고 있으며, 향후 더욱 늘어날 가능성이 높다. 북한이탈주민은 남한 주민에 비해 정신건강문제를 경험할 가능성이 높으며, 이에 대한 정확한 평가를 통해 대처 방안을 마련하는 한편 장기적으로는 통일 후의 북한 주민의 정신건강문제에 대한 예측 및 평가가 필요하다. 본 연구에서는 NK-CIDI를 제작하여 그 신뢰도 및 타당도를 검증하였으며 신뢰도는 충분하다고 평가되었다. 그러나 불안장애를 비롯한 절반 이상의 진단의 타당도는 충분히 확보되지 못하여 향후 제한점을 보완한 연구가 필요하다. NK-CIDI는 북한 주민의 정신건강을 진단 평가하는 최초의 평가도구로서 향후 북한 주민의 정신건강문제 연구의 기초가 될 것으로 기대된다.

중심 단어 : 정신과적 진단 · 신뢰도와 타당도 · 역학 · 북한.

Acknowledgments

본 연구는 국립서울병원 2014년도 학술용역사업의 지원으로 이루어졌음. 또한 보건복지부 보건의료연구개발사업의 지원으로 이루어졌음(HI12C0035).

Conflicts of Interest

The authors have no financial conflicts of interest.

REFERENCES

- 1) Korea Hana Foundation [homepage on internet]. Seoul: Korea [updated 2014 Dec 1; cited 2014 Dec 15]. Available from: <http://www.dongposarang.com/>.
- 2) Jeon WT, Min SK, Lee MH, Lee ES. Adjustment of North Korean defectors in South Korea. J Korean Neuropsychiatr Assoc 1997;36:145-161.
- 3) Kim BC. Mental health of North Korean defectors. Ment Health Policy Forum 2010;4:19-37.
- 4) Freedman DX. Psychiatric epidemiology counts. Arch Gen Psychiatry 1984;41:931-933.
- 5) Robins LN, Helzer JE, Croughan J, Ratcliff KS. National Institute of Mental Health Diagnostic Interview Schedule. Its history, characteristics, and validity. Arch Gen Psychiatry 1981;38:381-389.
- 6) Regier DA, Myers JK, Kramer M, Robins LN, Blazer DG, Hough RL, et al. The NIMH Epidemiologic Catchment Area program. Historical context, major objectives, and study population characteristics. Arch Gen Psychiatry 1984;41:934-941.
- 7) Robins LN. Psychiatric epidemiology--a historic review. Soc Psychiatry Psychiatr Epidemiol 1990;25:16-26.
- 8) Aderibigbe YA. Psychiatric epidemiology in cross-cultural perspective: a review. Eur Arch Psychiatry Clin Neurosci 1995;246:37-46.
- 9) Robins LN, Wing J, Wittchen HU, Helzer JE, Babor TF, Burke J, et al. The Composite International Diagnostic Interview. An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. Arch Gen Psychiatry 1988;45:

- 1069-1077.
- 10) Wittchen HU, Robins LN, Cottler LB, Sartorius N, Burke JD, Regier D. Cross-cultural feasibility, reliability and sources of variance of the Composite International Diagnostic Interview (CIDI). The Multicentre WHO/ADAMHA Field Trials. *Br J Psychiatry* 1991;159:645-653, 658.
- 11) Wittchen HU. Reliability and validity studies of the WHO--Composite International Diagnostic Interview (CIDI): a critical review. *J Psychiatr Res* 1994;28:57-84.
- 12) Cho MJ, Hahm BJ, Suh DW, Hong JP, Bae JN, Kim JK, et al. Development of a Korean Version of the Composite International Diagnostic Interview (K-CIDI). *J Korean Neuropsychiatr Assoc* 2002;41:123-137.
- 13) LEE DW. Validity of the Korean version of the composite International diagnostic interview [dissertation]. Seoul: Seoul National University;2003.
- 14) Ha JH. Understanding of Emotional Communication Pattern of North Korea Defectors. *Konkuk Univ Humanit Res Inst* 2012;5:303-329.
- 15) Prince M. Measurement validity in cross-cultural comparative research. *Epidemiol Psychiatr Soc* 2008;17:211-220.
- 16) Ghimire DJ, Chardoul S, Kessler RC, Axinn WG, Adhikari BP. Modifying and validating the Composite International Diagnostic Interview (CIDI) for use in Nepal. *Int J Methods Psychiatr Res* 2013; 22:71-81.
- 17) Composite International Diagnostic Interview (CIDI core) Version 2.1. Geneva, Switzerland: World Health Organization;1997.
- 18) Tsuang MT, Tohen M, Jones P. Textbook of psychiatric epidemiology. 3rd ed. Hoboken, NJ: Wiley-Blackwell;2011.
- 19) Helzer JE, Robins LN, Taibleson M, Woodruff RA Jr, Reich T, Wish ED. Reliability of psychiatric diagnosis. I. A methodological review. *Arch Gen Psychiatry* 1977;34:129-133.
- 20) Spitzer RL, Cohen J, Fleiss JL, Endicott J. Quantification of agreement in psychiatric diagnosis. A new approach. *Arch Gen Psychiatry* 1967;17:83-87.
- 21) Fleis JL. Statistical Methods for Rates and Proportions. 2nd ed. New York: John Wiley and Sons, Inc;1981.
- 22) Robins LN. Epidemiology: reflections on testing the validity of psychiatric interviews. *Arch Gen Psychiatry* 1985;42:918-924.
- 23) Anthony JC, Folstein M, Romanoski AJ, Von Korff MR, Nestadt GR, Chahal R, et al. Comparison of the lay Diagnostic Interview Schedule and a standardized psychiatric diagnosis. Experience in eastern Baltimore. *Arch Gen Psychiatry* 1985;42:667-675.
- 24) Helzer JE, Robins LN, McEvoy LT, Spitznagel EL, Stoltzman RK, Farmer A, et al. A comparison of clinical and diagnostic interview schedule diagnoses. Physician reexamination of lay-interviewed cases in the general population. *Arch Gen Psychiatry* 1985;42:657-666.