مجلة العلوم الشاملة

Journal of Total Science

Higher Institute of Science & Technology

Raqdalen, Libya

البحوث المنشورة باللغات الأجنبية

Research Papers in Foreign Languages

Volume (5), Issue (19), (MAR. 2022)

Contents

Research Papers Published in Foreign Languages

Journal of Total Science
Volume (5), Issue (19((Mar. 2022)
ISSN: 2518-5799

مجلة العلوم الشاملة
المجلد (5)، العدد (19)، (مارس2022)
ردمد: 2518-5799

The consequence of CD2 associated protein mutation on podocyte phenotype

Keria H. Sheha[1], Ramadan M. Sarrab[2]

[1]El Jemyl high tech institute of medical sciences, [2]Renal medicine unit, Tripoli children Hospital

Abstrac

The functional unit in the kidneys is called the nephron and the filtration unit in this nephron is called the glomerulus. The tiny capillaries in the glomerulus **are each wrapped with a filter membrane where the blood filtration process occurs.** The glomerular filtration barrier consists of the glomerular basement membrane, glomerular endothelial cells and glomerular epithelial cells, or podocytes. The glomerular ultrafiltration barrier, containing highly terminally differentiated podocytes. A number of molecules which associate with nephrosis and podocyte damage have been described. One of these molecules is CD2-associated protein (CD2AP) which is a crucial protein for the slit-diaphragm assembly and function. In spite of the fact that CD2AP knockout causes nephrotic syndrome in mice and the heterozygous +/− mouse is prone to proteinuria, little is known about the relevance of this molecule in human renal pathology. we studied the effect of CD2AP mutation on the human podocyte phenotype. We described the dramatic effects of CD2AP mutation on the cells morphology and the expression of the mesenchymal, epithelial and other markers. We found that in contrast with wild type, CD2AP mutant podocytes acquired dedifferentiation characteristics.

Introduction

In a normal human adult (70kg), the kidneys filter about 180 liters of blood per day to give 1.5 – 2 liters of urine which pass through ureters to the bladder. The functional unit in the kidneys is called the nephron and the filtration unit in this nephron is called the glomerulus. The tiny capillaries in the glomerulus **are each wrapped with a filter membrane where the blood filtration process occurs.** The glomerular filtration barrier consists of the glomerular basement membrane, glomerular endothelial cells and glomerular epithelial cells, or podocytes. Podocytes have branching and interdigitating processes, and filtration takes place through slits between these processes. The slit diaphragm, a critical component of the filtration barrier, is an ultra-thin zipper-like structure

thatbridges the gap between interdigitating podocytefoot processes. These three layers of glomerular filtration barrier separate the fluid in glomerular capillaries from that in Bowman's capsule.

A large body of evidence suggests that the podocyte foot processes and slit diaphragm are pivotal components of the glomerular filter, and disruption of their integrity is a critical event in the development of proteinuria and nephrotic syndrome in a variety of inherited and acquired glomerular disorders[1, 17, 18]. Many genetic studies have underscored that podocyte slit diaphragm associated proteins, such as nephrin and podocin, play an essential role in establishing the size-selective filtration barrier of the kidney, and mutations or deletions of the genes encoding these proteins are consequently associated with the development of proteinuria in both animal models and patients[3, 9, 11, 19]. However, mutations in the slit diaphragm-associated proteins are rare in most common forms of chronic kidney diseases such as diabetic nephropathy.

CD2-associated protein (CD2AP) is initially identified as an adapter protein in T lymphocyte, which is critical for stabilizing contacts between T cells and antigen-presenting cells [6]. In the kidney, CD2AP is expressed primarily in podocytes. CD2AP knock-out mice exhibited defects in foot processes, accompanied by mesangial cell hyperplasia and extracellular matrix deposition. Supporting a role for CD2AP in the slit diaphragm, CD2AP associates with nephrin, the primary component of the slit diaphragm [7]. Recent studies revealed that CD2AP is involved in the regulation of actin cytoskeleton, and lowered CD2AP expression is associated with the disruption of cytoskeleton and apoptosis in podocytes [13]. Another study indicated that podocytes lacking CD2AP are more susceptible to transforming growth factor-β1 (TGF-β1)-induced apoptosis and reconstitution of CD2AP may reverse this process [10].

Methdology

Conditionally immortalized human podocyte cell culture

Several conditionally immortalized human podocyte cell lines were used. Nephrectomy specimens were taken from a three year old child with hydronephrosis [16] and a four year old child with Wilms' tumor (Saleem et al, unpublished). CD2AP mutant podocytes were derived from a child with severe progressive mesangial sclerosis which was

found to be due to a CD2 associated protein mutation,heterozygous A599V exon 16 missense mutation (**Saleem et al, not published**

General culture conditions

All cell culture was performed under sterile condition in a Microflow Biological Safety Cabinet (MDH, Andover, Hampshire). Cells were incubated in 5% CO_2 in air, at either 33˚C or 37˚C, in a Hereaus Function Line CO_2 incubator. Cells were cultured in 12, 6-well plates or 75cm^2, 175 cm^2 flasks (all Iwaki, Tokyo, Japan).

Immunocytochemistry

The immunostaining of the cells can be divided into three main steps: - cells preparation, fixation and antibody binding and detection. Firstly, cells were prepared by growing them on sterile cover slips or plastic tissue culture dish. Secondly, media was removed from cells and washed twice with 1xPBS. The cells were fixed 100µl of 2% solution of paraformaldehyde for 10 minutes at room temperature. The formaldehyde was removed and cells were then washed once with 1xPBS for five minutes. Permeabilisation of the fixed cells was achieved by incubating the cells in 0.3% triton x-100 in PBS for 10 minutes and then rinsing once gently in washing solution (4% FCS and PBS). Cells were then blocked by using the 100µl block solution (which consists of 2% FCS and 2% BSA in PBS) for 30 minutes. Next the blocking solution removed and the coverslips transferred to parafilm on the base of a humidity chamber. Finally, antibody binding and detection were achieved by indirect antibody detection using a labeled secondary reagent that binds specifically to the primary antibody. Immunostaining was imaged using light microscopy connected to the Leica DFC350 FX monochrome digital camera.

Electrophoresis

Plates with set gels were loaded into a mini gel electrophoresis tank and the tank was filled with 1x running buffer, then connected to power pack (Bio-Rad 170-3846). Electrophoresis was performed at a constant voltage and an initial setting of 100V and once a linear running front had established, this was increased to 150V. The samples were ran until samples were almost to the bottom of the gel.

Immunodetection

This process detects proteins using specific antibodies. To block the non specific binding sites on the PVDF membrane, it is incubated in 5% bovine serum albumin (BSA) in 1x Tris-Buffered Saline Tween-20 **(**TBST) for one hour at room temperature using laboratory shaker. The primary antibody which identifies the protein of interest was added to the membrane and incubated at 4˚C overnight under gentle rotation. The following day, the unbound primary antibody was washed away with at least 5 times of at least five minutes each wash with 1xTBST at room temperature. Then the membrane was incubated at room temperature for one hour on a rotator in horseradish peroxidase (HRP) conjugated to secondary antibody diluted in 3% BSA in 1xTBST. Then the membrane was rinsed again as above. SuperSignal West Femto Maximum Sensitivity solution was added to the blots for 5 minutes. The membrane was run through the developing machine ChemiDoc-it imaging system (UVP, Upland California, USA).

Results

Morphological transformation due to CD2AP mutation in podocytes

Wild type podocytes had typical cobblestone morphology when grown on gelatin-coated tissue culture flasks at 33˚C (Fig.1 A). Whereas, CD2AP mutant podocytes when cultured at 33˚C the proliferating cells were noted to have a mesenchymal, elongated shape when subconfluent (not shown), however when entirely confluent assumed a more classic epithelial, cobblestoned appearance (Fig.1 B).Shifting the cells to 37˚C resulted in arrest of proliferation of cells within 2 days, and over a period of 7 to 14 days, wild type cell bodies enlarged in an irregular, very flat, and relatively rounded shape, and they tend to grow as groups of cells that preserve cell contact with their neighbours (Fig.1 E). In contrast, (Fig.1 D, F) the morphology of CD2AP mutant podocytes at nonpermissive (37˚C) culture condition was entirely different from the wild type. Mutant cells from day 3 at 37˚C become elongated and start losing their cell-cell contact and by day 6^{th} at nonpermissive showed fibroblast like morphology with spindle shape, cells seem to overgrow each other.

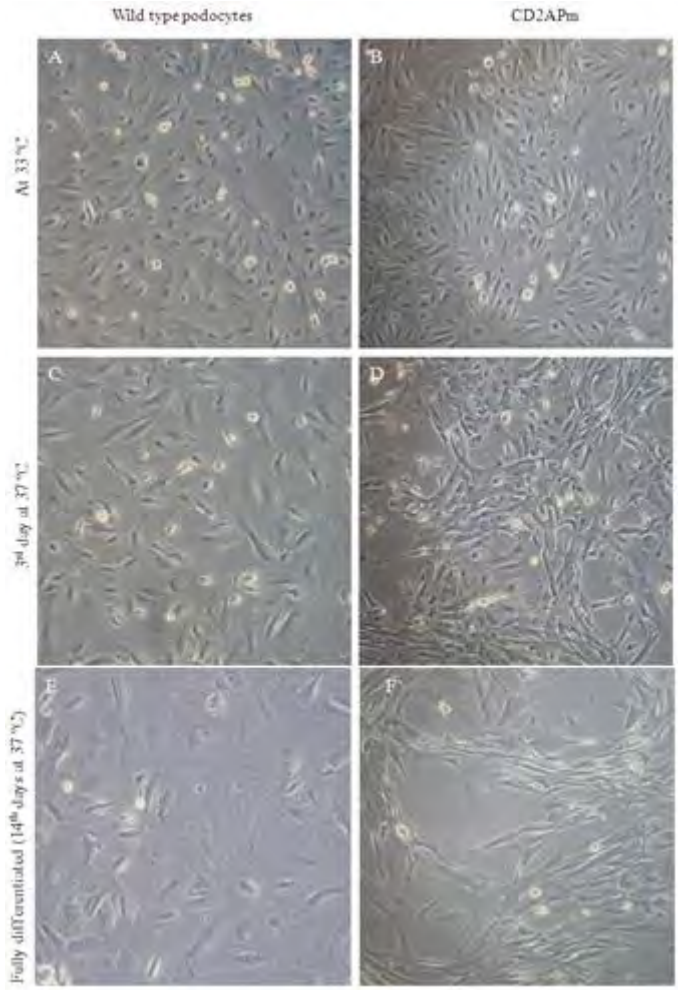


Figure 1: light microscopy images show morphology of conditionally immortalized human wild type podocytes and CD2AP mutant (CD2APm) podocytes growing at 33 ˚C, after 3 days at 37 ˚C and when become fully differentiated (at day 14th at 37 ˚C).

Confirming the expression of podocyte-specific proteins in cultured CD2AP mutant cells

The result shown in figure 2 A, demonstrates bands consistent with the predicted molecular weight of nephrin (180KDa) in protein extracted

from wild type and CD2AP mutant podocytes. The other sensitive and specific marker for differentiated podocytes is podocin (42KDa), and again this was found to be expressed in both in mature wild type and CD2AP mutant podocytes. CD2AP (75KDa) was similarly expressed in wild type and CD2AP mutant podocytes. WT1 is a transcription factor involved in the nephrogenesis and podocyte differentiation. Figure 2 B shows that WT1 expressed in CD2APm podocytes with level similar to wild type podocytes. Therefore, nephrin, podocin, CD2AP and WT1 are expressed in both wild type and CD2AP mutant podocytes at a comparable level.

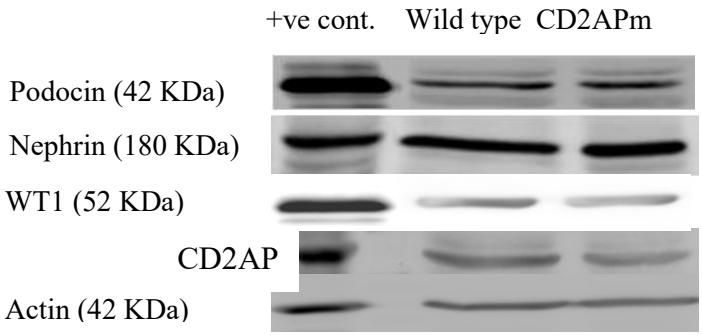


Figure: 2show the expression of podocyte markers (nephrin, podocin, CD2AP and WT1) in wild type and CD2AP mutant (CD2APm) podocytes.

Validating the expression of podocyte specific proteins in CD2AP mutant podocytes by immunoflourescence

In wild type podocytes, (Fig 3A) nephrin expression at the cell periphery, but also cytoplasmic distribution. In contrast in the mutant cells, nephrin was distributed primarily in the cytoplasm. Similarly, CD2AP and podocin was detected by immunostaining both in wild type and mutant podocytes (Fig 3 B&C), and was negative in mesangial cells. WT1, as can be seen in (Fig 3 D) is expressed in both podocyte cell lines and again was negative in mesangial cells. These results verify that nephrin, podocin, CD2AP and WT1 were expressed in wild and mutant podocytes but not in mesangial cells and thus serve as excellent markers to confirm that the mutant cells have podocyte origin.

Expression of epithelial markers in cultured CD2AP mutant podocytes

To evaluate the expression of the epithelial markers occludin, E-cadherin and cytokeratin in CD2AP mutant podocytes, western blot analysis was performed for CD2AP mutant podocytes and wild type podocytes using antibodies for the above mentioned proteins. Breast epithelial cells were used as a positive control. Figure 5

shows that occludin was expressed in epithelial cells and that its expression in both

wild type and mutant podocytes was very faint. As expected, E-cadherin expressed strongly in epithelial cells whereas, in both cell lines of podocytes its expression was

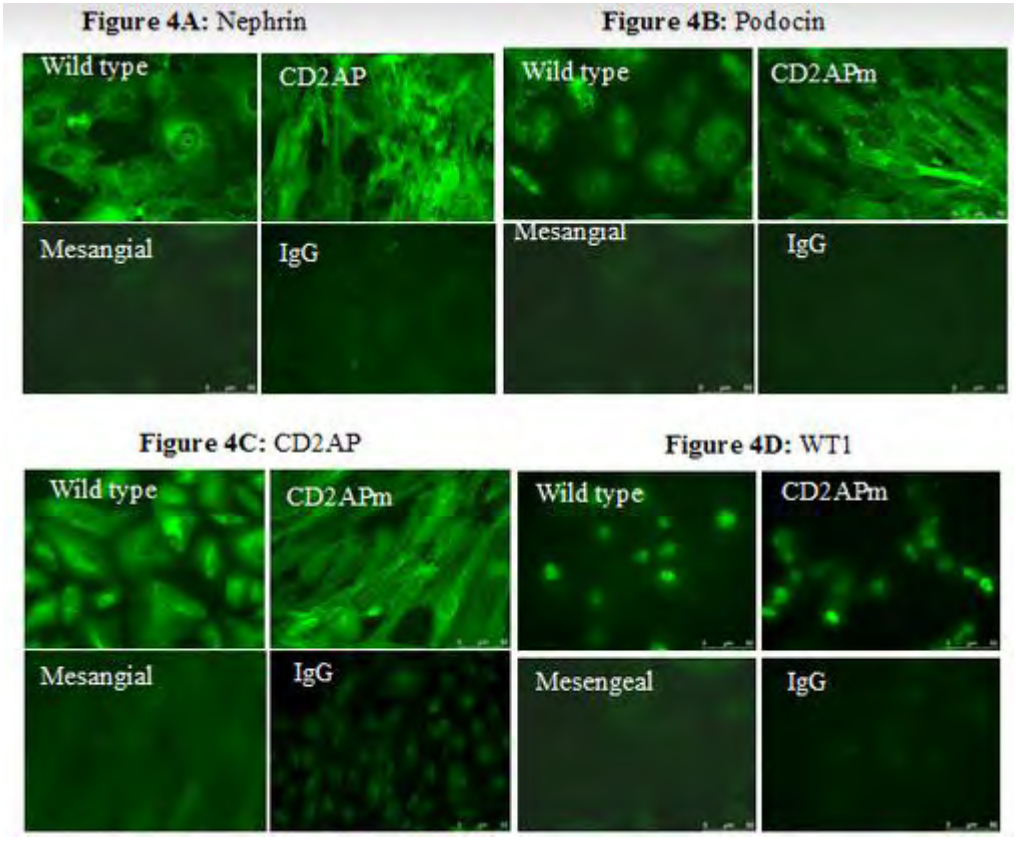


Figure 3: Specific indirect immunoflourescence shows the expression of the podocyte specific molecules nephrin (A),(CD2AP (B), podocin (C), and WT1 (D)) in wild type and CD2APm podocytes. Mesangial cells were used as a negative control.

completely absent. Interestingly cytokeratin which abundantly expressed in epithelial cells, in wild type podocytes its expression was very weak, and in CD2AP mutant podocytes, it was heavily expressed.

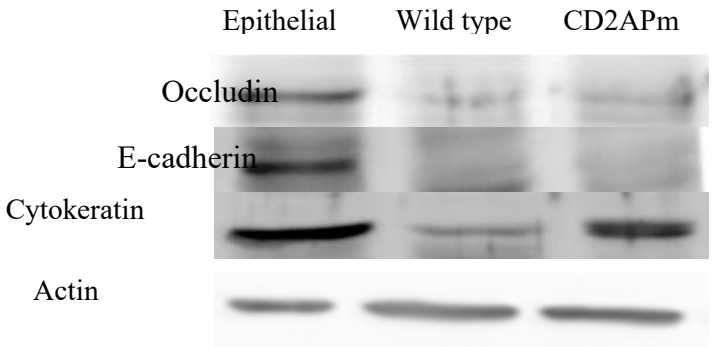


Figure 5: shows expression of epithelial markers occludin, E-cadherin and cytokeratin in CD2AP mutant podocytes and wild type podocytes, breast epithelial cells were used as a positive control and actin was used as a loading control.

Mesenchymal markers expression in cultured CD2AP mutant podocytes

To evaluate the effect of the CD2AP mutation on the mesenchymal markers expression in cultured podocytes. Expression of α SMA and fibronectin were examined using western blotting, (figure 5 A & B) showing expression of α-SMA and fibronectin in differentiated wild type and CD2AP mutant podocytes. It is clear that α-SMA and fibronectin were upregulated in the mutant podocytes in comparison with the wild type, and as expected they were overexpressed in mesangial. These results indicate that the CD2AP mutant podocytes overexpress the mesenchymal markers, α-SMA and fibronectin.

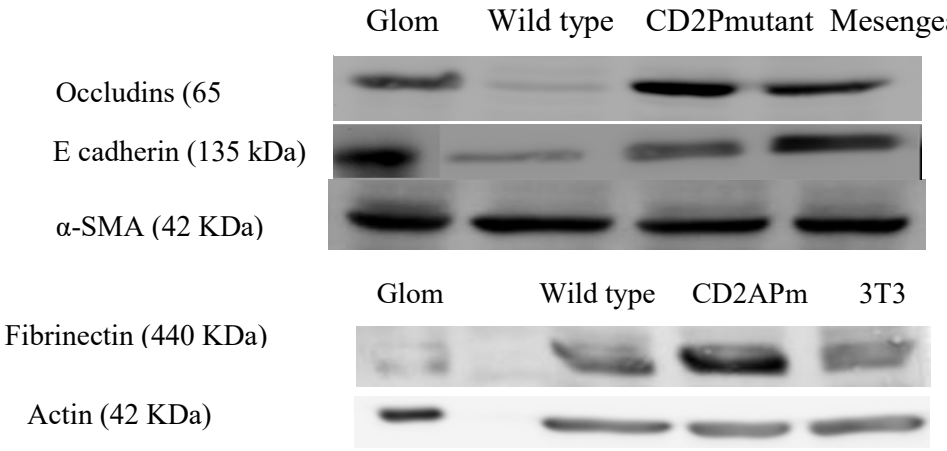


Figure5: Shows expression of mesenchymal markers (α-SMA, fibronectin and FSP1) in wild type podocytes, CD2AP mutant podocytes and mesangial cells.

Prove the expression of mesenchymal markerproteins SMA and fibronectin in cultured CD2AP mutant podocyte by immunostaining

Figure 6A, shows α-SMA stained strongly positive in linear pattern in the cytoplasm of CD2AP mutant podocytes and mesangial cell. In contrast, α-SMA was faint and barely detected and the stress fibres were completely absent in cultured wild type podocytes. Similarly, fibronectin was overexpressed in CD2AP mutant podocytes and expectedly in mesangial cells in comparison with type podocytes where its expression was less prominent. (Figure 6 A) Wild type podocytes were characterized by thick bundled cortical actin and bridging structures between adjacent cells. The staining for actin was also robust along the cell borders and in the cytoplasm of the cultured CD2AP mutant podocytes (Figure 6 B). Mouse IgG was used as a negative control.. The major morphological differences between cobblestone-wild type podocytes and the relatively large, fibroblast like cells CD2AP mutant podocytes was achieved by substantial rearrangements of cytoskeletal architecture (Fig. 6 B).

The consequence of CD2 associated protein mutation on podocyte phenotype

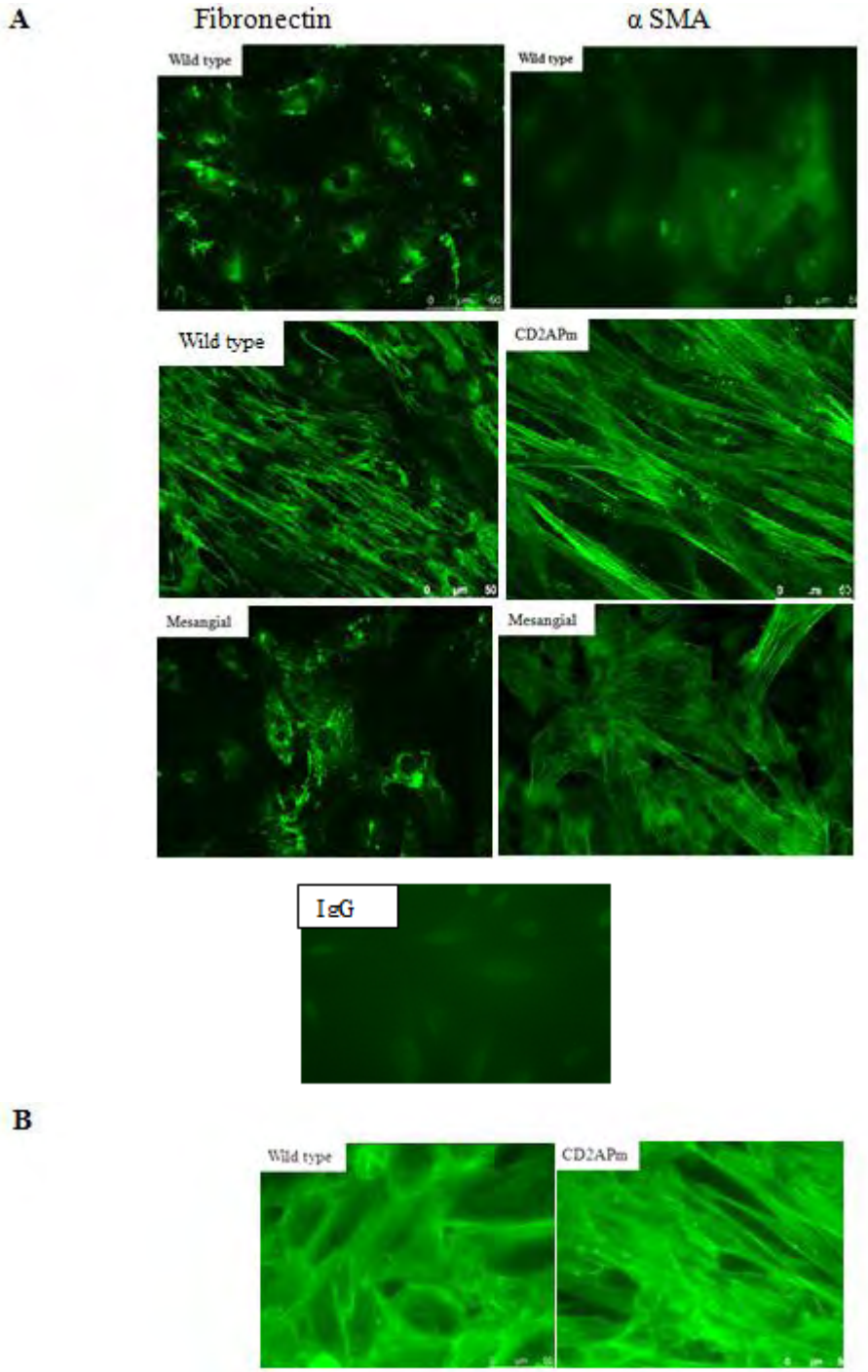


Figure 6: Specific indirect immunoflourescence of wild type podocytes, CD2AP mutant podocytes and mesangial cells. Cells were grown separately under same cell culture conditions and stained for α-SMA and fibronectin

Misexpression of PAX2 in CD2AP mutant podocytes

Figure 7A, staining of PAX2 in human podocytes with CD2AP mutation shows massive upregulation of PAX2 in nucleus of these cells. In contrast, PAX2 immunostaining was virtually negative in cultured wild type podocytes. In mesangial cells expression was negative as well. To validate the immunostaining results, western blot for PAX2 was performed in the same cell lines. Figure 7B shows the presence of intense band on western blotting consistent with the predicted molecular weight for the PAX2 protein (42 kDa) in the mutant podocyte. Whereas, in wild type podocytes and mesangial cells western blot analysis showed absence of PAX2 bands in these cells lysate (Figure 7B).

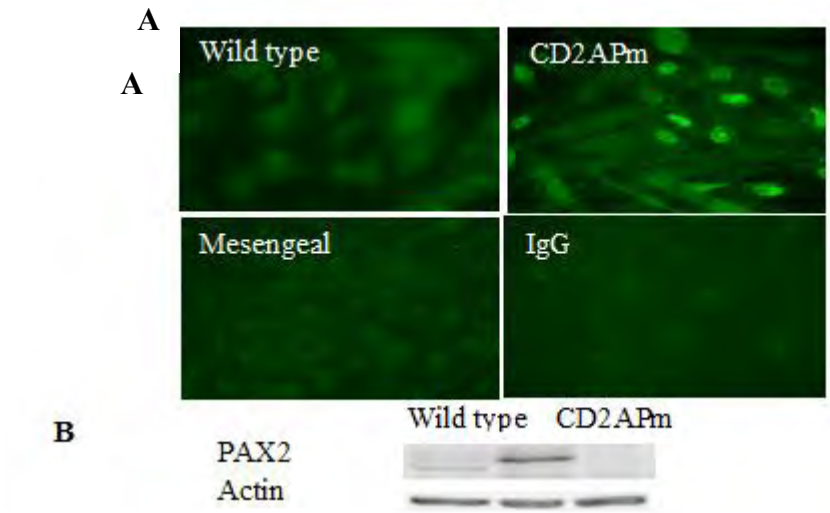


Figure 7: Shows PAX2 expression in wild type and CD2AP mutant podocytes. (A) Immunoflourescence , (B) Western blotting results

Junctional protein zonula occludens-1 (ZO1) expression in CD2AP mutant cells

Western blotting results (Figure 8A) shows the expression of ZO1 in wild type and CD2AP mutant podocytes and mesangial cells. ZO1 was similarly expressed in wild type and CD2AP mutant podocytes and mesangial cells. To examine the localization of ZO1 in CD2AP mutant podocytes, immunostaining of wild type and mutant podocytes was done. Interestingly, immunostaining for ZO1 showed (Figure 8B) that in wild type podocyte it was expressed at the cell surface, particularly at cell-cell contacts (figure 8B, left image). In contrast, CD2AP mutant podocytes loss this junctional pattern and rearranged in random way in the cytoplasm.

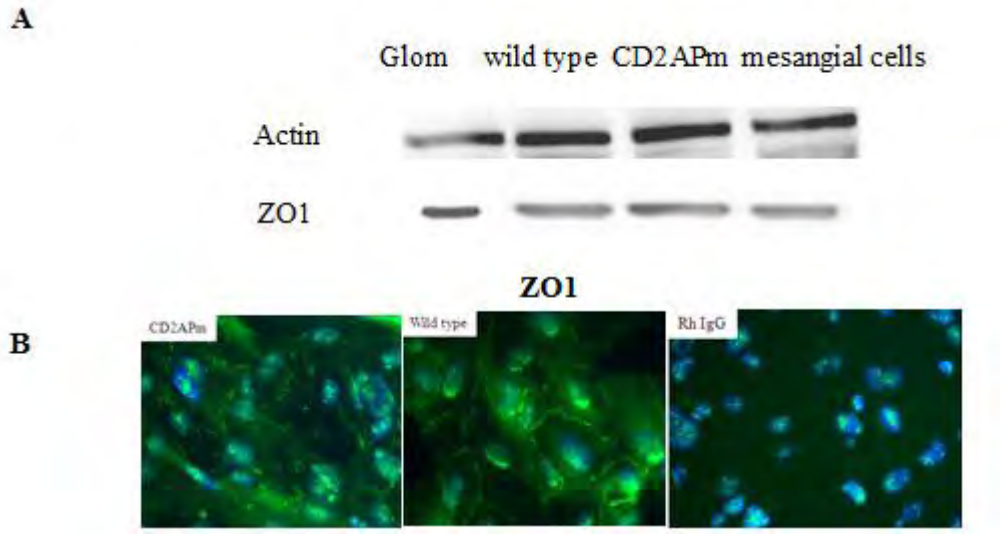


Figure 8: shows the expression of ZO1 in wild type and CD2AP mutant podocytes. (A) Western blotting (B) Immunoflouescence.

Discussion

During metanephric kidney development, substantial changes in podocyte phenotype are observed. Podocyte precursor cells arise from induced mesenchymal renal stem cells. During the "S-shaped body stage" of glomerular development, the presumptive podocytes acquire characteristics of epithelial cells, including apically located tight junctions[22]. As podocytes enter the subsequent "capillary loop stage,"

dramatic changes in phenotype occur. The cells begin to establish their complex cell architecture, including the formation of the characteristic foot processes, and the tight junctions are replaced by the slit membrane bridging, the filtration slits between the developing foot processes[22]. Thus, differentiation of podocytes represents an epithelial-to-mesenchymal transdifferentiation

We have described podocytes derived from the kidney of patient with diffuse mesangial sclerosis and CD2AP mutation, a heterozygous A599V exon 16 missense mutation, We examined the phenotypic characteristics of immortalized human CD2AP mutant podocytes derived from this patient including the cell morphology changes and the expression of different protein markers in comparison with wild type podocytes.

Our findings showed that at the permissive temperature of 33˚C, the wild type podocyte grew in cobblestone morphology (Figure 1). Shifting the cells to 37˚C resulted in arrest of proliferation of cells over a period of 7 to 14 days[16], the cell bodies enlarged in an irregular shape, with the formation of processes both short and more rounded and also long, spindle-like projections (Figure 1) similar to those previously described for human podocyte cultures [2, 16]. In contrast, CD2AP mutant podocytes at permissive temperature of 33˚C grew with slight elongated morphology when sub confluent,however when entirely confluent assumed a more classic epithelial, cobblestoned appearance (figure 1 B). Interestingly, the differentiated CD2AP mutant podocytes were consistently of a more elongated, essentially fibroblastic appearance; unlike the more rounded shape of wild-type cells. This finding suggested that the cytoskeletal architecture of the CD2AP mutant podocytes is altered. The possibility that CD2APinfluences cytoskeletal architecture is supported by recent studies in podocytes, in which deletion or mutation of α-actinin-4, synaptopodin, CD2AP, Nck1 and 2 in the human or rodent podocyte are associated with alterations in podocyte cytoskeletal architecture and/or junction formation[4, 6, 12, 14].To provide evidence that these are podocyte in origin, a comparison between wild-type and CD2AP mutant podocytes was initiated.Here, the expression of the podocyte specific proteins were compared, proteins critical to proper podocyte function: podocin[16], nephrin[16] and CD2AP[8]. These proteins have been proposed to interact with each other within podocyte foot processes to stabilize the slit diaphragm[19]. The other mature podocyte specific marker is WT1[2]. Western blotting analysis showed

that the podocyte specific proteins, CD2AP, podocin, nephrin and WT1 were expressed at comparable level in wild type and mutant podocytes and also the immunoflourescence examination of the same cells supports that these makers expressed in both cell lines with some variation in protein distribution between the wild type and CD2AP mutant podocytes. These findings provide evidence that the putative CD2AP mutant podocyte are of podocyte origin.

Podocyte dysfunction is a principal feature of many proteinuric kidney diseases[20, 23]. The results presented in this study demonstrate that podocytes are capable of undergoing EMT due to CD2AP mutation, and this phenotypic conversion causes the podocyte to lose its specialized podocyte features and to acquire new mesenchymal markers. It is conceivable that podocytes after EMT will abandon their complex morphological architecture and relinquish their highly specialized functions, which undoubtedly impairs the integrity of glomerular filtration barrier, leading to the onset of proteinuria. Therefore, our present findings provide a novel and rational illumination of some cellular events leading to podocyte dysfunction and proteinuria in response to CD2AP disturbance.

We found that the mutant cells loss their junctional pattern of ZO1, which rearranged instead in random way in the cytoplasm, and acquired mesenchymal markers, α-SMA and fibronectin. In agreement with our result it has been reported that CD2AP, and ZO1 are important components of the slit diaphragm cell adhesion complexes, and their default may lead to podocyte dedifferentiation and impair the integrity of the slit diaphragm [17, 24].

First we have shown that the podocytes specific proteins expressed in CD2AP mutant podocytes at similar level to that in wild type podocytes and the mutant cells overexpressed mesenchymal markers in comparison to the wild type. Moreover, the epithelial markers expression in the mutant cells was examined including E-cadherin, occludin and cytokeratin. we have observed the absence of E-cadherin protein expression in the mutant and wild type podocytes; this is in line with the findings of Schnabel and Reiser [1, 5]. Occludin which is one of characteristic protein of tight junctions, was very weakly expressed in the wild type and mutant podocytes; however Saitou found that occludin was not expressed in podocytes [25]. One important podocyte marker is WT1,

a tumour suppressor gene that is widely expressed in epithelial cells of the early nephron and is restricted to podocytes in the mature glomeruli. One of the WT1 targets is the transcription factor PAX2, the expression of which is normally down-regulated by WT1 in the mature glomeruli[21]. In glomerulopathies characterized by proliferation of podocytes, such as collapsing nephropathy and diffuse mesangial sclerosis as well as crescentic glomerulonephritis, abnormal expression of WT1 and increased expression of PAX2 results in dysregulation of podocytes which regress to an immature phenotype, re-enter the cell cycle, and undergo proliferation[15, 26]. Interestingly in our data, the western blotting and immunoflourescence analysis showed that, in contrast to wild type podocytes, CD2AP mutant podocyte inappropriately express PAX2, whichis normally expressed during early nephrogenesis[27]. Moreover, WT1 expressed at comparable level in wild type and mutant podocyte. These results may indicate that there is a link between the slit diaphragm protein, CD2AP, and the nuclear protein, WT1.

Summary of results

Because CD2AP mutation podocyte experienced many phenotypic changes which include

- Fibroblast-like morphology and disruption of cell-cell contacts.
- Normal expression of characteristic podocyte markers and this suggest that these cells are podocyte origin.
- Similar to wild type, CD2AP mutated podocytes do not express E-cadherin protein and very weakly express occludin but they acquired expression of cytokeratin that was not found in the wild type podocytes.
- In comparison with wild type, CD2AP mutant podocytes over expressed mesenchymal markers fibronectin, α-SMA and FSP1 suggesting that this cells undergo to EMT changes.
- CD2AP mutant podocyte inappropriately expressed PAX2 and disrupted of normal recruitment of ZO1atcell–cell junction.

Future work

The study of effects of CD2AP mutation on podocytes has raised acrucial question for research.The pathogenesis of many glomerular diseases such as FSGS is thought to result from podocyte dysregulation leading to dedifferentiation and true hyperplasia of these cells. it will be interesting to study the effect of drugs such as cyclosporine which used to treat these conditions on the phenotype of CD2AP mutant podocytes.

References

1. Schnabel, E., J.M. Anderson, and M.G. Farquhar, 1990.*The tight junction protein ZO-1 is concentrated along slit diaphragms of the glomerular epithelium.* J Cell Biol, **111**(3): p. 1255-63.
2. Mundlos, S., et al., 1993. *Nuclear localization of the protein encoded by the Wilms' tumor gene WT1 in embryonic and adult tissues.* Development. **119**(4): p. 1329-41.
3. Tryggvason, K., 1999. *Unraveling the mechanisms of glomerular ultrafiltration: nephrin, a key component of the slit diaphragm.* J Am Soc Nephrol, **10**(11): p. 2440-
4. Jones, N., et al., 2006.*Nck adaptor proteins link nephrin to the actin cytoskeleton of kidney podocytes.* Nature, **440**(7085): p. 818-23.
5. Reiser, J., et al., 2000.*The glomerular slit diaphragm is a modified adherens junction.* J Am Soc Nephrol, **11**(1): p. 1-8.
6. Shih, N.Y., et al., 1999.*Congenital nephrotic syndrome in mice lacking CD2-associated protein.* Science, **286**(5438): p. 312-5.
7. Dustin, M.L., et al., *A novel adaptor protein orchestrates receptor patterning and cytoskeletal polarity in T-cell contacts.* Cell, 1998. **94**(5): p. 667-77.
8. Shih, N.Y., et al., *CD2AP localizes to the slit diaphragm and binds to nephrin via a novel C-terminal domain.* Am J Pathol, 2001. **159**(6): p. 2303-8.
9. Schwarz, K., et al., *Podocin, a raft-associated component of the glomerular slit diaphragm,* 121. Schwarz, K., et al., *Podocin, a raft-associated component of the glomerular slit diaphragm,*

interacts with CD2AP and nephrin. J Clin Invest, 2001. **108**(11): p. 1621-9.

10. Schiffer, M., et al., *A novel role for the adaptor molecule CD2-associated protein in transforming growth factor-beta-induced apoptosis.* J Biol Chem, 2004. **279**(35): p. 37004-12.

11. Kim, J.M., et al., *CD2-associated protein haploinsufficiency is linked to glomerular disease susceptibility.* Science, 2003. **300**(5623): p. 1298-300.

12. Huber, T.B., et al., *Bigenic mouse models of focal segmental glomerulosclerosis involving pairwise interaction of CD2AP, Fyn, and synaptopodin.* J Clin Invest, 2006. **116**(5): p. 1337-45.

13. Lehtonen, S., F. Zhao, and E. Lehtonen, *CD2-associated protein directly interacts with the actin cytoskeleton.* Am J Physiol Renal Physiol, 2002. **283**(4): p. F734-43.

14. Kaplan, J.M., et al., *Mutations in ACTN4, encoding alpha-actinin-4, cause familial focal segmental glomerulosclerosis.* Nat Genet, 2000. **24**(3): p. 251-6.

22.15 Guo, J.K., et al., *WT1 is a key regulator of podocyte function: reduced expression levels cause crescentic glomerulonephritis and mesangial sclerosis.* Hum Mol Genet, 2002. **11**(6): p. 651-9.

16. Saleem, M.A., et al., *A conditionally immortalized human podocyte cell line demonstrating nephrin and podocin expression.* J Am Soc Nephrol, 2002. **13**(3): p. 630-8.

17. Benzing, T., *Signaling at the slit diaphragm.* J Am Soc Nephrol, 2004. **15**(6): p. 1382-91.

18. Tryggvason, K. and J. Wartiovaara, *Molecular basis of glomerular permselectivity.* Curr Opin Nephrol Hypertens, 2001. **10**(4): p. 543-9.

19. Roselli, S., et al., *Early glomerular filtration defect and severe renal disease in podocin-deficient mice.* Mol Cell Biol, 2004. **24**(2): p. 550-60.

20. Shankland, S.J., *The podocyte's response to injury: role in proteinuria and glomerulosclerosis.* Kidney Int, 2006. **69**(12): p. 2131-47.

21. Willis, B.C., R.M. duBois, and Z. Borok, *Epithelial origin of myofibroblasts during fibrosis in the lung.* Proc Am Thorac Soc, 2006. **3**(4): p. 377-82.

22. Reeves, W., J.P. Caulfield, and M.G. Farquhar, *Differentiation of epithelial foot processes and filtration slits: sequential appearance of occluding junctions, epithelial polyanion, and slit membranes in developing glomeruli.* Lab Invest, 1978. **39**(2): p. 90-100.

23. Batlle, E., et al., *The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells.* Nat Cell Biol, 2000. **2**(2): p. 84-9.

24. SJ, S., *The podocyte's response to injury: role in proteinuria and glomerulosclerosis.* Kidney Int, 2006. **69**: p. 2131–2147.

25. Saitou, M., et al., *Mammalian occludin in epithelial cells: its expression and subcellular distribution.* Eur J Cell Biol, 1997. **73**(3): p. 222-31.

26. Yang, Y., M.C. Gubler, and H. Beaufils, *Dysregulation of podocyte phenotype in idiopathic collapsing glomerulopathy and HIV-associated nephropathy.* Nephron, 2002. **91**(3): p. 416-23.

27. Dressler, G.R. and E.C. Douglass, *Pax-2 is a DNA-binding protein expressed in embryonic kidney and Wilms tumor.* Proc Natl Acad Sci U S A, 1992. **89**(4): p. 1179-83.

Journal of Total Science
Volume (5), Issue (19((Mar. 2022)
ISSN: 2518-5799

مجلة العلوم الشاملة
المجلد (5)، العدد (19)، (مارس2022)
ردمد: 2518-5799

DETECTING CORONARY ARTERY DISEASE BY USING ARTIFICIAL NEURAL NETWORK AND GENETIC ALGORITHM

Mrs.NAJAH ABDULHAMED NSER

Higher Institute of Science and Technology Al Azizia

Abstrac

Building up of plaque on the coronary arteries from inside cause the Coronary Artery Disease (CAD), which has a high mortality rate not only in Old ages but also in Middle ages too. To detect this condition efficiently, we need to perform invasive operations, expose to X-rays radiation and the time and cost of examination consideration. Consequently, a lot of researches have been conducted to seek alternative modalities using machine learning and data mining. In this paper, ANN, C 4.5 and CART techniques have developed and presented to classify the potential affected CAD cases from normal ones. This paper uses the Z-Alizadeh Sani dataset for the CAD disease. This dataset contains information on 216 CAD patients among the data of another 87 normal persons. For each person, fifty-five features were collected. We can divide the collected features into 31 logical features and 23 continuous features. Chi-Squared Automatic Interaction Detection (CHAID) was used to select features with only good discriminating capability. The significant resultant features were used to evaluate the presented ANN, C 4.5 and CART classifiers accuracy. We observed that the ANN trained classifier by the dataset features represent the highest specificity, accuracy, and sensitivity of 0.968, 0.950, and 0.838 respectively, and positive predictive value of 95% on, the dataset of Sani Z-Alizad

Keywords— CAD; Coronary Artery Disease; ANN; ANN training; ***Dataset;***

Introduction

This paper is interested in studying the impact of the computer-aided disease detection based on some historical information for other patient suffered from the same disease before. Many researches have collected and shared a dataset for some diseases based on their remarks and spreadsheets for real people suffered from the same disease. Among the most important disease to be taken care of is the Coronary Artery Disease (CAD).

This paper uses the Z-Alizadeh Sani dataset for the CAD disease. This dataset contains information on 216 CAD patients among the data of another 87 normal persons. For each person, fifty-five features were collected. We can divide the collected features into 31 logical features and 23 continuous features. These features represents the patients' data regarding symptoms, physical examinations results and demographic characteristics, echocardiography, the tests of the laboratories, and electrocardiography. The study is concerned with representing the sensitivity, specificity, and accuracy of the ANN repetition against the C 4.5 and CART according to many fault trails.

Moreover, the study is concerned as well with the impact of the number of samples on the methods with the solvers of the linear equations. According to the results of the study, there will be an enhancement in the future studies related to the ANN, C 4.5 and CART; this enhancement as well will help in getting more efficiency and reliability simulation.

The accurate prediction of the patient either having the disease or not will not only decrease the cost of the needed operations before diagnostic the disease but also will relax the need for the other limitations such as radiation, invasive operations, and time of examination.

In the previous work named "Computer Aided decision making for heart disease detection using a hybrid neural network-Genetic Algorithm," the researchers work on a set of 303 patients where 216 of them suffered from CAD [1]. This Data set contains 54 features for each patient.

The researchers first use SVM for features selection to reduce the dataset of 54 correlated features into a minimized dataset of uncorrelated features. They divide the dataset in a manner that they use 90% for training and 10% for test: They apply the genetic algorithm on the reduced variable training set iteratively, to select the best individual as an initial weight for the Artificial Neural Network (ANN).

They use many iterations on the weights of the ANN until a satisfying stop criterion achieved. They use back propagation with MLP structure and a sigmoid exponential function on the ANN. Concerning GA, they use two points of crossover to pass from parents to children then a fixed mutation probability on the generated children ($P_m = 0.2$).

To measure the performance of their algorithm, they introduce three metrics the sensitivity, specificity, and accuracy. In their paper, the sensitivity, specificity, and accuracy are 97, 92, and 93.85 respectively. However, we can easily see that they fall into the trap of overfitting because they used the input to the training samples to ANN as the output of another method.

The motivation for this paper is to enhance the results of the calculations of the ANN, C 4.5 and CART without falling into the trap of overfitting by developing MATLAB programs that professionally prepare the dataset before interring the feature selection process. In order to get the effect of the ANN and its impact on the convergence of the fast diagnosis of the disease.

There are many benefits of applying that model to the study that it is previously eliminating the effect of dependent features from the dataset and consequently enhance the results. Without the need to perform invasive operations, expose to X-rays radiation and the cost and time of examination consideration.

The paper will focus on the problem that calculations of the huge datasets will have more significant sensitivity, specificity and accuracy results, which will have more substantial impacts on the efficiency and reliability of the calculations.

This study is developing a new model for showing the effect of the ANN, C 4.5 and CART. On the other hand, we focus to avoid the overfitting for our model. To get accurate results, we will execute a Mount Carlo simulation about a 60-trial experiment on the test part of the dataset (i.e., 10% of the dataset) in order to ensure the output of the tests are without jitters or uncertainty.

The remainder of this paper organization is illustrated in this paragraph. Section 0 carries the background discussion. In section III, we reviews the related work. The CAD detection different techniques is introduced in Section IV. In section V, we present our proposed method. Sections VI and 0 are the results and the conclusions of the paper, respectively.

BACKGROUND

Dataset Preparation

In order to know the importance of dataset preparation, Martin Goodson who is a data science consultant told a good story about bad data collection that causes death. Machine learning (ML) healthcare project collects data through patient records. ML automatically sort patients according to their health conditions; those who have the lowest death risk and should take antibiotics at home and those who are at high risk of death and should be in the hospital. However, there was one of the most dangerous conditions, where doctors always send patients in this condition to intensive care to minimize death rates. Consequently, ML sends these patients home while they have the most critical condition (i.e., ML heavily depends on a dataset, not the situation). It is the most acute merit that makes algorithm training possible and made machine learning so popular in recent years. On the other hand, a machine will be nearly useless or perhaps even harmful, if doctors cannot make sense of data records despite the collection of actual terabytes of information and data science expertise.

In the machine learning process, data preparation is the most important step due to the flawed dataset. Data preparation, in a nutshell, is a set of procedures that make dataset more suitable for machine learning. In broader terms, the data preparation also includes setting the right data collection mechanism up. Hence, most of the time spent on machine learning is consumed on data preparation procedures.

The Adopted Dataset

This paper uses the dataset of Sani Z-Alizadeh, which contains information on 216 CAD patients among a sample of 303 subjects. Each subject fill in a sheet contain fifty-five features. These features can be divided into 31 logical features and 23 continuous features. Moreover, they cover the patients' data such as physical examinations tests and results,
symptoms, electro- and echo-cardiography, and demographic characteristics. Table 1 shows the chosen features. In the dataset, the patient is diagnosed with coronary artery disease (CAD), if at least one of the right coronary arteries (RCA), left circumflex (LCX), and left anterior descending (LAD) has stenosis of higher than 50%.

Feature Selection

There are four main goals for the Feature selection techniques which are: (1) facilitation of interrupting the models by researchers or users due to their simplifications, [1] (2) saving training times, (3) the burden of dimensionality is avoided, (4) Generalization is enhanced by variance reduction (i.e., reducing the overfitting) [2].

The most feature selection distinguishing characteristic technique is the data elimination without losing a relevant part of information. The eliminated data have either some irrelevant or redundant features [2]. Irrelevant and redundant features are two different situations; because relevant features may be redundant in the existence of another correlated stronger relevant features [3].

Feature selection and feature extraction are two distinctive techniques. Feature extraction technique is used to bring new features from the original features functions, while a features subset is returned in the feature selection technique. The techniques of feature selection are mostly preferred to be used where there are many features and relatively few samples (or data points). For instance, analysis of written text, or having DNA microarray data that contains hundreds of samples, and thousands of features are typical cases for deploying feature selection technique.

Subset Selection

Features subset is estimated using subset selection algorithms for group suitability. Wrappers, Filters, and Embedded are algorithms of subset selection. Wrappers search within the possible features space using search algorithm, then run the model on each subset to assess the subset. However, wrappers have disadvantages of expensive computations and have model overfitting risks. Filters are as similar as Wrappers from the search approach, however, a simpler filter is evaluated instead of evaluating against a model as in wrappers. Embedded algorithms are embedded in and specific to a model.

Greedy hill climbing is the most known search approach that estimates features subset iteratively, then the subset is enhanced and checked if it has a refinement over the previous or not. Scoring metric is essential for evaluation of subsets to grade subset of features. Exhaustive search at some operator require stopping point definition that leads to selection of

the highest score subset features which are close to the satisfactory point of the feature subset, which is generally impractical. The stopping criterion depends on algorithm used, for instance, a score threshold should be exceeded, or program's runtime reach the timeout limit, are examples of possible stopping criteria.

Alternative search-based techniques are another search approach that focuses on the targeted projection pursuits. Its way of operation is to find high score data in low dimensional projection, therefore, the features with the highest projection score in lower dimensional space are selected.

Problems are classified by using two popular filters; correlation and mutual information filters. Both filters are not considered as true matrices or even represent distance measurement in mathematical sense. So, they are considered as score instead, due to their failure in computing any actual distance complied with triangle inequality. The scores compute lie between features set (candidate feature) and the output category desired. In contrast, in some cases filter can be true matrix when they are simple function of mutual information.

Neural network

Artificial neural network (ANN) purpose is to solve problems as if the human brain is the one that solves them. ANN is performed on different applications, including machine translation, computer vision, social network filtering, speech recognition, playing video games and medical diagnosis. However, over time ANN was deviated from biology to execute certain tasks. ANN consists of weights between the output and input variables, which makes the behavior of the network depends on the interaction between output variables and input ones. Generally, ANN is interconnected group of (units, nodes or cells) such that each group is made of three layers, each layer made of processing cells called neurons, as follow:

– The input layer is the first layer that contains the raw received data entered the network.

– The second layer is a hidden layer or layers. The structure of the neural network may have several hidden layers that their performance is set by inputs raw data and their interactions with the remaining layers. Time of activation for a hidden unit is defined by the weights between input raw data and hidden units.

– The output layer is the final one. The performance of the output layer depends on the hidden layers activity and weights [10]. Each designer defines the number of neurons per layer and the number of layers mainly using the try and error process.

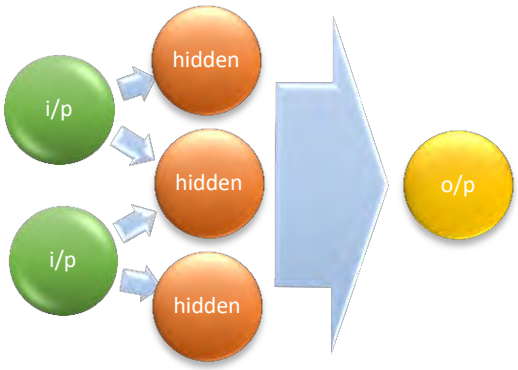


Figure 1 ANN is a group of nodes interconnected with each other

Genetic Algorithm

John Holland was the first one to introduce genetic algorithm that gives approximate optimal solutions for search issues. It is an evolutionary algorithm that uses biology terms such as mutation and inheritance. In genetic algorithms, from combining the appropriate responses of a generation by letting only the living organism's best fit to survive, the optimal solution is obtained [5].

In genetic algorithms, some operators are deployed to produce more qualified members from combination of qualified members of current generation [5]. Here is brief explanation on principle operation of some operators:

– Selection operator: stochastic universal sampling (SUS) and Roulette wheel (RW) in genetic algorithm are examples on selection operator. In this operator, the next population are generated from the current generation members that are more compatible with each other. In this operator, a member of a generation is selected, based on the fitness criterion, in order to take part in the process of reproduction [6].

– Substitution operator: such operator helps in propagating and transferring the members from a generation to the following.

- Recombination operator: such operator substitutes a chosen generation two members substrings with each other using the intersection. Uniform, two-point and single-point crossovers are examples of recombination approaches [7].
- Mutation operator: This operator produces new member by changing the genes of a member of the current generation. Boundary, Gaussian, uniform and non-uniform operators are mutation examples on these methods [8]. For example, a normal distribution random value is summed with the Gaussian operator selected gene.

RELATED WORKS

As discussed in the previous section, there are many methods to extract features from datasets. Data mining extracts of hidden information, relationships and valid patterns from huge datasets [9]. Even though data mining is employed in marketing, insurance, and banking in order to reduce the cost and improve the quality [10]. We target using such data mining in the medical field to achieve the same metrics (i.e., cost reduction and quality improvement) and of course, save lives. Recent years witness joining the medicine field the interest surge to utilize data mining and machine learning to early diagnose diseases [11, 12].

Coronary Artery Disease Classical Diagnosis

The most spread disease for causing death is cardiovascular disease, specifically CAD condition that is the most common condition resulted from this disease. The patient is diagnosed with coronary artery disease (CAD), if at least one of the right coronary arteries (RCA), left circumflex (LCX), and left anterior descending (LAD) has stenosis of higher than 50% [13].

Different methods are used to diagnose CAD, angiography is the most usable method due to its precision. However, this method has drawbacks such as it is expensive and the caused side effects. So, researchers have sought for long time to develop new methods that diagnost such modalities.

According to [1, 3, 14, 15] CAD can be detected by using differenct tried and tested techniques clinically. Invasive coronary roentgenography is the best CAD designation technique. In contrast, during the process of that examination technique, the inserted tube may

take away some plaque from the artery walls or might pierce associate degree artery and lead to embolism, which leads to increased stroke chances. The Exercise Stress Test (EST) is a foremost unremarkably practiced CAD detection non-invasive techniques. In contrast, the required heart rate in EST test is higher than the ability of most patients to reach [16]. Moreover, EST has only 66% CAD detection sensitivity as recorded experimentally [17, 18]. Al Moudi et al. [19] performed scientific reviewing to the CAD detection published research to the diagnostic value of PET/CT, PET, and SPECT. Al Moudi reveals these imaging modalities mean values of accuracy, specificity, and sensitivity of: 88%, 83%, and 85% for PET/CT; 89%, 89%, and 91% for PET; and 83%, 76%, and 82% for SPECT.

However, a restricted examinations repeating is prohibited in these techniques because of radioactive material usage. Examination time and high price are the drawbacks of these techniques. Consequently, safe non-invasive, efficient, radiation-free, speedy and inexpensive strategies are required for CAD diagnose.

The Heart Rate Variability (HRV) technique has been used to diagnose several cardiovascular and non-cardiovascular illnesses [21, 22, 23] by measuring beat to beat intervals within the heart rate [20]. CAD patients and healthy subjects are distinguished from each other in terms of frequency and time domain measurements of HRV such that CAD patients have lower HRV [24]. But these measures are not always effective in distinguishing between healthy and CAD cases appropriately because of the noise sensitivity. Based on the cardiac system nonlinearity assumption, some techniques have been developed to show that the derived HRV nonlinear parameters are helping the detection of CAD [25, 26].

Risk Factors

In fact, an expansive number of examinations have been led in this field [27, 28], with the UCI datasets [29] being the most much of the time used sets. A few other PC based methods for CAD discovery have been condensed in this segment. We saw that the traditional exactness of recognition is just around 70%. So as to enhance this precision, a few research planned a PC supported methodologies on echocardiography of the coronary vein, to extract a few nonlinear features, and utilize the resultant critical features in directed classifiers. A few elements are added

to advance a coronary heart occasion. These hazard variables might be characterized into two classifications, not modifiable and modifiable.

The Not Modifiable Category

This classification incorporates factors that can't be modified by intercession, for example, age, sexual orientation, activities, family ancestry, and hereditary characteristics [2, 29]. Such a category is imperative in separating the malady location highlights. However, no future safety measures can be extricated from them for the following ages. Also, no treatment is accessible to adjust for the impact of these components.

The Modifiable Category

Then again, modifiable hazard factors are those for which either treatment is accessible or in which the proportion of the population exposed can be reduced. Set up, modifiable hazard factors for CAD presently incorporate diabetes, hypertension, smoking, triglycerides, high- and low-density lipoprotein, and cholesterol [30, 31].

Risk Factors Prevention Guidelines

The American Heart Association (AHA) grew new rules and surveyed various hazard factors for CAD [32]. Moreover, results from the European Action on Secondary and Primary Prevention by Intervention to Reduce Events (EUROASPIRE) uncovered the risk factors that are important by different reviews crosswise over European nations [33]. Data mining encourages data investigation utilizing data examination strategies with complex calculations to find obscure patterns. Such calculations incorporate decision trees that have been utilized widely in medical field. The EUROASPIRE research with EUROASPIRE studies (I, II, III) included different European inhabitance and furthermore incorporated extra hazard factors, for example, overweight.

All EUROASPIRE studies were looked into together, and consolidated outcomes were extricated [34]. A fact that outcomes generally, Doctors' recommendations and advices are not followed by their patients. The risk factors that are modifiable after EUROASPIRE survey have changed and concluded as follows [35]:

- For smoking subjects 14% compared to 16% after and before the event;
- For higher blood pressure subjects 22% compared to 26% after and before the event;
- For higher total cholesterol subjects 34% compared to 31% after and before the event;
- For low-density lipoprotein subjects 45% compared to 31% after and before the event.

The EUROASPIRE important risk factors in their survey are cholesterol, blood pressure, and smoking [36, 37]. The risk factor between 15 countries and the use of cardioprotective medication treatments results in high variations in the results [38]. Furthermore, Europe targets raising the preventive care standards with considerably high potential to reduce the death and recurrent disease risks in CAD patients.

Besides, extra perceptions that could be elicited from the database examined in this investigation in regard to the hazard factors that are non-modifiable with correlation to EUROASPIRE study [39] are as follows:

- Female subjects are 14% compared to EUROASPIRE survey result of 24.7%;
- Less than 50 years old subjects are 9% (males only) compared to EUROASPIRE survey result of 23.1%;
- Between 51 and 60 years old subjects are 28% compared to EUROASPIRE survey result of 33.8%;
- Between 61 and 70 years old subjects are 39% compared to EUROASPIRE survey result of 43.1%; and
- Between 71 and 84 years old subjects are 24%.

Computer Based Techniques for CAD Detection

The goal of a few examinations was to build up a data mining framework dependent on various innovations for the appraisal of CAD, including logistic regression [40], affiliation rules [41, 42], fuzzy modelling [43], neural systems [44], and decision trees. The evaluation of CAD-related

hazard factors focusing on the decrease of CAD occasions. The data mining investigation was done for eliciting rules dependent on the previously mentioned hazard factors.

Framingham Risk Score

The cardiovascular 10-year risk estimation of a person is predicted by gender algorithm like Framingham Risk Score. Such score was created dependent on information got from the Framingham Heart Study, to assess the 10-year danger of creating coronary illness. The Framingham Heart Study data have been used to develop such score to help estimating the risk of CAD for 10-year [45]. To evaluate the 10-year cerebrovascular occasions, cardiovascular sickness chances, heart failure, and peripheral artery illness were hence included as ailment results for the year CAD score [46].

A few examinations utilized the Framingham formula to portray the inhabitance in a region or nation [47], though different investigations inspected the highlights of accessible Framingham-based hazard estimation [48]. Wang et al. [49] utilized the hazard factors sex, age, HDL, cholesterol, diabetes, smoking, and blood pressure to foresee CAD. They utilized the Framingham work and reasoned that the conventional hazard factors have distinctive degrees of effect or potentially that different elements are adding to chance.

Decision Tree Algorithms

In this thesis, we use the training dataset significant features to build a tree that defines the classes' extraction rules [50]. Then after the determination of the class, we apply these rules to the dataset testing part. Podgorelec et al. [51] pointed to algorithms based on decision trees because of their properties of giving high classification reliability, accuracy and effectivity in the while keeping the gathered knowledge simply presented, which particularly useful in medicine decision-making processes.

Ordonez [52, 53] examined decision trees and affiliation guidelines to foresee CAD dependent on the hazard factors age, cholesterol, smoking, and sex. Ordonez [54] utilizing the C4.5 decision tree algorithm and affiliation rules for the expectation of cardiovascular malady dependent on 24 hazard factors archived that set rules corporate less complex predictive principles than decision tree rules [44]. Our gathering on a

comparable database with this examination [56] additionally explored the benefits of the associated rules in the investigation of CAD hazard factors. The outcomes with respect to the most vital hazard factors were comparable.

Gamberger et al. [57] utilized a choice help technique to target high-hazard groups for CAD utilizing hazard factors like hypertension, smoking, and cholesterol. Decision-tree-based programming instruments were produced in [58, 59] to help in the analysis of CAD. Rao et al. [60] introduced a system to make organized clinical information for CAD. Zavrsnik et al. [48] utilized decision trees and made the ROSE apparatus to be applied in cardiology.

Besides, Polat et al. [61] created decision- tree-based models for the arrangement of CAD, accomplishing a right characterization score of 82%. Additionally, Pavlopoulos et al. [62] utilized the C4.5 calculation decision tree [42] to examine the diverse heart sound features, which help clinicians in improving CAD diagnostic.

Logistic Regression

Logistic regression analyzes dataset with independent variables statically to seek finding a reliable output. The estimated result is forced to convey in just two conceivable results (i.e., dichotomous).

The dependent variable in logistic regression is either dichotomous or binary, i.e., it contains only coded data like 1 (pregnant, success, TRUE, etc.) or 0 (non-pregnant, failure, FALSE, etc.). In order to find the most fit logistic regression model, we need to find the correlations in a set of variables (independent, explanatory or predictor) and between the binary interest characteristic of the outcome or response variable (i.e., dependent variable).

Tsien et al. [42] logistic regression building trees and likewise utilized grouping three distinct models for myocardial infarction (MI) and furthermore analyzing the centrality of these models. Also, their investigation showed that characterization trees that have points of interest over logistic regression models, may perform correspondingly to logistic regression models to find patients with MI.

Artificial Neural Network

Zeinab Arabasadi et al. [62] works on a set of 303 patients where 216 of them suffered from CAD. This Data set contains 54 features for each patient. The researchers first use SVM for features selection to reduce the dataset of 54 correlated features into a minimized dataset of uncorrelated features. They divide the dataset in a manner that they use 90% for training and 10% for test: They apply the genetic algorithm on the reduced variable training set iteratively, to select the best individual as an initial weight for the Artificial Neural Network (ANN). They use many iterations on the weights of the ANN until a satisfying stop criterion achieved. They use back propagation with MLP structure and a sigmoid exponential function on the ANN. Concerning GA, they use two points of crossover to pass from parents to children then a fixed mutation probability on the generated children. ($P_m = 0.2$). Three metrices were introduced to measure algorithm's performance:

$$Sensitivity = \frac{TP}{TP + FN} \quad (1)$$

$$Specificity = \frac{TN}{TN + FP} \quad (2)$$

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \quad (3)$$

Where the TP is the mean of the correctly identified number of samples, FP is the mean of the incorrectly identified number of samples, TN is the mean of the correctly rejected number of samples, and FN is the mean of the incorrectly rejected number of samples.

PROPOSED METHOD

In this section, we investigate the CAD risk evaluation based on decision trees, ANN and data mining. Our aim is to find the extracted CHAID classification rule risk factors. The extracted features will help enabling the better patient management to target events reduction, and also, decreasing of the therapy cost, due to the interventions expected restriction in only the necessary cases.

We will work on the dataset composed of 303 samples, among those samples 216 suffer from CAD. For each sample, we have collected 55 variables most of them are binary variables. This dataset is the same used on the previous work so a comparison between results can be made.

Chi-Squared Automatic Interaction Detection

Instead of using SVM for feature selection, we will select Chi-Squared Automatic Interaction Detection (CHAID) algorithm for this task. CHAID originally proposed by Kass in 1980, it is the classification tree oldest methods. Therefore, the definition of done (DOD) for this step is to have uncorrelated variables with small numbers instead of the initial number of correlated variables.

Preparing the Dataset

The dataset preparation is done by firstly changing any continuous predictors into categorical ones by approximately divide the continuous distributions until an observations with equal number is resulted as some categories distributions. Then the least significantly dependent variables is omitted by cycling through the predictors and determining the predictor pair for each categories that is different. Finally, split the smallest p-value predictor, which gives the split with the most significant effect and then continuing that process until all splits are performed.

Separate the Train and Test Data

We randomly divide the 303 samples dataset in a manner to use 90% for feature selection and ANN training (i.e., 266 ~ 276 samples at random) and rest 10% for testing our ANN output. The cases that are chosen for feature selection are cross-tabulated for each pair of them. Table 1 shows the general idea of cross-tabulating between two variables (e.g., A and B).

Remove Dependence

We test the dependence among the 55 features columns using the chi-square function between every two features. The output p-value determines the mutual dependence ratio from 0 (i.e., no mutual dependence) to 1 (i.e., total dependence). A feature is excluded if another selected feature has a total dependence with it. Z-Alizadeh Sani dataset has seven dependent features to be excluded. Although we exclude these features according to the training part of the dataset, we have to exclude

them also from the test part of the dataset before beginning with the test procedures.

Feature Selection

After removing the dependent variables, the chi-square function is computed to check the percentage of dependence between each variable and the disease state variable. The variable higher dependence on the disease state variable means a higher probability of suffering from CAD for subjects having that variable. Then we sort the p-value between all variables and the disease state variable in descending order. For all p-values that are higher than the mean value of p plus quarter the variance of p, we include them in the chosen features. Otherwise, we discard the variable for lack of dependency.

Artificial Neural Network

To choose the best individuals of the ANN or CART tree, we will use GA before training the ANN.

Genetic Algorithm

Genetic Algorithm (GA) generates children from the initial reduced variables dataset in the following two steps:

Crossover

Where we will use the multipoint crossover in the following manner: (1) Some crossover points are randomly selected between 1, 2 or 3 points, and (2) The position of crossover points is randomly selected.

Then we make a comparison in term of number of iteration before stability, between the proposed method in GA and the method that use fixed crossover points (2 points) and location. RW method was deployed for parent selection, and for substitution, the current generation best 10% individuals were chosen to supplant with the next generation worst individual. Recombination was done in each iteration such that the two-point crossover used with $P_c = 1$. Figure 2 demonstrates a crossover with simply two-point.

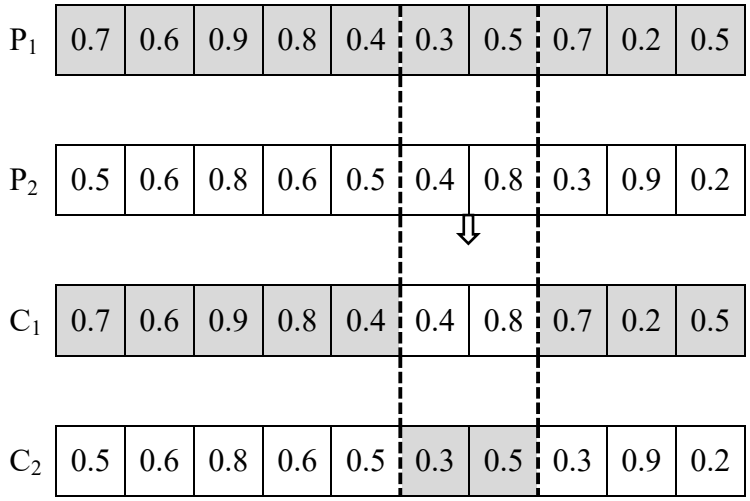


Figure 2 crossover with two-point. The selected parents are P_1 and P_2 while the generated childrens are C_1 and C_2.

Mutation

Where the Gaussian operator is used. The mutation in each gene has a probability of $P_m = 0.2$. Figure 3 shows the simple mutation. Changing in Weights is from −1 to 1 and σ equals 1. Each weight is shown in one gene of the chromosome while the chromosome carries all the weights of ANN.

Figure 3 mutation. The selected parent is P. the mutation probability is P_m, and the generated child is C.

Training of the ANN

We use error back propagation algorithm with sigmoid exponential function and MLP structure. The initial weights employed in the network structure is a factor that affect the ANN performance. Therefore, the primary weights is enhanced in the proposed model to ameliorate the neural network performance.

In this study, genetic algorithm is used to identify the initial weights of neural network, and neural network is identified by using the data assigned for training. In this network, we deployed a single hidden layer in a feedforward structure. The output layer is disease diagnosed pertains. We have a single output specifying the patient CAD condition, a single hidden layer with six neurons and 28 inputs in the input layer. Six hidden neurons are selected due to its performance on our training data

Overfitting

An overfitting hypothesis h in the hypothesis space $\mathcal{H}$ can be detected if the training data has another hypothesis h' in the same hypothesis space $\mathcal{H}$. Such that, in the training examples, h' has higher error compared to h. In contrast, in the practical instances, h has higher error compared to h'. In other words, overfitting treats noise data as important as other data in the training dataset and consequently produce a function that exactly fits them. Then, in testing the noise data are different which produce a low-quality output from that function. Instead, we must focus on the real data without the noise, which has a little smaller quality output in training but with higher output in testing.

FINDINGS AND RESULTS

We use MATLAB 2016b to generate custom-made codes to test our chosen classification methods. We select the needed features using CHAID, and classification by (ANN, C 4.5, and CART). Finally, we compare the results given by (ANN, C 4.5, and CART). The metric used for this comparison is the values of sensitivity, specificity, and accuracy.

Importing the Dataset

We import the chosen dataset (the dataset of Sani Z-Alizadeh), which contains 303 rows (information on patients). Fifty-five columns contain features that were collected for each patient. While importing we divide the features into 23 feature that contains numbers (directly imported) and

31 logical features that contains either binary or chars (e.g., 1, 0, yes, y, no, n, etc.) which needed to be converted into numbers before importing as shown in Table 1.

Separate the Train and Test Data

As shown in Appendix A, we have separated the train and test data from the imported dataset in a random manner. We choose a probability of 90% for selecting the rows in the train set, and the remainder is added to the test set. Such selection ensures the fairness and high quality output test results. We have selected 267 rows at random for the train and 36 rows for the test.

Remove Dependence

We test the dependence among the 55 columns (i.e., features) using the chi-square function between every two features. The output p-value determines the mutual dependence ratio from 0 (i.e., no mutual dependence) to 1 (i.e., total dependence). A matrix with 55×54 dimensions is created with its diagonal is the self-dependence (i.e., always equals one), each $a_{ij}\ and\ a_{ji}$ represent the mutual dependence between column i and column j. Column j feature is excluded if $a_{ij}=1$ and $i<j$ (i.e., Column j feature has a total dependence on Column i feature).

Table 1 Conversion of logical features into notation numbers

N	Female	Cad	N	0
Y	male	Normal	Mild	1
			Moderate	2
			Severe	3

Table 2 seven dependent features to be excluded

Feature	Sex	BMI	Obesity	BP	Atypical	FBS	Neut
Feature Number	4	5	11	18	28	39	51

Z-Alizadeh Sani dataset has seven dependent features to be excluded as shown in Table 2. Although we exclude these features according to the training part of the dataset, we have to exclude them also from the test part of the dataset before beginning with the test procedures.

Feature Selection

After removing the seven dependent features, the chi-square function is computed to check the percentage of dependence between each feature and the disease stat feature. The resultant array of 48×1 is created, each a_i represent the mutual dependence between feature i and the disease stat feature. Features with higher dependence on the disease stat feature means a higher probability of suffering from CAD for subjects having that features. Then we sort the p-value between all features and the disease stat features in descending order. For all p-values that are higher than the (mean value of p plus quarter the variance of $p \approx 0.75$), we include them in the chosen features. Otherwise, we discard the feature for lack of dependency (shown in gray in the table).

Artificial Neural Network

We generate a MATLAB code for the training of the ANN with the selected training data (i.e., $267\ row \times 28\ feature$). We notice that the output each time we run the code is slightly difference because of the small number of samples (i.e., 267). Therefore, we decided to use the Monte Carlo simulation to decrease the jitter in the output. Due to the intervention of random variables that makes hard prediction probability for various outcomes in a process, Monte Carlo simulations are used for modeling. The technique's advantages are to grasp the impact of risk, and the uncertainty in prediction and forecasting models.

The results show that the accuracy, sensitivity, and specificity are 0.950, 0.838, and 0.968 respectively. By halving the input rows, we get 0.825, 0.618 and 0.898 for accuracy, sensitivity, and specificity. While by quartering, we get 0.819, 0.533, and 0.933 for accuracy, sensitivity, and specificity as shown in Figure 7.

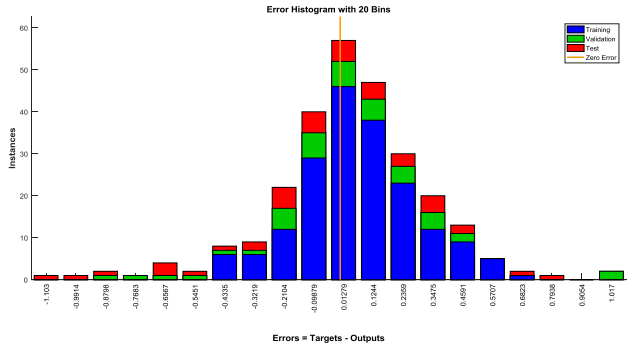


Figure 4 ANN errors

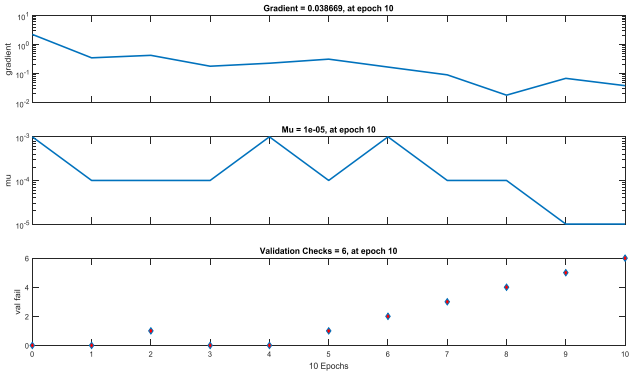


Figure 5 ANN gradient, mu and validation failed

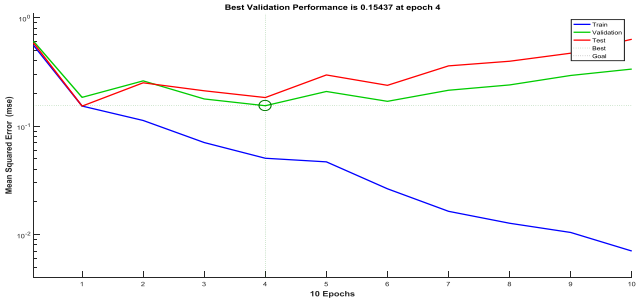


Figure 6 ANN Performance

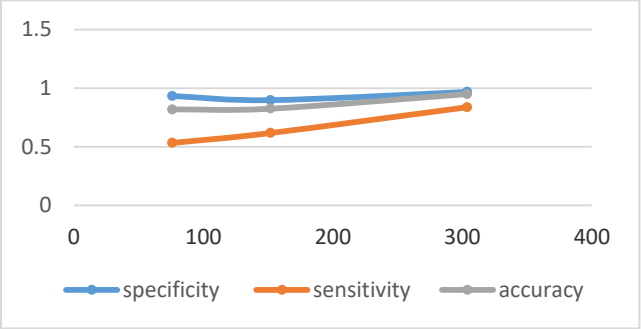


Figure 7 accuracy, sensitivity, and specificity versus the number of samples

Conclusion

In this work, we have augmented the neural network performance by proposing a CAD diagnostic tool with the aid of computers. The datasets used to test the CAD presence are not only for affected patients but also for normal subjects to track the performance improvement. The neural network performance for CAD detection, which can represent the highest specificity, accuracy, and sensitivity detection of CAD for 95%. Clinical use required high accuracy techniques such as the proposed one to make it suitable for assisting cardiologists and to make a good estimation on the CAD absence or presence. Especially, detection of CAD by that algorithm is without the need to perform invasive operations, expose to X-rays radiation and the cost and time of examination consideration (i.e., the majority of diagnostic side effects). We expect that the decision trees algorithms based on the process of data mining will continue helping to identify subgroups with high risk for the development of events in the future and may be a determining factor in the treatment choice (i.e., surgery or angioplasty). In addition, rules and models can help reducing the morbidity of diabetes (CAD) and hopefully mortality.

As a future work, our proposed method performance may be enhanced if we use methods other than genetic algorithm. There are a lot of evolutionary, intelligence and powerful algorithms, besides the genetic algorithm, such as particle swarm optimization and evolution strategy.

On the other hand, other neural networks parameters could be compared and tested. This work parameters need optimiziation in terms

of factor of momentum and of learning. Finally, this algorithm must check other features from newer datasets. However, larger datasets need more investigation with other criteria.

References

[1] S. Kara, F. Dirgenali, "Computer Aided decision making for heart disease detection using a hybrid neural network-Genetic Algorithm," In Proceedings of International Journal of Expert systems with applications, Elsevier, vol. 32, no. 2, pp. 632–640, 2007.

[2] A. Rajkumar, G.S. Reena, "Diagnosis of heart disease using datamining algorithm," In Proceedings of Global journal of computer science and technology, vol. 10, no. 10, pp. 38–43, 2010.

[3] A. T. Azar, S. A. El-Said, A. E. Hassanien, "Fuzzy and hard clustering analysis for thyroid disease," In Proceedings of International Journal of Computer Methods and Programs in Biomedicine, Elsevier, vol. 111, no. 1, pp. 1–16, 2013.

[4] C. Nikias, A. Petrnhjm, Higher-Order Spectral Analysis, Digital Signal Processing Handbook, Englewood Cliffs, NJ, 1999.

[5] H. G. Lee, K. Y. Noh, K. H. Ryu, "A data mining approach for coronary heart disease prediction using HRV features and carotid arterial wall thickness," In Proceedings of International Conference on Bio Medical Engineering and Informatics, IEEE Computer Society, vol. 1, no. 1, pp. 200–206, 2008.

[6] H. G. Lee, K. Y. Noh, K. H. Ryu, "A datamining approach for coronary heart disease prediction using HRV features and carotid arterial wall thickness," In Proceedings of International Conference on Biomedical Engineering and Informatics, IEEE, vol. 1, no. 1, pp. 200–206, 2008.

[7] H. G. Lee, K. Y. Noh, K. Ho, "Mining bio signal data: coronary artery disease diagnosis using linear and nonlinear features of HRV, emerging technologies in knowledge discovery and datamining," In Proceedings of International Conference on Knowledge Discovery and Data Mining in Computer Science, vol. 48, no. 19, pp. 218–228, 2009.

[8] H. G. Lee, W. S. Kim and others, "Coronary artery disease prediction method using linear and nonlinear feature of heart rate variability in three recumbent postures," In Proceedings of Journal of Information Systems Frontiers, Springer, vol. 11, no. 4, pp. 419-431, 2009.

[9] I. Babaoˇglu, O. Fındık, M. Bayrak, "Effects of principle component analysis on assessment of coronary artery diseases using support vector machine," In Proceedings of International Journal of Expert Systems with Applications, Elsevier, vol. 37, no.3, pp. 2182–2185, 2010.

[10] I. Babaoglu, O. Findik, E. Ulker, "A comparison of feature selection models utilizing binary particle swarm optimization and genetic algorithm in determining coronary artery disease using support vector machine," In Proceedings of International Journal of Expert Systems with Applications, Elsevier, vol. 37, no. 4, pp.3177–3183, 2010.

[11] G. Manis, S. Nikolopoulos, A. Alexandridi, C. Davos, "Assessment of the classification capability of prediction and approximation methods for HRV analysis," In Proceedings of International Journal of Computers in Biology and Medicine, Elsevier, vol. 37, no. 5, pp. 642–654, 2007.

[12] I. Babaoglu, O. Findik, M. Bayrak, "Effects of principle component analysis on assessment of coronary artery diseases using support vector machine," In Proceedings of International Journal of Expert Systems with Applications, Elsevier, vol. 37, no. 3, pp. 2182–2185, 2010.

[13] I. Babaoglu, O. K. Baykan, N. Aygul, K. Ozdemir, and others, "Assessment of exercise stress testing with artificial neural network in determining coronary artery disease and predicting lesion localization," In Proceedings of International Journal of Expert Systems with Applications, Elsevier, vol. 36, no. 2, pp. 2562–2566, 2009.

[14] J. F. Box, "Guinness, gosset, fisher, and small samples," In Proceedings of Journal of Statistical Science, JSTOR, vol. 2, no. 2, pp. 45–52, 1987.

[15] D. Itchhaporia, R. Almassy, L. Kaufman and others, "Artificial neural networks can predict significant coronary disease," In Proceedings of International Journal of the American College of Cardiology, vol. 25, no. 2, pp. 328A, 1995.

[16] I. H. Witten, E. Frank, Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, 2016.

[17] I. Maglogiannis, E. Loukis, E. Zafiropoulos, and others, "Support vectors machine-based identification of heart valve diseases using heart sounds," In Proceedings of International Journal of Computer methods and programs in biomedicine, Elsevier, vol. 95, no. 1, pp. 47–61, 2009.

[18] J. Han, M. Kamber, J. Pei, Data Mining: Concepts and Techniques, Morgan Kaufmann, San Francisco, CA, 2005.

[19] J. Soni, U. Ansari, D. Sharma, S. Soni, "Predictive data mining for medical diagnosis: an overview of heart disease prediction," In Proceedings of International Journal of Computer Applications, vol. 17, no. 8, pp. 43–48, 2011.

[20] J. T. Bigger Jr., J. L. Fleiss, R. C. Steinman, L. M. Rolnitzky, et al., "RR variability in healthy, middle-aged persons compared with patients with chronic coronary heart disease or recent acute myocardial infarction," In Proceedings of journal of Circulation, vol. 91, no. 7, pp. 1936–1943, 1995.

[21] K. Karamanos, S. Nikolopoulos, G. Manis, et al., "Block entropy analysis of heart rate variability signals," In Proceedings of International Journal of Bifurcation and Chaos, vol. 16, no.9, pp. 2093–2101, 2006.

[22] L. Hu, G. Hong, J. Ma, X. Wang, H. Chen, "An efficient machine learning approach for diagnosis of paraquat-poisoned patients," In Proceedings of International Journal of Computers in biology and medicine, Elsevier, vol. 59, no. 14, pp. 116–124, 2015.

[23] M. A. Karaolis, J. A. Moutiris, D. Hadjipanayi, C. S. Pattichis, "Assessment of the risk factors of coronary heart events based on data mining with decision trees," In Proceedings of

IEEE Transactions on information technology in biomedicine, vol. 14, no. 3, pp. 559–566, 2010.

[24] M. Alamoudi, Z. Sun, N. Lenzo, “Diagnostic value of SPECT, PET and PET/CT in the diagnosis of coronary artery disease: a systematic review,” In Proceedings of Biomedical Imaging and Intervention Journal, vol. 7, no. 2, pp. 333-335, 2011.

[25] D. T. Larose, “Decision trees,” In Proceedings of International journal in Discovering Knowledge in Data: An Introduction to Data Mining, Wiley Interscience, pp.108-126, 2004.

[26] M. K. Biswas, T. Ghose, S. Guha, P. K. Biswasa, “Fractal dimension estimation for texture images: a parallel approach,” In Proceedings of Pattern Recognition Letters, vol. 19, no. 4, pp. 309–313, 1998.

[27] M. Karimi, R. Amirfattahi, S. Sadri, S. A. Marvasti, “Noninvasive detection and classification of coronary artery occlusions using wavelet analysis of heart sounds with neural networks,” In Proceedings of 3rd IEEE International Seminar on Medical Applications of Signal Processing, pp. 117–120, 2005.

[28] M. Mirmehdi, X. Xie, J.S. Suri, Handbook of Texture Analysis, Imperial College Press, London, UK, 2009.

[29] M. Shouman, T. Turner, R. Stocker, “Using decision tree for diagnosing heart disease patients,” In Proceedings of the Ninth Australasian Data Mining Conference, Australian Computer Society, vol. 121, pp. 23–30, 2011.

[30] P. Gifani, H. Behnam, Z. A. Sani, “Noise reduction of echocardiographic images based on temporal information,” IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 61, no. 4, pp. 620–630, 2014.

[31] R. Alizadehsani, J. Habibi, B. Bahadorian, H. Mashayekhi, and others, “Diagnosis of coronary arteries stenosis using data mining,” In Proceedings of International Journal of medical signals and sensors, vol. 2, no. 3, pp. 153–159, 2012.

[32] D.T. Larose, Decision Trees, in Discovering Knowledge in Data: An Introduction to Data Mining, Wiley Interscience, NJ, USA, 2004.

[33] K. Antanavicius, A. Bastys, J. Bluzas, et al., "Nonlinear dynamics analysis of electrocardiograms for detection of coronary artery disease," In Proceedings of International Journal of Computer Methods and Programs in Biomedicine, Elsevier, vol. 92, no. 2, pp. 198–204, 2008.

[34] K. Polat, S. Güne, "A hybrid approach to medical decision support systems: combining feature selection, fuzzy weighted preprocessing and AIRS," In Proceedings of International Journal of Computer methods and programs in biomedicine, Elsevier, vol. 88, no. 2, pp. 164–174, 2007.

[35] K. Polat, S. Sahan, S. Güne, "Automatic detection of heart disease using an artificial immune recognition system (AIRS) with fuzzy resource allocation mechanism and k-nn (nearest neighbor) based weighting preprocessing," In Proceedings of International Journal of Expert Systems with Applications, Elsevier, vol. 32, no. 2, pp. 625–631, 2007.

[36] K. Qtsuka, G. Cornélissen, F. Halberg, "Circadian rhythmic fractal scaling of heart rate variability in health and coronary artery disease," In Proceedings of Journal of Clinical Cardiology, Wiley Online Library, vol. 20, no. 7, pp. 631–638, 2009.

[37] K. Srinivas, B.K. Rani, A. Govrdhan, " Applications of data mining techniques in health care and prediction of heart attacks," In Proceedings of International Journal on Computer Science and Engineering, vol. 2, no. 2, pp. 250–255, 2010.

[38] K. C. Chua, V. Chandran, et al., "Analysis of epileptic EEG signals using higher order spectra," In Proceedings of International Journal of Medical Engineering and Technology, Taylor and Francis, vol. 33, no. 1, pp. 42–50, 2009.

[39] K. C. Chua, V. Chandran, U. R. Acharya, C. M. Lim, "Cardiac state diagnosis using higher order spectra of heart rate

variability," Journal of Medical Engineering and Technology, vol. 32, no. 1, pp. 145–155, 2006.

[40] N. A. Setiawan, P. A. Venkatachalam, A. F. M. Hani, "Diagnosis of coronary artery disease using artificial intelligence based decision support system," International Conference on Man-Machine Systems, Batu Ferringhi, Penang, 2009.

[41] N. Lavesson, A. Halling, M. Freitag, et al., "Classifying the severity of an acute coronary syndrome by mining patient data," The Swedish AI Society Workshop, Linköping University Electronic Press, no. 35, pp. 55–63, 2009.

[42] N. Lavraˇc, "Selected techniques for data mining in medicine," In Proceedings of International Journal of Artificial intelligence in medicine, Elsevier, vol. 16, no. 1, pp. 3–23, 1999.

[43] R. Alizadehsani, J. Habibi, M. J. Hosseini, H. Mashayekhi, and others, "A data mining approach for diagnosis of coronary artery disease," In Proceedings of International Journal of Computer methods and programs in biomedicine, Elsevier, vol. 111, no. 1, pp. 52–61, 2013.

[44] R. M. Haralick, K. Shanmugam, I. Dinstein, "Textural features for image classification," In Proceedings of IEEE Transactions on Systems, vol. 3, no. 6, pp. 610–621, 1973.

[45] R. O. Bonow, D. L. Mann, D. P. Zipes, P. Libby, Braunwald's Heart Disease: A Text-book of Cardiovascular Medicine, Elsevier Health Sciences, 9ed, vol. 2, 2012.

[46] S. A. Izad Shenas, B. Raahemi, M. Hossein Tekieh, and others, "Identifying high-cost patients using data mining techniques and a small set of non-trivial attributes," In Proceedings of International Journal of Computers in biology and medicine, Elsevier, vol. 53, no. 18, pp. 9–18, 2014.

[47] S. Arafat, M. Dohrmann, M. Skubic, Classification of coronary artery disease stress ECGs using uncertainty modeling, Congress on Computational Intelligence Methods and Applications, IEEE, pp. 1-4, 2005.

[48] S. Zhonghua, "Multislice CT angiography in coronary artery disease: technical developments, radiation dose and diagnostic value," In Proceedings of World Journal of Cardiology, vol. 26, no. 10, pp. 333–343, 2010.

[49] D. F. Specht, "Probabilistic neural networks," In Proceedings of International Journal of Neural Networks, Elsevier, vol. 3, no. 1, pp.109–118, 1990.

[50] M. Sugeno, Industrial Applications of Fuzzy Control, Elsevier Science Pub. Co, New York, NY, USA, 1985.

[51] T. Ojala, M. Pietikäinen, T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," In Proceedings of IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 971–987, 2002.

[52] Tan, P-N and others, Introduction to Data Mining, Pearson Education, India, 2007.

[53] U. R. Acharya, O. Faust, N. A. Kadri, and others, "Automated identification of normal and diabetes heart rate signals using nonlinear measures," In Proceedings of International Journal of Computers in biology and medicine, Elsevier, vol. 43, no. 10, pp. 1523–1529, 2013.

[54] U. R. Acharya, S. V. Sree, M. M. R. Krishnan, et al., "Automated classification of patients with coronary artery disease using grayscale features from left ventricle echocardiographic images," In Proceedings of international journal of Computer methods and programs in biomedicine, Elsevier, vol. 112, no. 3, pp. 624–632, 2013.

[55] U. R. Acharya, N. Kannathal, S. M. Krishnan, "Comprehensive analysis of cardiac health using heart rate signals," In Proceedings of international journal of Physiological Measurement, vol. 25, no. 5, pp. 1139–1151, 2004.

[56] V. K. Sudarshan, U. R. Acharya, and others, "Data mining framework for identification of myocardial infarction stages in ultra-sound: a hybrid feature extraction paradigm (Part 2)," In

Proceedings of International Journal of Computers in biology and medicine, Elsevier, vol. 71, no. 20, pp. 241–251, 2016.

[57] W. Kim, S. Jin, Y. Park, H. Choi, "A study on development of multi-parametric measure of heart rate variability diagnosing cardiovascular disease," In Proceedings of Journal of World Congress on Medical Physics and Biomedical Engineering, Springer, vol. 14, no. 7, pp.3480–3483, 2007.

[58] Y. Isler, M. Kuntalp, "Combining classical HRV indices with wavelet entropy measures improves to performance in diagnosing congestive heart failure," In Proceedings of international journal of Computers in Biology and Medicine, Elsevier, vol. 37, no. 10, pp. 1502–1510, 2007.

[59] Z. Alizadeh Sani, A. Shalbaf, H. Behnam and others, "Automatic computation of left ventricular volume changes over a cardiac cycle from echocardiography images by nonlinear dimensionality reduction," In Proceedings of International Journal of digital imaging, Springer, vol. 28, no. 1, pp. 91–98, 2015.

[60] Z. Zhao, C. Ma, "An intelligent system for noninvasive diagnosis of coronary artery disease with EMD-TEO and BP neural network," In Proceedings of International Workshop on Education Technology and Training and International Workshop on Geoscience and Remote Sensing, IEEE, vol. 2, no. 2, pp. 631–635, 2008.

[61] M. G. Tsipouras, T. P. Exarchos, D. I. Fotiadis, et al., "Automated diagnosis of coronary artery disease based on data mining and fuzzy modeling," In Proceedings of IEEE Transactions on Information Technology in Biomedicine, vol. 12, no. 4, pp. 447–458, 2008.

[62] T. Ojala, M. Pietikäinen, D. Harwood, "A comparative study of texture measures with classification based on feature distributions," In Proceedings of International Journal of Pattern Recognition, Elsevier, vol. 29, no. 1, pp. 51–59, 1996.

Journal of Total Science
Volume (5), Issue (19((Mar. 2022)
ISSN: 2518-5799

مجلة العلوم الشاملة
المجلد (5)، العدد (19)، (مارس2022)
ردمد: 2518-5799

The Environmental and Occupational Health impacts of Unconventional Petroleum Industry in Libya

[MUSTAFA HUSAIN AM ELMAGHRBI]

[ABDURAHIM MUSTAFA E GHANIA]

[Higher Institute of Sciences and Technology – Azizia]

Abstrac

The purpose of the study involved the development of an understanding related to The Environmental and Occupational Health Influences of the Unconventional Petroleum Industry in Libya using secondary research based on qualitative methodology. It was found through the study that this kind of industry has occupational hazards include exposure of truck drivers to large amounts of silica moreover the workers are at the risk of exposure to oil and gas from spills that may cause damage to the respiratory system, and liver, furthermore it also has the potential for causing decreased immunity, and an increased risk of cancer,In addition to soil and surface water pollution, oil leakage into groundwater, and noise from machinery, Libya is recommended to perform voluntary initiatives that is a significant set of tools that range from international agreements and programmes to national policies, and legislation as well as regulation, and financial sector that has an enhanced involvement in the fulfilment of the requirements associated with lending and investment that are vital for the improvement of the practises related to the sustainable development and the performance of industrial activities, and implementation of the Total Quality Management (TQM) program that may contribute to the improvement of the environmental and occupational health systems that in turn, will promote improvement in the quality of environment in Libya.

Introduction

The process used for the mining of oil and gas with the help of unconventional methods from underground formations has been identified to involve injection of increased volumes in relation to freshwater, and chemical substances, as well as proppants at elevated pressure into profound formations of geology as it promotes the enhancement of the mining of hydrocarbons, it may be regarded as notable that this technique has been recognised to have provided an increased potential for the manufacture of domestic oil and gas that has been identified to be a significant method for the reduction in the level of dependency on foreign energy sources, and the replacement of dirtier

energy sources including coal, as well as the generation of advance jobs and the development of economy but simultaneously, unconventional oil and gas (UOG) has been observed to be a major source of the development of a variety of threats to health and the environment, that includes concerns associated with the quality and quantity of water, in addition to the problems related to health and safety of the workers, in addition to the problems related to quality of air and leakage of methane, as well as the effects of health that includes the populations that are vulnerable, it further involves physical threats, and influences on the community, as well as the influence of climatic change (APHA, 2018).

Research Aim

This research is aimed to develop an understanding related to The Environmental and Occupational Health Influences of the Unconventional Petroleum Industry in Libya

Research Objectives

- To explore the Unconventional mining in Petroleum Industry in Libya
- To identify the Environmental Influences of Unconventional Petroleum Industry in Libya
- To study Occupational Health Influences of Unconventional Petroleum Industry in Libya
- To provide recommendations for improved practises related to Oil mining

Research Question

What are the influences of the Unconventional Petroleum Industry on the Environmental and Occupational Health in Libya?

Literature Review

Hallett and Clark-Lowes (2016) found that the initial licences for the exploration of the hydrocarbons were granted in 1955 in Libya while the investigation was initiated in Cyrenaica, and the areas near to the Algerian frontier that rapidly substituted to Sirt Basin after several findings at Zahrah-Hufrah, and Zaltan as well as Amal, by the end of 1963, nearly 20 billion barrels were found related to the reserves that were identified as recoverable and the manufacture was initiated in 1961

that grew to about 3 million barrels daily in 1969 that led to the finding of 14 giant fields, while the Libyan revolution in 1969 caused the reduction of activity as the result of the establishment of the state participation by the government in all licences and nationalised the Libyan interests of several organisations and in 1974 a dominant position was held by state companies while the manufacture reduced to nearly 1.3 million barrels daily, this level was identified to be constant for several years, it may also be regarded as notable that the drilling for exploration also degenerated from over one hundred explorations of wildcat wells every year that are nearly to 30, whereas, the introduction of the sharing licences was performed in 1974, whereas, they received minor modifications until 2005. The authors have further discussed the finding of The Bouri field that occurred in 1977 and was recognised to be a major offshore oil field found in the Mediterranean, whereas the occasions and approvals associated with politics have caused a restriction to growth during the 1980s to 1990s, but in 2005, an increase in the activity was observed with the removal of approvals and the formation of an advance structure related to licencing that led to a variety of important findings in the Murzuq Basin and offshore while the level of manufacture increased to greater than 1.8 million barrels daily and the explorations of wildcat wells increased 50 on an annual level for the first time since 1969.

According to Blake (2016), the manufacture of oil and gas (O&G) may be regarded as a catalyst for the growth of the economy that suggests considerable profits as a source of energy has increased affordability, and as a feedstock for chemicals that promotes its utilisation in a variety of products, furthermore, it has also led to increased scrutiny by the policymakers, and regulators, as well as environmentalists, and advocates for public health that cite both potential effects in relation to environmental and public health. It has been indicated by the authors that constant efforts have been made for the improved understanding in relation to the potential problems, and the maximisation of the benefits, as well as the management of the threats that have been observed to be complicated as a result of the complexity of the operations associated with the industry, and the inconsistency of research, moreover, the complex inter reliance of site characteristics, and the utilisation of the chemical, as well as the manufacture, has been

identified to have enhanced interpretations in associations with the implications associated with the environment and health.

Unconventional methods for petrol and natural gas mining have demonstrated a rapid global expansion as several countries have been expressing interest in the development of such sources (Werner et al., 2015). Krupnick and Echarte (2017) have discussed that a variety of studies previously conducted have demonstrated an association between health effects and unconventional development of oil and gas development through the collection of information in association with self-identified effects of health that contribute to the assessment of the relationship that exists between exposure to unconventional development and reported or perceived outcomes related to health. The authors have added that the self-reported symptoms have been recognised to be potentially subject to increased bias, as the people that live near shale development have more knowledge related to its potential effects on the health of people and may therefore be more attuned to their own symptoms. It has been recognised by the authors that several studies discuss the influence of unconventional development of oil and gas on the employees that work on-site that are expected to be greater in comparison to the experiences of the communities given the relatively close proximity of workers to burdens and possibly the increase in the levels of their exertion that involves air pollution.

It has been identified by Zahari et al. (2012) that the petroleum sector has been observed to be the most important sector in various countries that including Libya, The National Oil Corporation (NOC) of Libya established in 1970 and it has been recognised to be dominant on the oil industry of Libya, along with various smaller subsidiaries that are known to have a mutual interpretation for about half of the output of the oil of the country, furthermore, NOC has been identified to have carried out enhanced operations in relation to exploration and manufacture in collaboration with a variety of other companies under agreements for investment in petroleum, and the marketing of the oil and gas on a local and global level. The authors have added that NOC of Libya has been recognised to be a state-owned company and one of the major companies of Africa involved in the manufacture of oil, while the policy has been recognised to be associated with the promotion of the Libyan economy through the development, and management, as well as the exploration of oil sources, in addition to this, Libya has been identified to be the

member of Organisation of Petroleum Exporting Countries (OPEC) that hold the main established reserves for oil and the most important countries that are involved in oil manufacture in Africa and on a global level, it is also noteworthy that the petroleum sector in Libya has enhanced importance as it may be regarded as the backbone of the Libyan economy.

Methodology

Research Method

The method of research chosen for this study is qualitative that may be explained as the type of method for market research in the field of business that tends to focus on the achievement of information through open-ended and conversational communication that assist in the understanding of the perceptions of people and the reasons for their thinking (Question Pro, 2018). Qualitative research involves the collection and analysis of the non-numerical data that may provide assistance in the enhanced identification of concepts, opinions, or experiences, additionally, it can be applied for the gathering in-depth insights in relation to a problem or the generation of advanced ideas for research (Bhandari, 2020).

Data Collection Method

Data collection is the procedure associated with the collection, and measurement as well as the analysis of the accurate insights for research with the help of standard validated techniques that may contribute to the evaluation of the hypothesis based on collected data, it may be regarded as the primary and most important step regardless of the field of research, moreover, data collection approach may be recognised to be different for various fields of study that may be dependent on the required information (Question Pro, 2018). The data collection for this study is based on the use of various databases that include Google Scholar, PubMed and Science Direct whereas, the other sources may include books and government websites.

Data Analysis

The analysis of data achieved based on the research topic is based on the use of critical analysis of data that may assist in having an enhanced understanding of the experiences of people in terms of health and

environment as well as occupational influences of the petroleum industry in Libya. Critical analysis may be referred to as critical discourse analysis or critical discourse studies that may be regarded as an approach for the performance of the research that may provide assistance in the investigation of the relationship between language and power through the examination of the "texts" and the creation and the reinforcement of the social inequality and hierarchy as discussed by Allen (2017). The author has further added that the purpose of critical analysis involves the exposure and for the development of a challenge that may be taken for granted in relation to the structures of power and offering alternate perspectives related to knowledge, and the development of the theory, as well as social reality, it may also be regarded as notable that Critical analysis continues to gain significant importance in academic research as a result of its interdisciplinary appeal and application.

Findings and Discussion

According to Nabih et al. (2011) Unconventional oil manufacture may be observed to have an expansion as a result of the decline in the conventional oil manufacture after 2020, moreover, the richest portion associated with the major unconventional oil sources has been anticipated to grow with the escalation of prices that have been recognised to have a potential to allow profitable operations. The evidence suggests that there is a requirement for the understanding and re-evaluation of the heavy oil reservoirs in Egypt and Libya that is identified as a sign to ascertain the possibilities related to the existence of heavy oil reserves at the edge of the Sirte basin in Libya. The authors have added that the occurrence of Heavy oil has been observed to exist in shallow and deep structure geologic settings that have been recognised to be unconventional on the basis of the accepted standards, yet they may be regarded as prospects related to economic interest in association with the modern methods in relation to manufacture, in addition to this, the encouraging potentiality of oil that is the proven oil of AP1 gravity ranging from 11.4° to 20.5° AP1 in deposits of both carbonate and sand of Miocene and pre Miocene that has a potential to add more oil reserves for future phases of exploration phases, whereas, the high prices of oil contributes to have a second look at old unprofitable wells, with the help of innovative technology to perform enhanced minings from the reservoirs.

It has been identified by Biltayib (2006) that exploration and manufacture of petroleum have been recognised to have the possibility for the development of several influences on the environment that is dependent upon the level of process, and the extent as well as the project complexity, addition to the nature and sensitivity associated with the surrounding environment as well as the planning effectiveness, it may further involve the prevention of pollution prevention, and the mitigation as well as control techniques, however, the activities related to oil and gas covering great areas involve emissions and discharges related to pollutants in all phases that may range from the initial seismic surveys until shutting down of fields and the removal of installations. The author has recognised that during exploration, noise from the aircraft, and helicopters involved in the survey, as well as seismic explosions, have the potential to cause the escape of animals from the area, whereas, the seismic crew may have a contribution to erosion while inappropriate waste disposal from base camps can cause water and food supply contamination as well as the degradation of the environment and in case of improper handling during drilling and manufacture, discharge of unwanted waste and toxic substances may be threatening to the environment and communities. The author further believes it to be notable that Groundwater has enhanced sensitivity in relation to the contamination that leads to the development of intense influences related to the health of the wildlife and people of that area, whereas, the inappropriate disposal of formation water extracted with the oil from the well that is comprised of oil, and increased levels related to chlorides as well as heavy metals contaminate the ground and surface water leading to severe influences on people, and animals as well as vegetation, moreover, the improper controls related to decommissioning and rehabilitation can cause soil and water contamination.

Kaden and Rose (2015) have identified that the problems related to environment and health in Unconventional Oil and Gas Development in Libya promote convincing viewpoints from enhanced views in relation to fundamental problems having increased relevance to the application of directional drilling and hydraulic fracturing, and the provision of an appropriate demonstration in relation to necessary information associated with the technological application associated with unconventional oil and shale gas development. The authors have indicated that the utilisation of the proficiency associated with the range of contributors in academia, and

non-governmental organisations, as well as the oil and gas industry, the problems related to environment and health in Unconventional Oil and Gas Development, may be regarded as a vital source for academics and professionals associated with the oil and gas, in addition to the environment, and health as well as safety industries and for policymakers.

Occupational hazards include exposure of truck drivers to large amounts of silica (Daher, 2015) moreover the workers are at the risk of exposure to oil and gas from spills that may cause damage to the respiratory system, and liver, furthermore it also has the potential for causing decreased immunity, and an increased risk of cancer, additionally, it may also have an influence on the process of manufacture and may lead to the higher levels of some toxic substances that include hydrocarbons and heavy metals.

For the improvement in its practises Libya is recommended to perform voluntary initiatives that may be regarded as a significant tool, that range from international agreements and programmes to national policies, and legislation as well as regulation, and the financial sector that has been recognised to have an enhanced involvement in the fulfilment of the requirements associated with lending and investment that serve the purpose for the improvement of the practises related to the viable growth and the performance of industrial activities, moreover, their effectiveness is based on their capacity for reaching beyond the regulations of the government and for the commitment of the industry of its own free for the development of goals associated with the improvement of environmental performance as discussed by Biltayib (2006). It has been further suggested by the author that the MEA agreement between several parties to take certain steps may contribute to increases in protection of the natural sources of the world and the promotion of enhanced environmental quality, this agreement is based on the international and regional conventions and protocols that may be applied for the management and regulation of the environmental influences associated with the oil industry.

Abusa and Gibson (2013) have also suggested that the Libyan Government in its stated efforts are required to promote motivation among the companies for the implementation of the Total Quality Management (TQM) program that may contribute to the improvement of

their business performance that in turn, will promote improvement in the economy of Libya.

Conclusion

It has been concluded that unconventional oil manufacture expanded as a result of the decline in conventional oil manufacture after 2020, moreover, the richest portion associated with the major unconventional oil sources may grow with the escalation of prices that have the potential to allow profitable operations. The occurrence of Heavy oil exists in superficial and profound structures related to geological settings that are unconventional on the basis of the accepted standards, yet they may be regarded as prospects of economic interest related to modern methods of manufacture. The promising possibility of the proven oil associated with AP1 gravity ranges from 11.4° to 20.5° AP1 in the deposits of carbonate and sand of Miocene and pre Miocene that has a potential to add more oil reserves for future phases of exploration phases, whereas, the high prices of oil contributes to having a second look at old unprofitable wells, with the help of advance technology from enhanced minings from the reservoirs.

For the improvement in its practises Libya is recommended to perform voluntary initiatives that is a significant set of tools, that range from international agreements and programmes to national policies, and legislation as well as regulation, and the financial sector that have an enhanced involvement in the fulfilment of the requirements associated with lending and investment that are vital for the improvement of the practises related to the sustainable development and the performance of industrial activities, moreover, their effectiveness is based on their capacity for reaching beyond the regulations of the government and for the commitment of the industry of its own free for the development of goals associated with the improvement of environmental performance.

References

Abusa, F.M. and Gibson, P., 2013. TQM implementation in developing countries: a case study of the Libyan industrial sector. *Benchmarking: An International Journal,* 20 (5), pp. 693-711.

Allen, M. ed., 2017. *Critical Analysis.* The SAGE encyclopedia of communication research methods. Sage Publications.

APHA, 2018. The Environmental and Occupational Health Influences of Unconventional Oil and Gas Industry.

Bhandari, P., 2020. *What is Qualitative Research? | Methods & Examples.*Scribbr.

Biltayib, B.M., 2006. Oil manufacture in Libya using an ISO 14001 environmental management system.

Blake, U., 2016. *The Unconventional Oil and Gas Process, and an Introduction to Exposure Pathways.* Environmental and Health Problems in Unconventional Oil and Gas Development. Elsevier. pp. 1-12.

Daher, E., 2015. *An analysis of global safety trends in the oil and gas industry-influences and challenges in the years ahead.* SPE/IATMI Asia Pacific Oil & Gas Conference and Exhibition. OnePetro.

Hallett, D. and Clark-Lowes, D., 2017. *Chapter 1 - History of Oil and Gas Exploration.* Petroleum geology of Libya (Second Edition). Pp. 1-33, Elsevier.

Kaden, D. and Rose, T.L. eds., 2015. *Environmental and health problems in unconventional oil and gas development.* Elsevier.

Krupnick, A.J. and Echarte, I., 2017. *Health influences of unconventional oil and gas development.* Sources for the Future (RFF). June.

Nabih, N.A., Youssef, A. and Abd Raboh, W.M., 2011. *Petroleum Geology and Potential Hydrocarbon Plays Of Heavy Oil Reservoirs: Examples From Egypt And Libya.* Offshore Mediterranean Conference and Exhibition. OnePetro.

Question Pro, 2018. Data Collection: Definition, Methods, Example and Design.

Question Pro, 2018. Qualitative Research: Definition, Types, Methods and Examples.

Werner, A.K., Vink, S., Watt, K. and Jagals, P., 2015. Environmental health influences of unconventional natural gas development: a review of the current strength of evidence. *Science of the Total Environment,* 505, pp.1127-1141.

Zahari, I.B. and Shurbagi, A.M.A., 2012. The effect of organizational culture and the relationship between transformational leadership and job satisfaction in petroleum sector of Libya. *International Business Research,* 5(9), p.89.

Journal of Total Science
Volume (5), Issue (19((Mar. 2022)
ISSN: 2518-5799

مجلة العلوم الشاملة
المجلد (5)، العدد (19)، (مارس2022)
ردمد: 2518-5799

A study of modulation technique in LoRaWAN Technology

Mr. MAHMUD A. ALBRNAT

Department of Electrical Engineering, Faculty of Engineering, Sabratha University, Libya

Abstract

power consumption, coverage and cost of manufacturing are the most critical performance metrics in Wireless Network Sensors and IoT applications, LoRa® is a Wide Area Network protocol that provides a platform for sensor networks being suitable candidate for smart city applications, the technology is relying on chirp spread spectrum technique acquiring unlicensed spectrum with low power consumption allowing sensors powered with batteries for 10 years and more, this paper provides in-depth study of LoRa Network modulation and compression of different modulation scenarios, changing Spread Factor SF, Code Rate CR and bandwidth which trades between data rate and range of transmitted signals, also a compression of similar already deployed technologies such as Bluetooth, RPMA and SigFox which used as an alternative to WSN applications .

Introduction

The emerge of IoT devices and application demanded large amount of data communication from small devices connecting each other efficiently, mentioning efficiency in telecommunication means handling carefully the tradeoff relationship between power and range seeking higher data rates and error free transmission. Low power consumption circuits with wide range connection is the desired target in engineering communication devices especially for application that require less human encounter for example sensors in remote places.

LoRa (short for long range) is a spread spectrum modulation technique adopting chirp spread spectrum (CSS) technology. Semtech's LoRa is a long range, low power and low cost wireless network that has become a desired platform of Internet of Things (IoT) [1]. the technology based on spread spectrum which uses the entire spectrum bandwidth making it immune to channel noise, fading and Doppler shift effect, long range transceivers with low power consumption will yield to smaller data Rate, however, this will not be a severe problem for certain applications.

LoRa devices interconnect globally through LoRaWAN® which is a network protocol allowing LoRa nodes to communicate globally through cloud service deployed and maintained by the LoRa Alliance® which -so far- functioning in more than 177 countries worldwide[2] enabling small devices and sensors to be connected to the networks through gateways with ranges from 5 to 15 km and powered using batteries which can last 10 to 15 years with a low cost of manufacturing.

Low Power Wireless Technologies

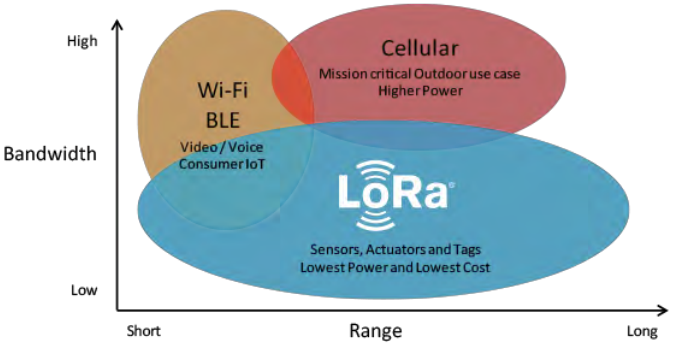


Figure 1 LoRa Range and Data rate compared to other Wireless technologies

There are several wireless technologies that utilizes license free ISM band with low power consumptions and provides a fairly acceptable data rates such as Bluetooth, ZigBee SigFox and Wi-Fi, wireless network sensors and IoT applications uses such networks as it provides the appropriate platform to work efficiently, however, these technologies do not provide the long range of connection which LoRa can provide, but they are still potential candidates to WSN and IoT applications.

LoRa Modulation

According to Shannon law for channel capacity. The capacity of transmitted data through a channel with the presence of noise is:

$$C = B \times log_2\left(1 + \frac{S}{n}\right)$$

Where:
C = channel capacity (bit/s)
B = channel bandwidth (Hz)
S = average received signal power (Watts)
N = average noise or interference power (Watts)
S/N = signal to noise ratio (SNR) expressed as a linear power ratio
From the above equation we can rearrange the $\log_2$into *e*which allows the equation to become:

$$\frac{C}{B} = 1.4333\ \frac{S}{N}$$

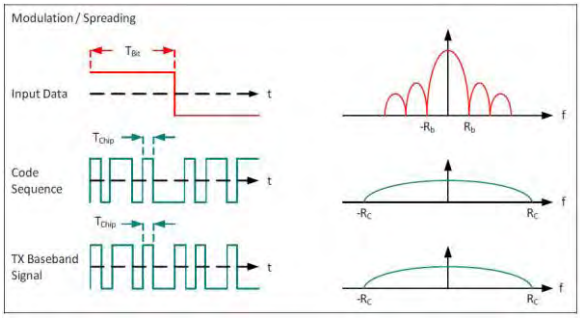
In spread spectrum modulation, signal to noise ratio is small as the signal through the channel is always below noise floor level. Assuming a noise level such that S/N << 1, Equation 2 can be re-written as:

$$\frac{C}{B} \approx \frac{S}{N}$$

Or in inverse situation:

$$\frac{B}{C} \approx \frac{N}{S}$$

From the above equation it's obvious that for a channel with fixed noise level and to perform an error free transmission you need to increase the bandwidth B.



Direct Sequence Spread Spectrum Modulation is simply multiplying the input data to code sequence which spreads the signal along the entire bandwidth as shown in the figure below:
At the receiver end the received signal it to be re-multiplied with the same code sequence to retrieve the original input data, however, synchronization and time offset issues between the transmitter and receiver is vital to avoid error, implementing a synchronization clock increases circuit complexity, power consumption and can add more cost.

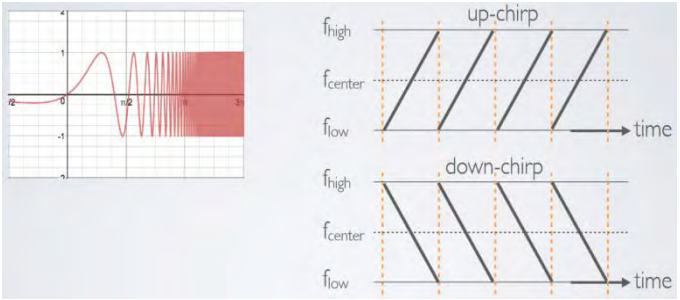


Figure 3: chirp frequency signals used in LoRa modulation

Chirp Spread Spectrum CSS

chirp spread spectrum (CSS) is a spread spectrum technique that uses wideband linear frequency modulated chirp pulses to encode information [2], chirps is a sinusoidal signal with increasing frequency from $F_{low}$to F_{high} (up-chirp) or decreasing (down-chirp). this technology was deployed in early 1940s in radar application, unlike other Spread Spectrum technologies DSSS and FHSS it doesn't require added pseudo-random elements to the signal to allow recognition of data at the receiver end, but rather relays on the linearity nature of the chirp signals, requiring less data transferred in data sequence, yet providing security and robustness against in-band jamming and channel degradation problems such as fading and multipath loss.

LoRa Modulation

LoRa technology adopts Chirp spread spectrum in modulation, it addresses traditional spread spectrum technologies issues providing a low-cost low-power alternative with the same signal robustness properties for long range signal transmission.

Varying chirp frequency signals achieve transmission almost equivalent in time offset between transmitter and receiver reducing the complexity of both transmitter and receiver design, the frequency bandwidth of chirps is equivalent the bandwidth of data signal [4].

data signal is chipped at a higher data rate and modulated onto the chirp signal.

modulation bit rate, R_b can be defined as:

$$R_b = SF \times \frac{1}{T_S} \quad Bit/sec$$

$T_S$is the symbol period and can defined as:

$$T_s = \frac{2^{SF}}{BW} \quad sec$$

SF = spreading factors (7 to 12)
BW = modulation band width (Hz)
Equation () can be rewritten as follows:

$$R_b = SF \times \frac{1}{[\frac{2^{SF}}{BW}]} \quad Bit/sec$$

Symbol rate $R_S$is the reciprocal of Symbol duration T_S

$$R_s = \frac{1}{T_S} = \frac{BW}{2^{SF}} \quad symbol/sec$$

Chip rate $R_C$is a direct relation of the spreading factor and defined:

$$R_C = R_S \times 2^{SF} \quad chips/sec$$

Mentioning that the chip rate is acquiring the whole bandwidth is defined as follows:

$$R_C = \frac{BW}{2^{SF}} \times 2^{SF} = BW \quad chips/sec$$

For the sake of safety transmission of bits an error correction code (parity check) is transmitted along data.

The coding rateCR refers to the proportion of transmitted bits that actually carry information. Coding rate can be 6/8, 4/8, CR of 4/8 means that every 8 bits transmitted 4 of them is carrying data information.
Introducing code rate into the equation we get:

$$R_b = SF \times \frac{code\ rate}{[\frac{2^{SF}}{BW}]} \quad Bit/sec$$

Where $code\ rate$= $[\frac{4}{4+CR}]$having a value from 1 to 4
From above, it's obvious that LoRa modulation is dependent on three important factors:

- Bandwidth BW
- Spreading factor SF
- Code rate CR

Bandwidth: LoRa uses three types of signal bandwidths 125,250,500 KHz and can adapt transmission between them, LoRa send modulated chirps at rate equivalent to the bandwidth used in chirps/sec/Hz so using 250 KHz will result 250 Kcps [6].

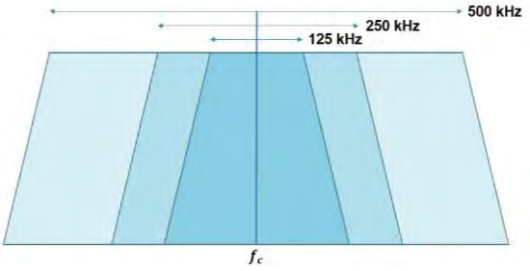


Figure 3: bandwidth of LoRa modulated signals

Spreading factor: LoRa uses spreading factors between 7 and 12 providing a tradeoff between range and data rate, every bit symbol is spread in 2^{SF} chips.

Spreading Factor (SF)	Chip Length 2^{SF}
7	128
8	256
9	512
10	1024
11	2048
12	4096

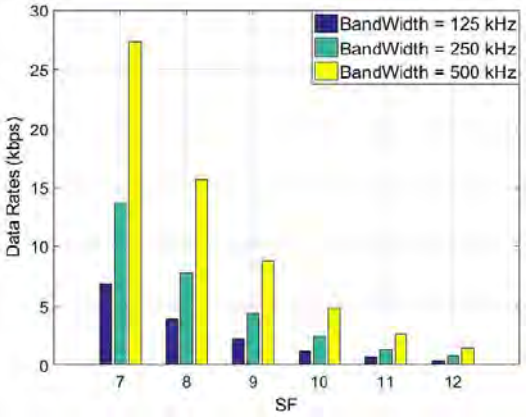


Figure 4:data Rate Vs Spreading Factors for different bandwidth

from this relation It's obvious that increasing spread factor will decrease the amount of data bits transferred, however, allowing one bit to conquer the whole bandwidth at a time increases the efficiency of transmission resulting in range extending, figure 4 shows the relation between spreading factor and data rate for CR=0.

Code Rate CR: to increase receiver sensitivity and safety of data transmission, a forward error correction FEC is carried after modulation, Code rate CR is a value expressing the amount of bits carrying information to transmitted bits, CR varies from 0 to 4, CR = 0 means no FEC is applied.

FEC can detect and sometimes correct errors, however, it consumes bits within the transmitted packet decreasing the effectiveness of data rate, figure 5 shows code rate used Vs data rate for different bandwidths.

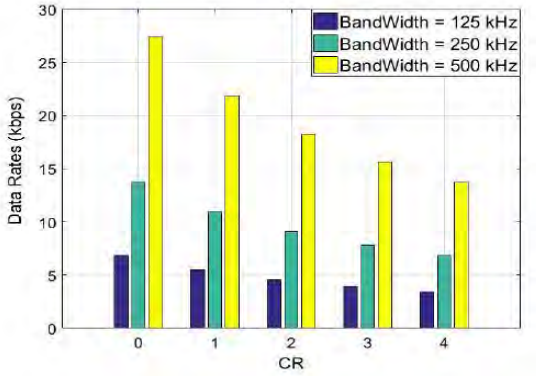


Figure 5: Data Rate Vs Code rate for different bandwidths

Packet structure:

Figure 6 shows Lora packet structure, LoRaWAN protocol allows 256 bytes as a maximum packet length, it is consisted of Permeable field, Header field, Payload, CRC[5].

Permeable field: is an eight up-chirp signals followed by two down-chirp signals and they are used to notify the start of transmission so the receiver will realize that the next beginning of data signals after 8 up 2 down message.

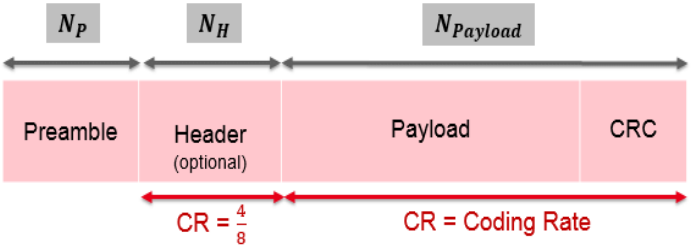


Figure 6: LoRa data packet [5]

Header Field: it is 4-byte field that contain CRC correction code, as mentioned in figure the header can be optional and will be dispensed for CR = 0

Payload: it is the data field which can holds from 0 to 255 bytes of data bits it contains:

- MAC header: holds information such as frame type (data or acknowledgment), protocol version and data direction (upstream or downstream).

- MAC payload: contains actual data.
- MIC: is a digital signature of the transmitted data.

CRCfield: is a 2-byte optional field containing cyclic redundancy check (CRC) bytes for error protection for payload

Time on Air

An important parameter in studying wireless networks is how long does the transmitted packet needs to reach receiver, this is can similarly expressed as latency in telecommunication systems which is the time required for the packet to reach destination and back, however, in LoRa networks, devices or sensor nodes usually only transmit in one-way direction, for this reason we would use the term *Time on Air.*

The *time on Air* for Lora packets is can be expressed as follows:

$$T_{packet} = T_{peramble} + T_{payload}$$

Where $T_{peramble}$is preamble duration and $T_{payload}$is payload duration. $T_{peramble}$ which descried as follows:

$$T_{peramble} = \left(\eta_{peramble} + 4.25\right) Ts$$

As mentioned above:

$$T_s = \frac{1}{R_S} = \frac{2^{SF}}{B.W}$$

From the above equations it is obvious that *Time on Air is* a direct relation of the spreading factor SF, the *Time on Air* for LoRa packets increases as packet length and Spreading factor increased,

figures 6 and 7 shows the relationship between spreading factor and Code Rate to the *Time on Air* for transmitted packets [6].

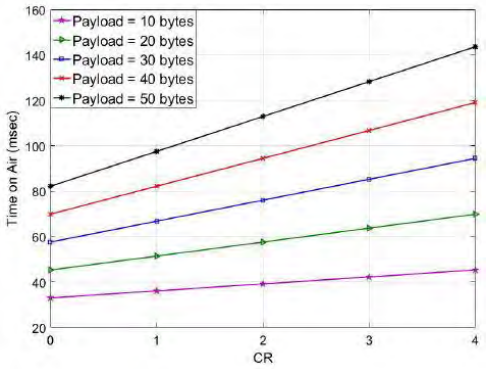


Figure 7:Time on Air Vs Code Rate for different payloads[6]

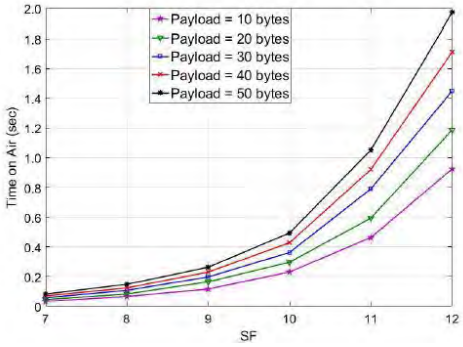


Figure 8: Time on Air Vs spreading factors for different payloads[6]

Conclusion:

Different LPWAN technologies are competing to supply the best service in the field of IoT applications that the world will demand in the future.

This paper focused on LoRaWAN, a robust low cost long range state-of-the-art technology that can provide the most prominent needs in wireless networks.

A compression of other wireless low power technologies was carried in this research.

An in-depth study of the modulation technique used in this technology, the overall performance was analyzed for different parameters such as Spreading factors and code rate presence, the effect of using different spreading factors to data rate and range, utilization of FEC codes and how using different bandwidths can effect time on air, transmission range and data rates.

References

1- https://www.semtech.com/lora/what-is-loraaccessed 10-1-2022
2- https://lora-alliance.org accessed 10-1-2022
3- IEEE Computer Society, (August 31, 2007). IEEE Standard 802.15.4a-2007. New York, NY: IEEE.https://ieeexplore.ieee.org/document/1091721
4- "LoRa Modulation Basics" (PDF). *Semtech*. Archived from the original (PDF) on 2019-07-18. Accessed 2022-01-10
5- TaoufikBouguera, Jean-François Diouris, Jean-Jacques Chaillout and RandaJaouadiEnergy consumption model for sensor nodes based on LoRa and LoRaWAN(2018)
6- Umber Noreen, Ahc`eneBounceur and Laurent Clavier, A Study of LoRa Low Power and Wide AreaNetwork Technology(2017)
7- https://iot4beginners.com/difference-between-lora-and-lorawan/ accessed 13-01-2022

Journal of Total Science
Volume (5), Issue (19((Mar. 2022)
ISSN: 2518-5799

مجلة العلوم الشاملة
المجلد (5)، العدد (19)، (مارس2022)
ردمد: 2518-5799

The problems That Face Students In Translating Phrasal Verbs

Mr. Nouraddeen Almagtouf Alwaheishi
Higher institute of science and technology Azzahra.
Mrs Aish Moulod Mansour
Higher Institute of science and technology Azzhra.
Mrs Mounira Hadia Mahfoud
Higher Institute of science and technology Al Azizia.
Mrs Ibtissam Al mabrouk Sahab Faculty of technical Engineering Janzour

Abstrac
This research focused on the misunderstanding and problems that faced students in translating phrasal verbs and finding the appropriate equivalent in the second language. The main tool for data collection was various examples that contained phrasal verbs with suggested phrasal verbs to know the way that obtained by students in translating phrasal verbs. This research also talked about the errors that committed by some students, these errors could be attributed to the lack of knowledge and depending on literal translation way.

الملخص

هدا البحث ركز على المشاكل التى تواجه الطلاب في ترجمة اشباه الجمل الفعلية وكيفيه اجاد المرادف المناسب في اللغة الثانية. حيث كانت الوسيلة الرئيسية لتجميع البيانات في تحليل مشاكل الترجمة عبارة عن أمثلة تحتوي على ترجمات مقترحة لمعرفة ماهى الطرق التى أعتمدها الطلاب في ترجمة الأخطاء التى وقع فيها بعض الطلاب نتيجة للفقر فالمعلومات ولاعتمادهم على الترجمة الحرفية.

1_ Introduction

Translation is an act through which the content of a text is transferred from the source language into the target language (Foster,1958). The translator needs to have a good knowledge of both languages, in addition to a high linguistic sensitivity as he should transmit the writers‟ intentions , original thoughts and opinions in the translated version as precisely and as faithfully as possible.

Translation is a process and a product. According to Catford (1965), “translation is the replacement of textual material in one language (SL)by equivalent textual material in another language(TL)”(p.20). This definition shows that translation is a process, in the sense that is an activity performed by people through time, when expressions are

translated into simpler ones in the same language (rewording and paraphrasing).

2_Issue of research

Most translators face problems in translating phrasal verbs, because phrasal verbs combine a verb with a preposition or adverb or both and that functions as a verb whose meaning is different from the combined meanings of the individual words. For example :

- You are welcome anytime you call in.
  The translator here unaware to the exact meaning for this phrasal verb , because he translated it literally as :تهاتفني
  While the exact meaning of this phrasal verb is :
  أهلا بك في أي وقت تزورني
- We fell out but we are friends.
  اختلفنا لكننا أصدقاء
  In this case, we have to avoid literal translation completely because phrasal verbs have different meanings can not be obtained by knowing the meaning of the simple verbs separated.

The translator used idiomatic translation, because he was aware of the structure phrasal verb and its context .

3_Aim of the research

Taking into consideration the importance of phrasal verbs, and recognizing the difficulties that may be posed while translating this type of figurative language, the main aims of this research are:

1- To find out the difficulties that facing the translators when translate the English phrasal verbs into Arabic.
2- To find out the strategies that should be followed to overcome these problems.

4_Hypothesis

The majority of English as well as translation students and translators have problems with translation, especially in translating phrasal verbs that constitute one of the major difficulties in translation, I think this is due to:

1- Misunderstanding of using phrasal verbs .
2- The lack of language competence .
3- Over depending on traditional way of translation(Direct Translation).

5_Methodology

This study consists of two parts; the first one is theoretical ,while the second is practical.

Part one provides a theoretical background about the phrasal verbs while Part two deals with comparative analytical study based on data collected from students translation of The Higher Institute Of The Science and Technology_ Al Zahra about five sentences consist phrasal verbs with their various standard translations, these standard translations were taken from the book " Fundamental Idioms in English " written by "Qahtan F El-Khatib".

6_Previous studies

Kolln,1998. Believe that „‟phrasal verbs define only those combinations that from an idiom, a phrase whose meaning cannot be predicated from the meaning of its parts‟‟‟.

1.1 Introduction

This chapter is concerned with phrasal verbs as being a very specific characteristic trait of the English language. It will be about the definition of the phrasal verbs, the importance of phrasal verbs and the problems of translating them from English into Arabic.

1.2 Definition of phrasal verbs

Placing prepositions and adverbs after some verbs in English, in order to obtain different and various meanings, is a very frequent trend in modern English, these verbs are called phrasal verbs (Thomson &Martinet,2009). According to Potter: phrasal verbs are made up of a verb and a particle .A particle can be an adverb (such as “out” or “away” for example “go out” , “put away” , or it can be a preposition (such as “with” or “form”)for example :”deal with” , “shrink form”.

Some phrasal verbs have two particles; both an adverb and a preposition for example :”get on with” ,”stand up for" (cited in McMillan, 2005:2).

Parrott(2000) defines multiword verbs as being “made of a verb (e.g. come, get, give, look, take) and one or more particles .

Particles are words that we use as adverbs and /or prepositions in other contexts(e.g. away, back, off, on, out)

(p.108). To illustrate his definition, Parrott has given the following examples:

1- I didn‟t come to until several hours after the operation.

2- I couldn‟t put up with the noise any longer.

According to some grammarians, multiword verb is the best name for this linguistic phenomenon to refer to all the forms phrasal verb can take (verb+ adverb+ preposition). These possible combinations are distinguished by two different names for some other grammarians . When the particle is a preposition, the combination is called a prepositional phrase. When the particle is an adverb and a preposition, the combination is called a phrasal prepositional phrase . The name multiword verbs seems the best name since it covers all these forms . Phrasal verbs are best described as a lexeme; that is a unit of meaning that may be greater than a word (Crystal,1995). They may be obtained from combining the same verb with different particles . Therefore, different meanings are obtained(e.g. break away, break in, break out). The same multiword verb may have more than one meaning (e.g. I broke the chapter down into small units, or the car broke down just outside London). Other verbs are combined only with one particle, and generally have only one meaning (e.g. log on).
(Parrott,2000). Jepersen,(1928)defines phrasal verbs as:

"(short) verbs are frequently used in connection with adverbs and preposition in a way that the meaning of the combination can be in no way deduced from the meaning of the meaning of each word separately, for instance: put in ,put off, put out, put up, make out, make for, make up, set down, set in, set out, set on, set up…."(Cited in fraser,1976:64).

This definition indicates the fact that the meaning of phrasal verbs cannot be obtained by knowing the meaning of the simple verbs separated from the meaning of the particles. The meaning of simple verbs differs from their meaning when they are related to any particle(Parrott,2000). As far as meaning is concerned, phrasal verbs can be classified into three types depending on how much change of meaning takes place with the particle(s).

1.3_ The importance of phrasal verbs.

Phrasal verbs are an interesting linguistic phenomenon. They were found in Middle English. They were also common in Shakespeare writings and often used to define verbs of Latin origins . In the eighteenth century, the famous lexicographer Samuel Johnson was among the first to consider these formations. According to Fraser(1976),they are: "one of the most characteristic traits of the English language""(P.64). phrasal verbs are a very specific linguistic feature in the English language; Moon stated: "very few languages have phrasal verbs like English"" (Cited in

McMillan, 2005:6) . The English language is characterized by the existence of numerous and various phrasal verbs.

Bolinger (1971) answered the question of why there are so many phrasal verbs in English, he said:

"they are words the everyday inventor is not required to reach for elements such as roots and affixes that have no reality for him , it takes only a rough familiarity with of other cases of *head* and *off* to make them available for *head off.* Virtually self suggestion when the occasion for them comes up which is not true for learnt notions like *intercept*""(P.xii).

This statement indicates that phrasal verbs are easy to be formed and invented. Besides , it has been noted that phrasal verbs are more expressive than the synonyms they replace. The following are contrasted; insult with to jump on, assault with to jump at .Insult and assault are less expressive than the two phrasal verbs they replace; whereas the two phrasal verbs :to jump on and to jump at are more expressive and stronger versions of the single word verbs they can be replaced by. Another aspect which is not often taken into consideration is the number of new nouns derived from phrasal verbs .They are possibly the source of most new nouns in English. Here are some examples of some nouns derived from phrasal verbs:

Runaway _ run away
Made-up _ made up
Layout _ lay out
Checkup _check up
Upkeep _ keep up
Drop out _ drop out

Phrasal verbs are important in the English language and for translators .Using phrasal verbs both in spoken and written language makes them sound natural and native alike; because they are used in natives both speech and writing.

They are more expressive than single word verbs as they are indisputably important for translators .They can be used in both situations; formal and informal. Fletcher says: "phrasal verbs are used across all types of texts, even where the writer or the speaker has the option of choosing a single word alternative.

Although phrasal verbs tend to enter the language through casual speech , in most cases, they progressively become accepted across a wide range of

texts until they reach even the most technical or conservative of text types""(Cited in McMillan, 2005:14).

7_Data Analysis

*7.1*_Introduction

The aim of this study is to examine the students' knowledge or skills when dealing with English idiomatic phrasal verbs. In order to analyze the students mistakes made in translating idiomatic phrasal verbs from English into Arabic. The following procedures were followed:

_ The original structure or items and their translations, the percentage of translating each structure or items and total number of mistakes made in each is presented in table.

_ The statistics are given in tabular formats.

7.2_Source of data

Mistakes in translating English idiomatic phrasal verbs into Arabic and the standard translations of mistakes are given. These standard translations were taken from the book " Fundamental Idioms in English " written by "Qahtan F El-Khatib".

7.3_Analysis of data:

-sentence one:

I whish I could think up a good excuse to give the teacher for not having prepared my homework.

The English sentence	The suggested translation	Total Number Occurrences	Percentage of Translation	Total Number of mistakes	Number of non-answered sentence
I whish I could think up a good excuse to give the teacher for not having prepared my homework	1_امل ان اجد 2_ اتمنـــى ان يكون لدي 3_ تمنيـت لـو اني فكرت 4_ ان اختلق	20 06 08 22	23% 15% 10% 30%	8 10 12 5	
Total		56	78	35	22

In this table ; the percentage of all participants was 78% and there are four translations of the phrasal verbs " *think up"*. Firstly was 23% for امل ان اجد. Secondly was 15% for لدي. Thirdly was 10% for تمنيت لو اني. Portly

was 30% for ان اختلق. While the standard Arabic translation of English phrasal verbs "think up "is اختلق. The mistakes that made by students were 35% because the lack of the background of the second language.

_ Sentence Two

I refuse to put up with his actions any longer.

The English Sentence	The Suggested Translations	Total Number of Occurrences	Percentage Translation	Total Number Mistakes	Number of non-answered sentence
I refuse to put up with his actions any longer.	ارفض التستر ارفض السماح لم ارد الاستمرار ارفض ان اتفق ارفض تصرفاته ارفـــــض ان اتحمل ارفض الرضوخ	02 06 01 03 02 09 02	13 16 17 08 30 04 03	02 06 01 03 02 00	
Total		25	89	14	11

All the participations were 89%. In this sentence was seven standard translations . The first one was 13% for ارفض التستر .The second one was 16% for ارفض السماح. The third one was 17% for لم ارد الاستمرار .The fourth one was 08% for ارفض ان اتفق. The fifth one was 30% for ارفض تصرفاته .The sixth one was 04% for ارفض ان اتحمل. and the last one was 03%for ارفض الرضوخ.

.

_Sentence Three

She is going to the airport to see him off

The English Sentence	The Suggested Translations	Total Number of Occurrences	Percentage of Translation	Total Number of Mistakes	Number of non-answered sentence
She is going to the airport to see him off	لرؤيته لتراه يغادر لرؤيـــة طائرتـــه تحط لرؤيته يهبط لرؤيته يرحل لوداعه	01 03 02 02 04 10	10 03 06 11 10 50	03 02 04 02 05 03	
Total		23	90	18	10

The participation was 90% . There were six suggested translations of the phrasal verbs " see off". The first one was 10% for لرؤيته, the second 03 for لتراه يغادر, the third 06% for لرؤية طائرته تحط , the fourth 11% for لرؤيته يهبط, the fifth 10% for لرؤيته يرحل, and the sixth 50% for لوداعه. The total number of mistakes was 18 while the number non-answered was 10.

-Sentence four

If you don't stand up for your rights, no one will do it for you.

The English sentence	The suggested Translations	Total Number of Occurrences	Percentage of Translation	Total Number of mistakes	Number of non-answered sentence
If you don't stand up for your rights, no one will do it for you	لم يحرص لم تطالب لم تحارب لم تنتظر لم تقف لم تدافع	02 05 01 01 02 14	08 20 04 04 08 56	02 05 01 01 02 00	
Total		25	100	11	00

All the participations 100% in this sentence. The first one was 08%, the second 20%, the third 04%, the fourth 04%, the fifth 08% and the sixth 56%. The total number of mistakes was 11 with a percentage of 44%.

-Sentence Five

we told him not to play tricks on anybody, but he never listened.

The English sentence	The suggested translations	Total number of occurrences	Percentage of translation	Total number of mistakes	Number of non-answered sentence
we told him not to play tricks on anybody, but he never listened	ان يصنع مقالب ان يمازح ان لا يخدع بعدم التلاعب مع بعدم اللعب مع ان لا يمارس خدعة سحرية ان لا يسخر ان لا يعبث	03 04 10 05 01 03 05 06	05 10 15 02 09 10 14 22	02 03 19 05 06 13 04 03	
Total		37	87	55	05

The first translation was 05% for ان يصنع مقالب , the second 10% for ان يمازح, the third 15% for ان لا يخدع , the fourth 02% for بعدم التلاعب , the fifth 09% for بعدم اللعب مع , the sixth 09% for ان لا يمارس خدعة سحريه , the seventh 10% for ان لا يسخر and the eight 20% for ان لا يعبث. The total number of mistakes was 55 and the number of non-answered sentence was 05%

8_Conclusion:

This study is focused on the students problems in translating English phrasal verbs into Arabic. These problems and errors happened for three reasons, firstly; the students should have a great information of both languages especially English language. Secondly; students should avoid literal translation way in rendering the meaning because phrasal verbs are related to the culture more than the meaning itself.

Thirdly and finally; back ground and broad knowledge are the main tool in transferring the exact meaning.

9_ References:

_ Martha Kolln, Rhetorical Grammar: Grammatical Choices, Rhetorical Effects . Publisher: Boston : Allyn and Bacon, 1998.

_Bolinger,D.(1971). The phrasal verb in English . Cambridge, MA Harvard University Press.

_ Foster, M. (1958). Translation : Mission to proxy- Elixir International Journal. Translation From/ into Farsi and English.

_ Martin Parrott. (2000). Grammar for Englih Language Teachers. Cambridge University Press.

_ Fraser, P. (1979). The Interpretation of Novel Metaphors. In A. Ortony et. Al. (Eds). Metaphor and Thought (2nd Ed). Cambridge University Press.

_ Catford,(1965). A linguistic Theory of Translation. Oxford University Press.

_ Thomson and Martinet, (2009). A practical English Grammar. Oxford University Press.

_ Otto Jespersen,(1928). Introduction, AN International Language .

_ Crystal,(1995). Glossary of Linguistic Terms-Sil International. Cambridge University Press.

_ Mcmillan. J. H. Schumacher. (2005) Investigation Educative a 5 Ed. Virginia Commonwealth University.

Journal of Total Science
Volume (5), Issue (19((Mar. 2022)
ISSN: 2518-5799

مجلة العلوم الشاملة
المجلد (5)، العدد (19)، (مارس2022)
ردمد: 2518-5799

Cadmium Telluride (CdTe) Compound and Silicon (Si) Semiconductor Materials Detectors for X-Ray

Mrs. Karima Mohamed Abdulla
Mrs Assma Musbah Said
Mrs Samah Aghtisa

University of Zawia.

Abstrac

There are different approaches and also materials to construct a detector for x-rays. the available semiconductor materials, Silicon (Si) is mainly used for charged particle detectors and soft X-ray detectors. Cadmium Telluride (CdTe) compound semiconductors for x-ray detectors have experienced a rather rapid development in the last few years, due to their appealing performance. In this paper we review the physical properties of semiconductor detectors for x-ray. In particular, we focus on compound semiconductor detectors. In addition, Si detector has better energy resolution achievement. However, it does not have good detection efficiency for higher energy above (40Kev). While CdTe detector supplies higher detection efficiency for x -rays about 100% for up to 100 keV in energy. The response functions of Si detector to reduce small x-ray, which has limited sensitivity for energy about 30keV. The XR-100T is high performance for x-ray and gamma ray detector by using ^{241}Am (59.5keV), so the cooler are mounted the input FET a feedback components to the Amptek to charge sensitive preamplifier. Both the Si and CdTe detectors are connected to the PX2 combined amplifier and power supply, that was provide the voltage needed to operate XR-100T CdTe. The PX2 T CdTe amplifier has a Rise Time Discrimination circuit, measure the FWHM energy resolution of ^{55}Fe (5.89) the peak in each spectrum hence a graph of energy resolution. Moreover the Rise Time Discrimination So the escape peak of ^{55}Fe was determined with Si detector was 11.66 keV, also k- sell x-ray energy was calculated to 5.77keV and the Rise Time Discrimination effect was studied on the spectrum when is it switched on and off. Its influence was found in decreasing the noise of the spectrum.

Keywords: Cadmium Telluride, Silicon (Si), materials, x-ray, radiation detector, semiconductor, cannel number, Energy.

Introduction

Silicon (Si) is a chemical element, and its atomic number 14. It is a hard, brittle crystalline solid with a blue-grey metallic luster, and is a tetravalent metalloid and semiconductor. It is a member of group 14 in the periodic table; carbon is above it; and germanium, tin, lead, and flerovium are below it. It is relatively unreactive. Because of its high chemical affinity for oxygen. Cadmium telluride (CdTe) is a stable crystalline compound formed from cadmium and tellurium [1].

Semiconductor detectors, originally developed for particle materials applications, are now widely used for x-ray spectroscopy in a large variety of fields, as x-ray fluorescence analysis, x-ray astronomy, where the applications as focal plane detectors and diagnostic medicine are of particular interest. semiconductor detectors have a several unique properties such as high detection efficiency, excellent energy resolution, and possibility of development of compact and tough detection systems, silicon (Si) detectors are the key detectors in the soft x-ray band (< 15 keV). In addition, among the compound semiconductors, cadmium telluride (CdTe) and cadmium zinc telluride (CdZnTe) have been considered very appealing for hard x-ray detectors and are now widely used for the development of spectrometer prototypes for medical and astrophysical applications [2].

X-rays are a form of electromagnetic radiation similar to radio waves, microwaves, visible light and gamma rays. X-rays are similar to gamma rays; the main difference is the producing way, X-rays are produced by electrons external to the nucleus. Traditionally X-rays had longer-wavelengths and lower energy than gamma rays but this is obsolete with modern X-ray production methods.

X-ray photons are highly energetic and have enough energy to break up molecules and hence damage living cells. When X-rays hit a material some are absorbed and others pass through. Generally, the higher the energy the more X-rays will pass through. It is this penetrating power that allows us to take internal images of the human body or objects. X-rays cannot be steered by electric and magnetic fields like alphas, betas or other charged particles [3].

The electromagnetic radiation includes x-rays and gamma rays, x-rays released in the rearrangement of electron shells of atomic that were produced by the deceleration of an electron radiate especially when it hits the target of high atomic number. The most important impact of the x-rays or gamma rays is very low energy so energy resolution for low energy gamma rays and x-rays irradiation on the cathode side is best.

There are many semiconductor detector which are using in research where an accurate measurement effect which is predominant mode of - interaction for gamma ray and x- ray of relatively low energy which also enhanced for absorber materials of high atomic number Z

silicon detector and cadmium telluride detector-each of which has advantages in certain application-were used for detection of x-rays [4,5].

Semiconductor detectors are the main substitute for the scintillation detectors in medical imaging systems. Silicon and germanium materials were old semiconductors studied by Van Heerden in 1945 as radiation detectors. These detectors, have a lower carrier creation energy and low scatter count. Unfortunately, Si with a low stopping power for high-energy photons is restricted to low-energy photons. Besides, Ge works at cryogenic temperatures because of its small bandgap. For these reasons, materials with a high atomic number and ability to function at room temperature such as CdTe are probable materials for using in X and gamma rays detectors.

XR- 100T Silicon detector Si detector with pettier cooling, have come into common use, delivering a resolution of 150 ev FWHM at 5.9 kev.

The Si detector combines a low atomic number with bandgap energy (1.12ev) for Si (300k) and 1.16 for Si (77k) also, the ionization energy is 3.61 ev per e-h pair for Si (300k) and 3.76 ev per e-h pair for Si (77k). Good energy resolution requires a relatively small bandgap, so, to excite an electron-hole pair it most reducing the average ionization energy. The number of charge carries excited per unit deposited energy is increased by using small ionization energy. Si detector has a high efficiency, is close to 100% for charged particles, but so is gas counters. For gamma rays and x-rays, efficiency is very poor because of the small sensitive volumes and the low Z number [5,7]. Figure 2 shows Cadmium telluride

(CdTe) detectors.

Figure 1 Si X-ray detector-XR- 100T Silicon detector [7]

Cadmium telluride (CdTe) detectors

The high atomic numbers of CdTe (48for Cd and 52 for Te) with a large band gap energy (1.52eV), and high density (5.85 g/cm3) with a high resistance (109 Ω) provide a better absorption characteristics [6], and ionization energy (4.43ev per eh pair). The probability of photon electric absorption per unit path length is 100 times larger than in Si for typical gamma-ray energies. The response function of a CdTe detector has some important difference from that of Si detectors and these differences must be understood and quantified to achieve accurate quantitative x-ray analysis results. Since Si detectors have been the detectors of choice for quantitative x-ray spectroscopy but they have limited sensitivity at energies above 30 Kev [4,7].
Energy resolution achievable in CdTe detector is generally not comparable with that obtainable in Si or Ge, because of rather poor collection efficiency of holes, Figure 2 shows Cadmium telluride (CdTe) detectors.
. Application of this material therefore most involve situation in which high gamma rays detection efficiency per unit volume is at a premium. The most CdTe detectors is detectors are relatively small and usually planar in design, as a result of the charge carrier transport difficulties [7,8].

Figure 2 Cadmium telluride (CdTe) detectors [7]

Electronic pulse rejection and correction techniques have been successfully applied to CdTe detectors to improve energy resolution. it can be applied effectively to simple pulse counting if spectroscopic information is not required and can also be operated in current mode in high gamma ray fluxes which is possible to temperatures up to 70 ^{o}c [7].

Theory

The intrinsic detection efficiency of a detector (εINT) generated by preamplifier, so preamplifier noise must also be minimised if the excellent resolution is to be realised in practice. Also being charge-sensitive, they must be low-noise devices and usually, based on field effect transistor (FET), So [5,7];

$$\varepsilon INT = 1\text{-}I(x)I_0 = 1\text{-}e^{-\mu x} \qquad (1)$$

Where I(x) is the intensity of photons, (x) a thickness of detector, (I_0) is the intensity initially incident on the detector, ,μ is the linear attenuation coefficient. So it can found the detection efficiency as function of the photon energy E:

$$\varepsilon INTE = 1\text{-}\ e^{-\mu E\chi} \qquad (2)$$

The energy resolution FWHM of Si detector for higher photon energies evaluates using expression:

$$R(E) = [R(5.9keV)^2\text{-}120^2 + 2440E]^{1/2} \qquad (3)$$

Where is R (5.9keV) the measured FWHM resolution at 5.9keV.

The energy resolution can be achieved with silicon detectors is almost as good as that obtainable with CdTe as shown in Figure 3.

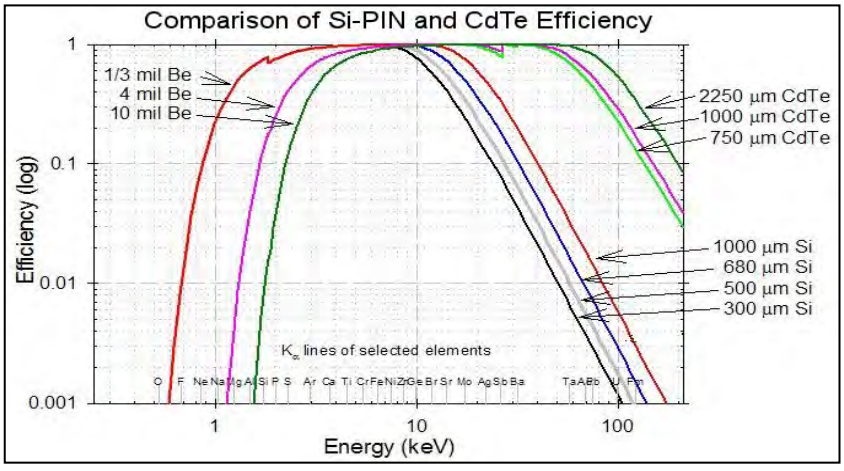


Figure 3 comparison of Si-PIN and CdTe Efficiency [7].

Silicon detector has good energy resolution performance, due to the c combination of a high quality Si detector. For the cadmium detector a higher bias voltage of 400v is required to ensure complete electron /hole collection.
Electron /hole Paris created by radiation which interact with the CdTe detector result in fluctuations of change collection time.

Experimental Methods

This experiment was 3 parts as following:

Part 1

It was used Si detector, which has a good resolution to measure x-ray spectrum. The following equipment was connected as in Figure

- XR-100T silicon detector;
- PX2 amplifier
- power supply and a separate shaping amplifier.

The output of the detector preamplifier was to the NIM mounted shaping time and the UNIPLR output of the shaping amplifier was connected to the MCA INPUT [7,8].

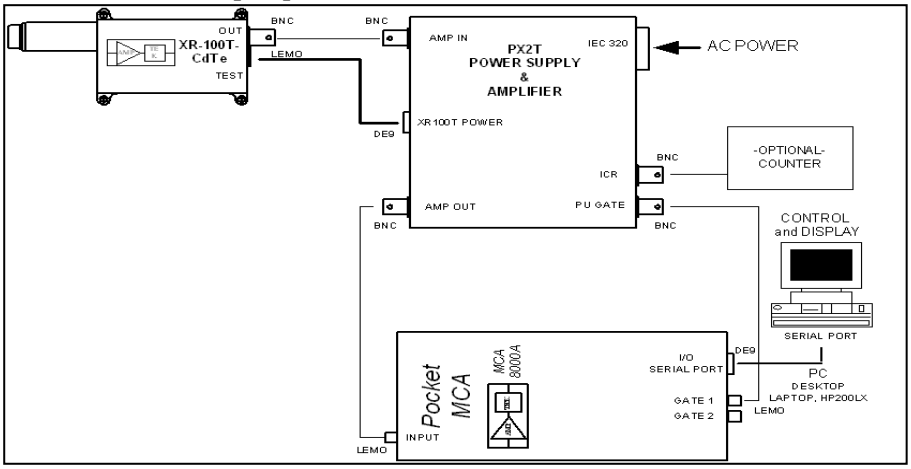


Figure 4 XR100T-CdTe/PX2T-CdTe/MCA800A [7]

Energy calibration of the silicon detector

- By using 241Am (59.5) kev source to produce x-ray and gamma ray peaks. The following steps done;
- value were set on shaping amplifier; coarse gain=30x, fine gain=1, shaping time=1μs and input polarity was negative.
- When acquired spectrum was appear, the value of fine gain was adjusted and changed to a 100x to get best spectrum.
- A copy of a ^{142}Am was printed and all peaks were identified.
- The energy calibration of the spectrum was done by using maestro MCA software and the information of each peak was recorded.
- A graph of energy (kev) vs channel number was plotted.

Peak intensities

By using Si detector to produce x-rays and gamma rays, to get good statistics from the ^{241} Am spectrum for longer time (10 minutes) was gained by using ROI as before. The peak intensities (net area) and peak energy of each ROI was recorded. The data will be used in the part 3.

Energy resolution of ^{55}Fe:

This part of experiment was carried out by using the low energy ^{55}Fe (5.89)kev. Check the MCA lower level discriminator low energy. The red plastic cap from the front the detector was removed. A new spectrum was acquired by placing ^{55}Fe source close to the detector. A series of the spectrum of the ^{55} Fe photon peak was obtained to acquire at different amplifier shaping times, in the range 0.5μs to 12 μs. The FWHM energy resolution of the mean peak spectrum was measured. The preceding part was used to record the centre and the FWHM of each energy peak to calculate energy resolution. Then the graph of energy resolution vs shaping time was plotted.
Finally, the shaping time was set to the maximum value (12μs) to acquire ^{55}Fe spectrum for several minutes. The experimental results show weak silicon escape peak below the (5.89) keV photon peak. The k-shell x-ray energy have been calculated in silicon, the result were compared with a data book value.

Part 2

The second part of this experiment was carried out by using XR-100Cd detector, whit PX2T-CdTe amplifier and power supply.

spectroscopy performance of the CdTe detector:

This section of experiment the internal shaping amplifier within PX2T-CdTe amplifier were used with a shaping time constant of 3 μs .the following settings on the PX2T-CdTe amplifier was checked:

- Gain is 2 and the Rise Time Discrimination (RTD) was active.
- 142Am source was used to get high spectrum of 142Am (59.5kev) in the middle of the spectrum.
- The spectrum of a142Am was printed and identified to compare with spectrum which was obtained from Si detector, than the calibration of the spectrum was done and the information of each peak was recorded.
- A plot of graph of the energy (kev) vs channel number was made to compare with the linearity and gain of silicon which was made in part 1.

peak intensities

for longer time (10munites) was get good statistics from 241Am spectrum by using the ROI as before, the peak intensities (net area) and peak energy of each ROI was recorded,(not error in the net area). and energy of each ROI were recorded to be used at in the part 3.

Rise Time Discrimination of the CdTe detector

- The ^{142}Am spectrum was acquired with the Rise Time Discrimination RTD off (non-active);
- A copy of this spectrum was printed to compare with previous spectrum acquired when the RTD was active.

Part 3

In final part of this experiment:

The peak intensity which was recorded previously from the Si and CdTe detectors, was used for plotting graph of Si detector efficiency vs photon energy, this was after calculate ratio of the Si/CdTe peak intensities for each peak, and plot this ratio as a function of peak energy.

Result and discussion

Energy calibration of ^{241}Am from Si and CdTe detector

This section of experiment is divided into parts, where the Si detector and CdTe detector with ^{241}Am (59.5keV) source was used which cause X-ray fluorescence that will larger absorption length than electron and will escape from the surface of the detector, the result will be gamma ray photon and escape peak as shown in Table 1&2.

Part 1

From the Si detector the x-ray spectrum energies were calibrated, all mean peaks were measured and the cannel number was determined from peak identification, Table 1.

Figure 4 shows the result obtained from x-ray spectroscopy

Table 1 channel number and peak energy for ^{241}Am determined by MAC

Energy calibration	Peak energy	Cannel number
^{241}Am	59.5	2069.93
^{241}Am	26.5	923.39
Np-Lx	20.7	734.07
Np-LX	17.7	614.65
Np-LX	13.9	488.00

Figure 4 Energy of Si detector vs channel number

part 2

The spectrum of source ^{241}Am (59.5keV) was done by using CdTe detector. The x ray spectrum was collected, the energy was calibre and the data was recorded. Table 2 shows the relationship between the cannel number and energy (keV).

Table 2 shows the relationship between the cannel number and energy (keV).

Peak identification	Peak energy(keV)	Cannel number
Am241	59.59	3399
Am241	26.3	1514.75
Np-Lx	20.7	1204.85
Np-Lx	17.7	1025.85
Np-Lx	13.9	810.72

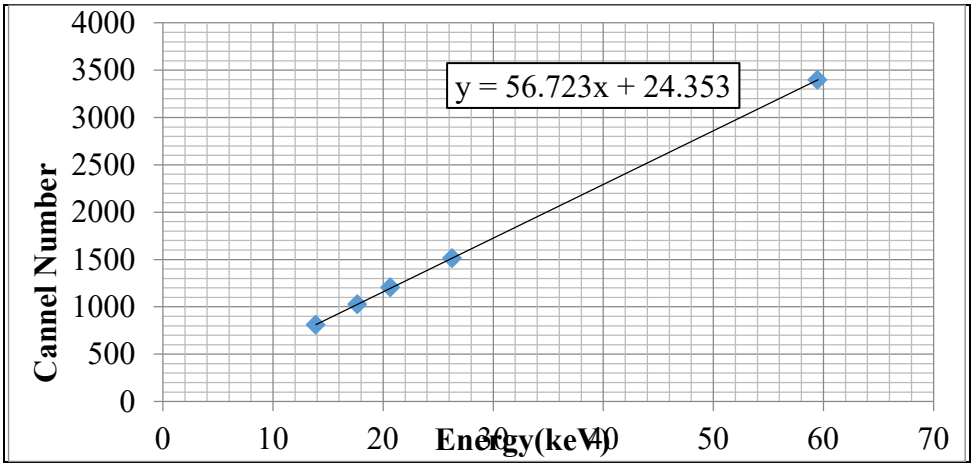


Figure 5 Energy (keV) of CdTe detector vs channel number.

Figures 4-5 show the relationship between the cannel number and energy (keV). From the Si and CdTe detectors respectively as it can been seen from the figures that the plotted graphs reveal a linear relationship between the cannel number and energy (keV) , which can be presented as y=ax+b, where y axis represent the channel number and x represent energy(keV).

Peak intensities and efficiencies

It was observed during the experiment that the ^{241}Am spectrum for a long time was acquired to get good statistics by using the same ROI from the Si and CdTe detector. The peak intensities (net area) and peak energy of each ROI was recorded. Table 3,4 show the relationship between the peak intensities (net area) and peak energy of ^{241}Am by using Si detector and CdTe detector.

Table 3 peak intensities and peak energy for Si detector

Peak identification	Peak intensities	Peak energy
Am241	2667±131	59,56
Am241	2137±168	26.17
Np-Lx	5359±289	20.86

Np-Lx	38332±608	17.53
Np-Lx	36445±658	14.00

Table 4 peak intensities and peak energy for CdTe detector

Peak identification	Peak intensities	Peak energy
Am241	66497±471	59.5
Am241	15264±318	26.25
Np-Lx	35008±532	20.82
Np-Lx	136941±861	17.68
Np-Lx	58944±547	13.9

Efficiency was calculated by using this relative:
Relative efficiency = peak intensities (net area) of Si / peak intensities (net area) of CdTe.
Figure 6 shows the ratio of Si/CdTe peak intensities agents the energy of each peak. it can seen clearly the high efficiency of CdTe detector; which was 100% over full energy range, so the Si detector was not as good efficiency as CdTt detector.

Table 5 Ratio of peak intensities of Si to the peak intensities of CdTe intensities peak vs Energy peak

Energy	peak intensities of Si (1)	peak intensities of CdTe (1)	Relative efficiency (Si/ CdTe) (1)	peak intensities of Si (2)	peak intensities of CdTe (2)	Relative efficiency (Si/ CdTe) (2)
59.5	2667+131	66497+471	0.04178	2667-131	6649-471	0.037868
26.3	2137+168	15264+318	0.14792	2137-168	15264-318	0.131741
20.7	5359+289	35008+531	0.15892	5359-289	35008-531	0.14705
17.7	38332+608	136941+861	0.28257	38332-608	136941-861	0.277219
13.9	36445+658	58944+547	0.62367	36445-658	58944-547	0.61282

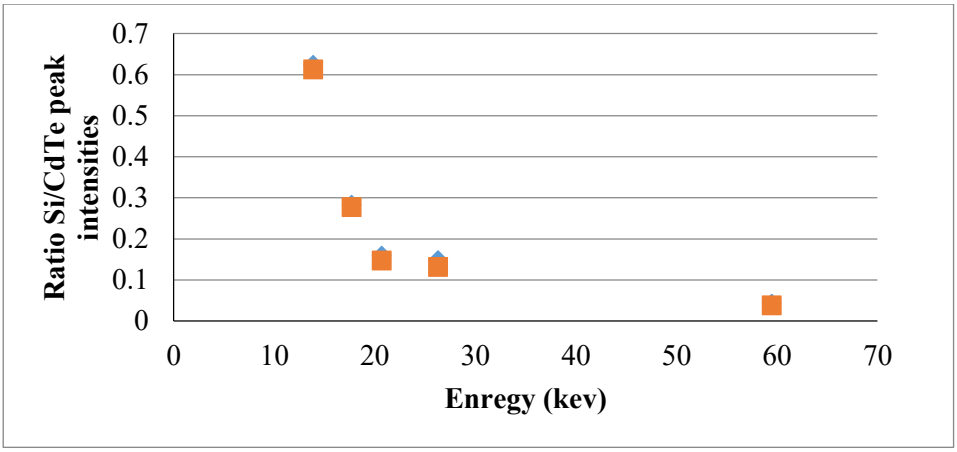


Figure 6 Ratio Si/ CdTe intensities peak vs Energy peak.

Energy resolution of ^{55}Fe

Table 5 represent the energy resolution of law energy ^{55}Fe (5.89keV) source, in the range of 0.5 µs to 12 µs by using Si detector. As calculated from equation 4:

$$R(E) = [R(5.9KEv)^2 - 120^2 + {}^{244}E]^{1/2} \quad (4)$$

Table 6 Energy resolution vs shaping time for ^{55}Fe

Shaping Time	Shaping Time
0.5	13.92
1	11.70
2	11.35
4	8.89
8	9.26
12	1

Figure 5 shows the relationship between energy resolution and shaping time, which was a clear in series spectrums were obtained with difference shaping time of (0.5,1,2,4,8,12)µs. The graph shows a good energy resolution of Si detector was with a long shaping time of amplifier due to reducing noise as it was increasing in shaping time.

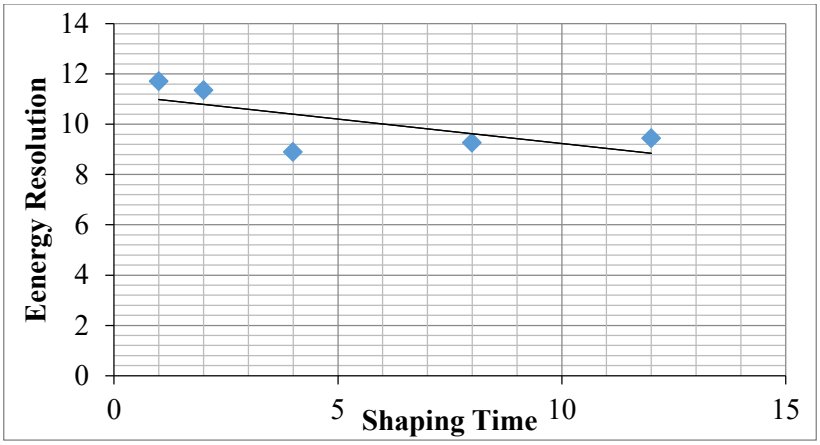


Figure 7 Show the shaping Time vs Energy Resolution

Rise time discrimination of PX2T-CdTe amplifier

In this part of experiment with the CdTe detector by using ^{241}Am when the (RTD) off (not active), the PX2T-CdTe amplifier had rise time discrimination (RTD) to refuse pulses that had a scrawny shape with lower energy resolution. However, when the (RTD) was active, the shaping pulses had clearly shape were internally gated and only pulses corresponding to good x-ray events were allowed to be sent to (MAC) for analysis.

Figure 7 shows the spectrum of ^{241}Am from CdTe detector with the (RTD) was off.

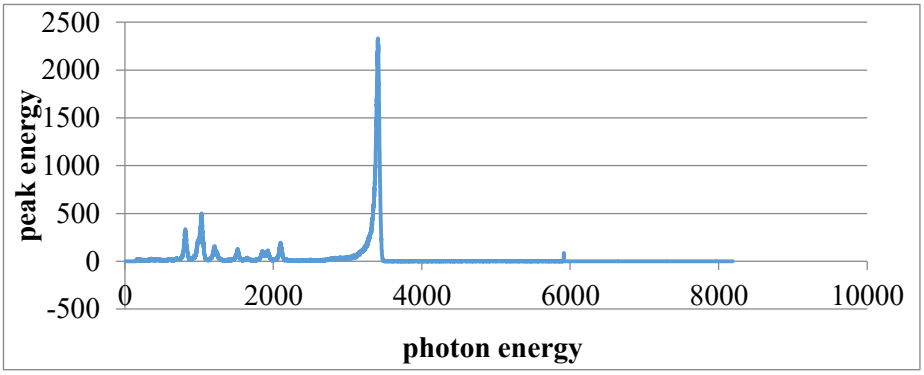


Figure 7 Function of peak energy vs photon energy.

Conclusion

The general features of x-rays and gamma rays spectra and the performance characteristic of have been investigated. The collected spectra of x-rays and gamma rays present. The Si detector and CdTe detector have advantages in certain application. CdTe detector has high atomic number to translate into significantly higher x-rays and gamma rays photon absorption has much higher than Si detector, also the detection efficiency per unit thickness has better than Si detector for x-rays with stopping power 100% efficiency up 50kV and 5% at 100kV and good operate at shorter shaping time, so the CdTe detector can be improved by cooling to reduce the leakage current is high enough to improve the energy resolution. While Si detector must be cooled in low noise amplifier, also the Si detector has better energy resolution than the CdTe detector lowest 25kev, for Si detector energy resolution of photon approximately 40keV; also Si detector has good spectrum characteristics and peak background. The PX2T-CdTe amplifier has a rise time discrimination circuit was practical to the linear amplifier to reduce tailing effect even further.

References

[1] J. R. De Laeter, J. K. Bohlke, P. De Bievre, H. Hidaka, H. S. Peiser, K. J. R. Rosman, and P. D. P. Taylor., 2000, "Atomic weights of the elements", International Union of Pure and Applied Chemistry. pp 683-800.
[2] B. Anders., 2014, " Comprehensive Biomedical Physics", 1st Edition, Elsevier.
[3] L. Carl-Magnus, H. Gillian., "X-rays" , The Australian Radiation Protection and Nuclear Safety Agency. Online on: https://www.arpansa.gov.au/understanding-radiation/what-is-radiation/ionising-radiation/x-ray. [15-01-2022]
[4] Knoll, G.f, 2000, "Radiation Detection and Measurement", Third Edition,(2000), John Wiley &sons.
[5] S. Jacob, 2002, "Radiation Protection", Third Edition.
[6] A. Samira, M. Babak, and P. I. Jalil., 2017, "Cadmium Telluride Semiconductor Detector for Improved Spatial and Energy Resolution Radioisotopic Imaging", World J Nucl Med.

[7] AMETEK material analysis division., Online on: https://www.amptek.com/internal-products/obsolete-products/px2t-cdte-power-supply-shaping-amplifier-for-the-xr100t-cdte. [02-02-2022].
[8] AMETEK material analysis division., Online on: file:///C:/Users/asma/Downloads/ Application-Note-AN34-Experiments-Nuclear-Science-Experiment-3.pdf. [23
-02-2022].

Journal of Total Science
Volume (5), Issue (19((Mar. 2022)
ISSN: 2518-5799

مجلة العلوم الشاملة
المجلد (5)، العدد (19)، (مارس2022)
ردمد: 2518-5799

A study on the impact of online flight reservations among Libyan companies

Mr. ABDALLA F.A BELIIAGI

Higher institute for sciences and technology – regdaleen

دراسة عن أثر حجوزات الطيران عبر الإنترنت في شركات الطيران الليبية

الملخص

لقد كان نمو استخدام تكنولوجيا المعلومات هائلاً في عالم الأعمال اليوم ولا شك أنه تم استخدامه على نطاق واسع في مجال صناعة الطيران. يعد نظام التذاكر الإلكترونية أحد الأمثلة على مثل هذا التطبيق للتكنولوجيا. هذه الدراسة عبارة عن بحث يركز على المستخدم ويهدف إلى تحديد العوامل التي تحد من قبول المستخدم لنظام التذاكر الإلكترونية. من خلال تكييف نموذج قبول التكنولوجيا (TAM) باعتباره عملًا فرعيًا نظريًا ، تفحص هذه الدراسة المتغيرات المدركة للفائدة وسهولة الاستخدام المتصورة كمفهوم لسلوك نية استخدام التكنولوجيا. علاوة على ذلك ، تبحث هذه الدراسة في تأثير الكفاءة الذاتية للكمبيوتر كمحدد تجاه الفائدة المتصورة وسهولة الاستخدام المتصورة.

من المتوقع أن تظهر النتائج التأثير الكبير للاكتفاء الذاتي للكمبيوتر في شرح نية التصرف على الرغم من سهولة الاستخدام المتصورة والفائدة المتصورة.

ABSTRACT

The growth of the IT usage has been tremendous in the world of business today and no doubt it has been widely used in the area of airline industry. The e-ticketing system is one of the examples of such application of the technology. This study is a user-focused research that aims to identify the factors that detemine acceptance of e-ticketing system by user. Adapting the technology acceptance model (TAM) as a theoretical franework, this study examines the variables perceived usefulness and perceived ease of use as antecedents of behavior intention to use the technology. Furthermore this study examines the effect of computer self-efficacy as determinant towards perceived usefulness and perceived ease of use.

Finding will be expected to show the significant effect of computer self-efficacy in explaining behavior intention though perceived ease of use and perceived usefulness.

Introduction

Due to the important role affixation plays in the process of vocabulary learning, special attention should be paid to the treatment of this linguistic phenomenon. According to Nation (2001), treatment lies in explicit presentation. In particular, we need an affective vocabulary learning strategy by using word parts to help learners remember the meaning of a word. This strategy requires learners learners to know the most frequent and regular affixes well, to be able to recognize them in words; and to be able to re-express the meaning of the word using the meanings of its word parts.

The field of psycholinguistics (which actually has very close ties with the area of language learning and processing) has given us an important concept related to explicit language learning: The more one manipulates, thinks about and uses mental information, the more likely it is that one will retain that information (Rechards,2000). In the case of affixation Richards (2000: 65) asserts that:

" affixes defiantly worth explicit attention from the teacher, because using word parts is one of three major strategies that can help students become independent vocabulary learners . This explicit learning can be achieved 0nly through constant and conscious exposure to the language which is highly important to acquire affixes knowledge."

Literature Review

Schmitt and McArthur (1979) emphasize the importance of what s termed rich instruction in vocabulary teaching. Instructional conditions should be arranged to provide opportunities for a maximum amount of processing of the words. Students should be given chance to manipulate words in varied and rich ways, for example , by describing how they relate to other words and their own familiar experience . To promote and reinforce deep processing , activities should include much discussion of the words and require students to create justification for the relationships and association that they discover.

Nation (2001) asserts that the learners need to know several things to make use of word parts. For receptive use, they have to be able to recognize that a complex word, such as unhappiness , is made of parts, and that these parts can occur in other words, such as unpleasant, happily and sadness. For productive use, the learner needs more detail awareness of the formal changes to the stem and the affix that can occur when they are combined to form complex word. These formal changes can affect

the pronunciation: *flirt / flirtation* (stress change), *quantity / quantify, describe / description They may also affect the written form: legal /illegal. Also, for productive use the learner needs to be aware which form class of stem can take certain affix.*

Before looking at activities to develop each of these kinds of knowledge, Nation(2001)suggests some general principles worth considering. Firstly ,it is probably the most efficient to deal with word parts after learners have already acquired substantial number of complex words as unanalyzed wholes. These can act as familiar items to attach their new knowledge of word parts to. Secondly it is important to see the development of knowledge of word parts as being a long-term process. Thirdly, the use of word parts in understanding and producing words is essentially a creative activities. Fourthly, there are large numbers of stems and affixes but some are much more useful than others.

Considering the instructions the students receive McKeown and Curtis(1987) pointed out that there is good deal of confusion about what affixes are and consequently about teaching them Thus the first thing is needed is full description and consideration of the sorts of elements that might be taught. They suggest a list including inflectional suffixes, derivational suffixes ,prefixes attached to regular English words, prefixes attached to non-English roots, and non English roots them-selves. McKeown and Curtis(1987:175) go on to say

" deciding on grade levels for teaching specific word parts is partly a matter of arbitrarily establishing certain grade levels for certain elements, and partly a matter of deciding when and for what sorts of students certain elements might be useful. For example, instruction in non-English roots may be appropriate only at the high school level.

Nation(2001) recommended a list divided into five stages ( ) of very usefull accessible affixes that learners could be introduced to at appropriate levels of their language development. Stage (1) can be used with low intermediate learners.

Allen(1983) asserts that when dealing with the affix that signal the grammatical function, the correspondence between forms and function, must be learned. For example, having learned the verbs accept , allow, and appear the class learns the related nouns: acceptance, allowance, and appearance. It is helpful to teach a group of (-ance) nouns like those three together, then on another day teach nouns like the (ment) suffix in common .

Rechards (2000) claims that if affixes are regular and behave as would be expected, such as the affix (-dom) always form nouns e.g **freedom / kingdom** then their acquisition should be facilitated. But affixes that are not regular can clearly cause problem. He suggests that, in general, exception should not be introduced until any rule and / or regularity in patterning has been acquired. This implies that the most regular affixes should be taught initially, and only after students are comfortable, with them should the more irregular affixes be focused on.

Ayers (1986) in his appraising of the role of affixes in inferring the meaning of unfamiliar words, he advocates the technique of word analysis. He claims that by developing this technique i.e breaking the word down into base, prefix, and suffix students will acquire a valuable method for gussing the meaning of strange words with the help of the context.

Pitman (2003) claimed that the best method in teaching this aspect of the language is to extract complex words from the text, write them in the board, and tell the students that these words are made up of word parts, and studying these parts increases vocabulary. The teacher breaks the words down into their parts and writes what each part means; defines the prefixes and suffixes of these words; quizzes the students on how to break down the given words into prefixes, roots ,suffixes and what each word part mean. After the quiz, or during the reading time in class, introduce a group of word parts for the following week, and so on. It does not matter what skill the teacher teaches.

Methodology

Data for this study were collected by using a questionnaire for language teachers. The subjects of the questionnaire are Libyan teachers of English of both sexes and all of them teach English in secondary schools.

Hypothesis

The study presupposes that the teachers, generally, have a limited focus on the teaching of prefixes and suffixes, For example, they do not explain in details what affixation is, its types, haw affixes change the grammatical class of the word or haw some prefixes signal negative forms, etc.

Teachers' Questionnaire:

The questionnaire aimed to establish the extent to which teachers focus on affixation in their teaching at secondary schools. It also aimed to determine the teachers' views regarding the student's ability with respect

to the teaching and understanding of affixation; moreover, to obtain information about their viewpoints towards this study.

Preparation of the questionnaire:

The questionnaire consisted of 10 multiple-choice questions ( ). It requires the teachers to select one of the given alternatives. These questions were based on my experience in the teaching field as well as on my background of the English curriculum for secondary school level.

Procedures of the Questionnaire:

The opinions of 20 teachers were elicited through a questionnaire. The results of each item and the percentages will be calculated and introduced in a separate table according to teacher's answers to obtain responses.

Analysis and Evaluation of the Questionnaire Results:

The first question asked whether the teaching of prefixes and suffixes is (a) very important (b) imp0ttant (c) not important, when teaching vocabulary. (see table 1).

Table 1: Answers to Q1 on the importance of using affixation in teaching vocabulary.

0ptions	Very important	%	important	%	Not important	%
N0 of teacher / percentage	13	65%	7	35%	0	0%

The majority of teachers (65) emphasize the importance of using affixation in vocabulary teaching. Table (1) Shows the results of this question.

The second question asked the teachers whether they (a) always (b) usually (c) sometimes (d) never teach affixation when teaching vocabulary. Table (2) shows that the majority of the teachers were not really concerned with teaching of affixation, except for when they came across it by chance as question four show

Table 2: Answers to Q2 on the frequency of teaching affixation when teaching vocabulary

0otions	always	%	usually	*%*	sometimes	%	never	*%*
No of teachers / **percentage**	***3***	***15%***	***3***	***15%***	***4***	***20%***	***10***	***50%***

In question three, the teachers were asked about their opinions on whether the teaching of affixation facilitates vocabulary learning. Half of the teachers (50%) answered "I think so" but non of the teachers answered " I don't think so". The other teachers had different viewpoints, such as 20% of them emphasize facilitation.Table (3) shows the results of this question:

Table 3: Answers to Q3 on whether the teaching of affixation facilitate vocabulary learning.

options	certainly	%	I think so	%	probably	%	I don't think so	%
No of teachers / percentage	4	20%	10	50%	6	30%	0	0%

This table is concerned with the way of how teachers deal with affixation when come across it in their teaching as required in question four, (40%) of them answered that they teach it, while (35%) neglect the teaching of this aspect of the grammar. However (10%) said it depends on the context of the word. Table (4) shows the results of this question

Table 4: Answers to Q 4 on how the teachers deal with affixation, when they come across it in their teaching.

options	Explain it in details	%	Teach it	%	Depends on the word	%	Neglect it	%
No of teachers / percentage	3	15%	8	40%	2	10%	7	35%

Question five asked the teachers how the textbooks deal with affixation. A high percentage of the teachers asserted that the textbooks neglect this aspect of grammar which may not encourage teachers to teach it.

Table (5) shows the answers of this question:

options	Very good	%	good	%	Not bad	%	poor	%
No of teachers / percentage	2	10%	6	30%	2	10%	10	50%

The answers to question six which asked the teachers about the percentage of the affixed words included in texts in their textbooks show that the majority of the teachers 65% confirmed that texts included only 20% of affixed words. Table (6) shows the results of this question:

Table 6: Answers to Q6 on percentage of the affixed words included in texts:

option	20%	%	40%	%	60%	%	80%	%
No of teacher / percentage	13	65%	6	30%	1	5%	0	0%

The seventh question is concerned with the students' ability to recognize the existence of affixation .(65%) of the teacher answered that some of the students could recognize the existence of affixation, and (35%) of them thought it was "few"The table bellow shows the results of this question.

Table 7: Answers to Q7 on the ability of the students to recognize the existence of affixation:

options	All of them	%	Some of them	%	Few of them	%	Non of them	%
No of teacher / percentage	0	0%	13	65%	7	35%	0	0%

In question eight, the teacher were asked if their students encountered any problems with any the productive suffixes. The majority of them (65%) emphasized that their students had faced difficulty with this aspect. Table (8) shows the results of this question:

Table 8:Answers to Q8 on the problems that the students have with any of the common prefixes or suffixes.

options	always	%	usually	%	Some times	%	Never	%
No Of teachers / percentage	3	15%	13	65%	3	15%	1	5%

Question nine asked the teachers whether their students (a) always (b) usually (c) some times (d) never misuse affixes. Most of the teacher (65%) answered "usually" and only one answered " never". Table (9) shows the results of this questions:

Table 9: Answers to Q9 on how often the students misuse affixes.

options	always	%	usually	%	sometimes	%	Never	%
No of teachers / percentage	3	65%	13	65%	3	15%	1	5%

Question ten asked the teachers whether their students (a) always (b) usually (c) some times (d) never ask them about words in reading comprehension built with affixes. Most of the teachers confirmed that students do ask about words in reading comprehension built with affixes.

Table (10) shows the answers of the teachers.

options	always	%	usually	%	sometimes	%	Never	%
No of teachers / percentage	3	15%	10	50%	4	20%	3	15%

Questionnaire Conclusion:

From the analysis of teacher's questionnaire, the following conclusions can be derived:

- Most of the teachers (65%) said that the teaching of prefixes and suffixes is very important.
- Half of the teachers (50%) indicated that they had never taught affixation when teaching vocabulary, where as the other half had different views. This result indicates that a high percentage of the teachers don't focus on affixation
- (50%) of the teachers think that teaching affixation facilitate vocabulary comprehension and (20%) said "certainly" whereas no one answered "I don't think so". This result assured that the

teaching of affixation is a very helpful way of teaching vocabulary.

- (40%) of the teachers said that they teach affixation while (35%) neglect to teach this aspect of grammar.
- Half of the teachers (65%) said that the textbooks deal poorly with affixation.
- Most of the teachers (60%) think that the percentage of the affixed words included in texts in their textbooks is only (20%). That is textbooks have not covered this grammatical aspect adequately.
- Most of the teachers (65%) think that some of their students are able to recognize the existence of affixes, but need teachers' encouragement and assistance.
- A good percentage (65%) of the teachers confirmed that their students have a lot of problem with affixes. This result encourage us to revise the way of teaching this problem and devote time to explain it.
- The majority of the teachers (56%) answered that their students usually misuse affixes when applying them, because they are not clearly taught.
- Finally, the highest percentage of the teachers (65%) said that their students ask about words built with affixes in reading comprehension, because such words sound some how odd to them.

From the above analysis and conclusions, it is clear that the teacher' responses give us a detailed view, on which we can rely. That is teachers generally have a limited focus on the teaching of prefixes and suffixes. For example, they don't explain in detail what affixation is, its types, how affixes change the grammatical class of the word or how some prefixes signal negative forms….etc.It is also clear, from the teachers responses, that the textbooks don't focus on this aspect.

Teaching Techniques: classroom activities:

Students' weakness in understanding affixation is attributed to the fact that teachers have not focused on them out in the text they are using. The teachers' prominent role is to help learners identify affixes

and their types and functions. Teachers should also highlight the strategy of using affixes in vocabulary expansion.

Nation (2001: 278) points out that the word parts strategy for learning new complex words involves two steps:

1- Break the unknown word into parts. This step requires learners to be able to recognize prefixes and suffixes when they occur in words.
2- Relate the meaning of the word new parts to the meaning of the root. This step requires learners to know the meanings of the common word parts.

There are several ways of learning the meaning of prefixes and suffixes and becoming familiar with their forms (Nation:2001). Learners should deliberately learn the meaning of most common affixes. The learning procedure can be the same as the deliberate learning of words using for example word cards. Time should be provided in class to make sure they are learned and simple tests should be given to monitor and encourage learning.

Before dealing with teaching techniques, it is a good idea to refer to the strategy of teaching affixation. To Pitman (2003), this strategy may proceed as follows:

1- Pointing out a complex word from the text.
2- Writing it on the board and asking students to record it in their notebooks.
3- Breaking the word down into its parts and write what each part mean.
4- Defining the prefixes and suffixes of the word.
5- Providing sufficient exercises and activities to enhance teaching.
6- Quizzing the students on how to break down words into prefixes, roots, suffixes and what each word part means.

After some affixes have been learned, there are various activities that can be used to help establish the knowledge. The researcher has chosen some of the affixes which are suitable to practice for this level of learners.

Technique (1)

Nation (2001) suggests, word-making and word –taking activity which involves learners using cards with affixes and stems on them and trying to put them together to make words.

Technique (2)

The other game-like activity, learners can teach each other prefixes and suffixes in pair work. One learner is the teacher and has a list of words with their prefixes and meanings of the prefixes listed. The learner folds his paper, so he can only see the list of meaning. The" teacher" says a word, says its prefix and then waits for the learner to find the meaning. The teacher gives the learner three chances to find the meaning in his list and then gives the answer. The "teacher" then moves on to the next word.

Technique (3)

An interesting affixation game-like activities can be played with intermediate students, is the one devised by (Morgan and Mario:1986)

Put the following affixes on the board:

In-		re-
Un-	-tion	-less
-ness		-er

2- Ask the students to suggest words that contain these elements.
3-Invite the students, individually or in pairs, to invent words of their own containing the elements, and then to use dictionaries to confirm their speculations.

Technique (4)

Forming adjectives from verbs and nouns by using suffixes. Thornbury (2002)

1- Teacher writes these suffixes on the board:

-y , -ic , -ful , -able , -al , -ive

And then write the following nouns and verbs:

Attract , colour , sympathy , romance , crime , hope , imagine , dirt.

2-Invite students in pairs to form adjectives from the nouns and verbs.

Technique (5)

Another technique is suggested by (Gaims and Stuart:1986) to remind students of several different meanings of the suffix (-ish)

1-Denoting nationality e.g Swedish =from Sweden

2-Approximately: with times, numbers e.g fiftyish = about fifty

3-Somewhat, rather tending Towards (informal) e.g. fattish =tending to be fat

greenish = a vague, green Colour.

4-Typical of , in the Manner of. e.g. foolish = like a fool, Typical of a fool.

The teacher puts the following adjectives on the board:

Spanish ,boyish , devilish , childish , snobbish , Irish , tallish , finnish , youngish , sixtyish , blackish.

Then invite the students, individually or in pairs, to divide the adjectives into the four categories according to the meaning of (-ish) in each case.

Nationality	Typical of, in the manner of	Approximately	Some what tending towards

The above mentioned activities and techniques aim to enhance the learners awareness of the notion of affixation competence by being consciously and fully involved in this area of language learning

Recommendation:

Based on the conclusions, the researcher suggests the following recommendation.

Firstly, due to the important role of affixation it should be given a top priority in every language course. The techniques that follow vary in relation to the level of the learners whether it is intermediate or advanced learners.

Secondly, teachers of English should focus on affixation in their due to the importance of this aspect of grammar in learning English as a

foreign language. For intermediate and advanced learners explicit presentation of the aspect of affixation makes students fully aware of the notion they are dealing with. It gives the learners some degree of conscious insight into this phenomenon so as to enlarge their experience of English through the language courses.

Thirdly, a significant implication for teachers is the importance to show students how their acquaintance with a word (enjoy, for example) can lead to learning more new words (like enjoyable, enjoyment, and enjoyably) where endings such as- ment- able -, and- ly signal different grammatical functions for members of word family. Such endings, which are sometimes called derivational suffixes, deserve attention, for two reasons. One is that students' comprehension of English can be greatly strengthened by recognizing familiar elements within words they have not seen before. The other is that their production of English sentences often depends on knowing correspondences between word forms and grammatical function (or parts of speech). Without that knowledge, the learner tends to use the wrong member of a word family – to say, for example, " That was a very enjoy party".

Fourthly, it is worth mentioning, that teachers should be well prepared for their task. Only teachers who have a clear understanding of the notion of affixation will be able to help learners in the best possible way. Thus enhancing the teachers' affixation knowledge and experience play a remarkable role in the process of teaching.

Fifthly, the teacher should also raise his students' power of observation so as to create a sense of affixation awareness which is an essential step to acquire affixation experience.

Sixthly, it is psycholinguistic believed that the more drills learners of English are exposed to, the better their performance will be and the less errors will be committed. It is, therefore, suggested that textbooks authors should take this fact in consideration when designing language courses and textbooks

To conclude, this study is expected to give a positive contribution to the notion of affixation in English. It is hoped that it draws the teacher's attention to the importance of affixation when teaching English as a foreign language. It is also hoped that the study will be of some value to textbook authors when they set out to design their language courses and textbooks.

Moreover, it is hoped that more research be conducted on this aspect to help us find the most possible effective teaching methods.

References

Ayers, Donald. English Words: From Latin and Greek Elements. The University of Arizona Press, 1986.

Gairns, Ruth and Stuart. Working with words: A guide to Teaching and Learning Vocabulary. Cambridge University Press, 1986.

Morgan, John and Mario. Vocabulary Oxford University Press, 1986.

Pittman, W. 2003. Building Vocabulary Through Prefixes, Roots and Suffixes. http:// ite slj. Org/

Thornbury, Scott. How to teach Vocabulary. London: Longman 2002

McArthur, Tom. Using English Prefixes and Suffixes. Collins: London. Glasgow, 1979.

McKeown, Margart and Curtis Mary. The Nature of Vocabulary Acquisition. New Jersey, 1987.

Nation,I.S.P. Learning Vocabulary In Another Language.

Cambridge University Press, 2001.

Rechard, Jack C. Vocabulary in Language Teaching. Cambridge University Press, 2000

مجلة العلوم الشاملة

Journal of Total Science

البحوث المنشورة باللغات الأجنبية

Research Papers in Foreign Languages